

Philipp, Maik

Lesen digital. Komponenten und Prozesse einer sich wandelnden Kompetenz

Weinheim ; Basel : Beltz Juventa 2025, 660 S.



Quellenangabe/ Reference:

Philipp, Maik: Lesen digital. Komponenten und Prozesse einer sich wandelnden Kompetenz. Weinheim ; Basel : Beltz Juventa 2025, 660 S. - URN: urn:nbn:de:0111-pedocs-330935 - DOI: 10.25656/01:33093; 10.3262/978-3-7799-5520-7

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-330935>

<https://doi.org/10.25656/01:33093>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen und das Werk bzw. diesen Inhalt nicht bearbeiten, abwandeln oder in anderer Weise verändern.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-Licence: <http://creativecommons.org/licenses/by-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to alter or transform this work or its contents at all.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft



Maik Philipp

Lesen digital

Komponenten und Prozesse
einer sich wandelnden Kompetenz

BELTZ JUVENTA

Maik Philipp
Lesen digital

Maik Philipp

Lesen digital

Komponenten und Prozesse
einer sich wandelnden Kompetenz

Der Autor

Maik Philipp, Jg. 1979, Dr. phil., ist Professor für Deutschdidaktik an der Pädagogischen Hochschule Zürich. Seine Arbeitsschwerpunkte sind der Erwerb und die (integrative) Förderung von Lese- und Schreibkompetenz, Lese- und Schreibmotivation, Selbstregulation beim Lesen und Schreiben sowie das materialgestützte Schreiben.

Dieses Buch wurde von der Pädagogischen Hochschule finanziell unterstützt. Teile des Buchs sind außerdem im Rahmen des vom Schweizerischen Nationalfonds geförderten Projekts »Writing, Reading, Integrating and Task Effects (WRITE)« (Projektnr. 205180) entstanden.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Der Text dieser Publikation wird unter der Lizenz **Creative Commons Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International (CC BY-NC-ND 4.0)** veröffentlicht. Den vollständigen Lizenztext finden Sie unter: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.de>. Verwertung, die den Rahmen der **CC BY-NC-ND 4.0 Lizenz** überschreitet, ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für die Bearbeitung und Übersetzungen des Werkes. Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Quellenangabe / Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Dieses Buch ist erhältlich als:

ISBN 978-3-7799-6220-5 Print

ISBN 978-3-7799-5520-7 E-Book (PDF)

DOI 10.3262/978-3-7799-5520-7

1. Auflage 2025

© 2025 Beltz Juventa

Verlagsgruppe Beltz

Werderstraße 10, 69469 Weinheim

service@beltz.de

Einige Rechte vorbehalten

Satz: Helmut Rohde, Euskirchen

Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza

Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag

(ID 15985-2104-1001)

Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

I	Lesen digital – eine Einleitung	<u>9</u>
1.1	Was will dieses Buch?	<u>9</u>
1.1.1	Lesen, Leseverstehen, Lesekompetenz	<u>10</u>
1.1.2	Digitales Lesen	<u>12</u>
1.2	Ein Verbund von Prozessen und Fähigkeiten digitalen Lesens – Versuch einer (vor-)strukturierenden Synopse	<u>15</u>
1.2.1	Fünf Modelle und ihre zentralen Komponenten	<u>15</u>
1.2.2	Die zentralen Komponenten digitalen Lesens	<u>17</u>
2	Komponente 1: Aufgabenmodell formieren	<u>21</u>
2.1	Funktionen	<u>22</u>
2.1.1	Kontext- und Aufgabenmodelle: die Rahmung und die Klärung der Leseaktivität	<u>24</u>
2.1.2	Kohärenzstandards: den Aufwand für den Kohärenzaufbau kalibrieren	<u>27</u>
2.1.3	Relevanz: spezifische Verwertungsabsichten an Informationen herantragen	<u>30</u>
2.2	Prozesse und Produkte	<u>32</u>
2.2.1	Kontext- und Aufgabenmodelle	<u>33</u>
2.2.2	Effekte der Aufgabenstellungen	<u>43</u>
2.3	Determinanten und interindividuelle Unterschiede	<u>84</u>
2.3.1	Sprachliche Fähigkeiten	<u>84</u>
2.3.2	Arbeitsgedächtniskapazität	<u>87</u>
2.3.3	Thematisches Vorwissen	<u>90</u>
2.4	Probleme	<u>92</u>
2.4.1	Problemfeld 1: Probleme bei der adäquaten Einschätzung von Informationen hinsichtlich ihrer Eignung in puncto Aufgabenrelevanz	<u>93</u>
2.4.2	Problemfeld 2: inadäquates Vorgehen bei der Aufgabenbearbeitung wegen mangelnder Aufgabenanalyse	<u>95</u>
2.4.3	Abschlussbemerkung	<u>96</u>
2.5	Fazit	<u>97</u>

3	Komponente 2: Informationsquellen evaluieren (Sourcing)	101
3.1	Funktionen	102
3.1.1	Das Beurteilen von Informationen und Informationsquellen – ein erster kursorischer Einstieg	103
3.1.2	Begriffsklärungen	104
3.1.3	Sourcing: mehrdimensionale und vor allem vernetzte Glaubwürdigkeitsevaluationen	118
3.2	Prozesse und Produkte	140
3.2.1	Ein vorgängiger Überblick über Operationalisierungen des Sourcings in Studien	140
3.2.2	Metadaten beachten bzw. benennen	147
3.2.3	Metadaten nutzen	158
3.2.4	Metadaten evaluieren	161
3.3	Determinanten und interindividuelle Unterschiede	199
3.3.1	Gruppen von Personen beim Sourcing (und weiteren Prozessen)	199
3.3.2	Thematisches Vorwissen (Produktperspektiven)	207
3.3.3	Thematische Überzeugungen (Produktperspektiven)	210
3.3.4	Epistemische Überzeugungen (Prozess- und Produktperspektiven)	212
3.3.5	Lesefähigkeiten (Produktperspektiven)	216
3.4	Probleme	219
3.4.1	Problemfeld 1: ungenügende Beurteilungen und Beurteilungsleistungen von Metadaten	219
3.4.2	Problemfeld 2: schwache Heuristiken bei der Beurteilung	222
3.5	Fazit	227
4	Komponente 3: Informationen verstehen und integrieren	233
4.1	Funktionen	234
4.1.1	Leseverstehen – Produktperspektiven	235
4.1.2	Kohärenz aufbauen: allgemein und mit multiplen Repräsentationen und Dokumenten – Prozessperspektiven	241
4.2	Prozesse und Produkte	273
4.2.1	Intertextuelle Integration und Aufbau von Dokumentenmodellen	273
4.2.2	Integration angesichts inkohärenter, widersprüchlicher und fehlerhafter Informationen	301
4.3	Determinanten und interindividuelle Unterschiede	321
4.3.1	Thematisches Vorwissen	321
4.3.2	Thematische Überzeugungen	325
4.3.3	Epistemische Überzeugungen	329
4.3.4	Abschlussbemerkung	358

4.4	Probleme	359
4.4.1	Problemfeld 1: mangelnde Kohärenzleistungen beim Aufbau umfassender (Dokumenten-)Modelle	359
4.4.2	Problemfeld 2: kognitiv ungünstig gestaltete Texte/ Dokumente aus Sicht der verstehenserschwerenden Arbeitsgedächtnisbelastung	364
4.5	Fazit	372
5	Komponente 4: Aufgabenprodukt herstellen	379
5.1	Funktionen	380
5.1.1	Zielbezogene Transformationen von Informationen beim Herstellen von Diskurssynthesen	381
5.1.2	Transformationsprozesse und -logiken beim Herstellen von Diskurssynthesen	384
5.2	Prozesse und Produkte	392
5.2.1	Darstellerische Vorbemerkungen	392
5.2.2	Perspektive 1: Zusammenhänge von Leseprozessen und Textprodukten bei Diskurssynthesen	397
5.2.3	Perspektive 2: Schreibprozesse im Rahmen lesebezogener Aufgaben am Beispiel der Diskurssynthesen	410
5.2.4	Perspektive 3: Zur Rekursivität und zum Zusammenhang von (glückenden) Prozessen der Diskurssynthese	428
5.2.5	Zusammenfassung: ein abstrahiertes idealtypisches Vorgehen beim Schreiben von Diskurssynthesen	438
5.3	Determinanten und interindividuelle Unterschiede	440
5.3.1	Leistungsunterschiede	440
5.3.2	Thematisches Vorwissen	443
5.3.3	Thematische Überzeugungen (Produktperspektive)	449
5.3.4	Lesefähigkeiten (Produktperspektive)	453
5.3.5	Schreibfähigkeiten (Produktperspektive)	456
5.4	Probleme	457
5.4.1	Problemfeld 1: geringe Abdeckung von Inhalten in Diskurssynthesen (Probleme im Auswählen)	457
5.4.2	Problemfeld 2: geringe Integrationsleistungen in Diskurssynthesen (Probleme im Organisieren)	460
5.5	Fazit	466
6	Komponente 5: Selbstregulation	471
6.1	Funktionen	472
6.1.1	Begriffliche und konzeptuelle Schärfungen zum Konstrukt „Selbstregulation“	474
6.1.2	Metakognition und Strategien als zentrale Elemente der Selbstregulation beim Lesen (digitaler Texte)	497
6.1.3	Exekutive Funktionen	504

6.2	Prozesse und Produkte	517
6.2.1	COPES als Beispiel für ein Selbstregulationsmodell mit stark metakognitiver Prägung	519
6.2.2	Empirische Bedeutsamkeit selbstregulatorischer Fähigkeiten für das Lernen und das verstehende Lesen	531
6.2.3	Konstruktiv-responsives Lesen als eine empirisch fundierte Idealbeschreibung selbstregulierten (digitalen) Lesens	542
6.3	Determinanten und interindividuelle Unterschiede	550
6.3.1	Leistungsgruppen in kognitiven Fähigkeiten	550
6.3.2	Epistemische Überzeugungen	558
6.3.3	Thematisches Vorwissen	567
6.4	Probleme	572
6.4.1	Problemfeld 1: missglückende Aufmerksamkeitslenkung (Probleme der Kontrolle)	573
6.4.2	Problemfeld 2: ungenaues Überwachen der eigenen Verstehensleistungen (Probleme des Monitorings)	581
6.4.3	Problemfeld 3: ungenügende spontane Strategieranwendung (Verbund von Strategien auf Meta- und Objektebene)	586
6.5	Fazit	588
	Literaturverzeichnis	595

I Lesen digital – eine Einleitung

1.1 Was will dieses Buch?

Das Thema digitales Lesen hat in der Forschung seit einiger Zeit geraume Aufmerksamkeit auf sich gezogen. Als ein prominentes Beispiel hierfür kann die Metaanalyse von Delgado et al. (2018) gelten. Diese systematische Sekundärauswertung verscrieb sich der Forschungsfrage, ob sich im Leseverstehen Effekte niederschlagen, wenn dieselben inhaltlichen Texte entweder ausgedruckt vorliegen oder am Bildschirm rezipiert werden, wobei dem Umstand, welches Endgerät genutzt wurde, Rechnung getragen wurde. Tatsächlich gab es einen studien-, altersgruppen- und kulturkreisübergreifenden Mediumseffekt zuungunsten der digital präsentierten Texte. Der ermittelte Mediums- bzw. Bildschirmunterlegenheitseffekt ist pointiert formuliert: Wer Texte am Bildschirm las, verstand sie schlechter. Dieses Muster kam dann besonders zum Tragen, wenn expositorische Texte den Lesegegenstand bildeten und wenn die Aufgaben innerhalb einer definierten Zeitvorgabe zu lösen waren. Dass ausgerechnet expositorische Texte vom Mediumseffekt betroffen sind, ist deshalb bemerkenswert, weil sie nachweislich in verschiedenen Leseleistungen – sei es das Memorieren von Inhalten, sei es das Bilden von Inferenzen, seien es globalere Verstehensmaße – Leserinnen und Lesern größere Schwierigkeiten bereiten als narrative Texte (Clinton et al., 2020; Mar et al., 2021). Damit stellt sich die Frage, ob das Lesen digital dargebotener – nicht einmal notwendigerweise die technischen Möglichkeiten der Medien ausnutzender – Texte unerwünschte Nebenwirkungen hat.

Weniger normativ gewendet wirft das metaanalytische Befundmuster von Delgado et al. (2018), zu dem auch andere Metaanalysen gelangten (Clinton et al., 2020; de-la-Peña, Chaves-Yuste & Luque-Rojas, 2024; Kong et al., 2018; Li & Yan, 2024; Salmerón et al., 2024; Schwabe et al., 2022), die Frage auf, ob es in den Prozessen digitalen Lesens – vor allem von Sachtexten – Eigenheiten gibt, die theoretisch erklärt, empirisch erforscht und didaktisch reflektiert werden sollten. Diese Frage bildet den Anlass des vorliegenden Buchs. Dessen Ziel besteht darin, das Lesen in digitalen Medien im Sinne eines internationalen, kognitionspsychologisch verortbaren Forschungsüberblickes zu konturieren und dies – ganz gemäß dem Untertitel – mit Blick auf die Komponenten und zentralen Prozesse sowie die Produkte zu untermauern. Aus diesem weiten und dynamischen Forschungsfeld werden insgesamt fünf Komponenten fokussiert, was impliziert, dass damit nur ein Ausschnitt bearbeitet wird. Dieser Ausschnitt jedoch erscheint klärungsbedürftig, da das Lesen häufig zielbezogen und selbstreguliert erfolgen sollte, um zu gelingen (s. Kap. 2 und 6). Außerdem stellen die epistemisch unsicheren und erst

durch Prüfungen mit Restunsicherheit als glaubwürdig erachteten Dokumente (Kap. 3) die Basis dafür dar, komplexe Kohärenzleistungen über Dokumenten- und Modalitätsgrenzen hinweg zu erbringen (Kap. 4), um auf dieser Basis wiederum eigene (Text-)Produkte zu erstellen (s. Kap. 5). Digitales Lesen ist also ein Problemlösevorgang mit vielen Freiheitsgraden und vielen Herausforderungen, was ihn allein aus kognitiver und metakognitiver Warte heraus zu einem aufwändigen Unterfangen macht, welches den Leserinnen und Lesern durchweg hohe Leistungen abverlangt.

1.1.1 Lesen, Leseverstehen, Lesekompetenz

Wenn es um das Thema Lesen geht, sind diverse Begrifflichkeiten im Umlauf, die sich historisch teils in ihrer Bedeutung stark verändern, zu rivalisierenden Schulen führen und in aller Regel in eine erhebliche Ausdifferenzierung gemündet haben (Alexander & Fox, 2018; Frankel et al., 2017). Aktuell ist dies besonders gut daran zu erkennen, dass das Literacy-Konzept massiv in Erscheinung tritt, das als metatheoretisches Konstrukt (sprachliche) Fähigkeiten vereint und dadurch Gefahr läuft, inhaltlich unbestimmt zu werden. Wenn man sich daher der Domänenspezifik zuwendet, dann gilt es, die Begrifflichkeiten zu definieren und in ihrem Verhältnis zueinander zu klären. Dabei geht es hier nicht um eine umfassende Aufarbeitung, sondern lediglich darum, einige für diesen Band wichtige Termini zügig zu bestimmen. Diese Termini sind von ihrem Verhältnis als konzentrische Kreise zu denken: Lesekompetenz ist das umfassendste Konstrukt, welches das Leseverstehen und das Lesen in sich vereint. Lesen bildet hingegen den Nukleus – auch für das Leseverstehen.

Was also ist zunächst das *Lesen*? Es wird verstanden als die *Fähigkeit, aus schriftsprachlichen Äußerungen im Sinne einer Interaktion von text- und lesen-denseitigen Faktoren Bedeutung zu konstruieren*. Damit ist Lesen ein aktiver Prozess der Bedeutungskonstruktion (Lenhard, 2019). Zusätzlich wird Lesen aber nicht nur als Bedeutungskonstruktion modelliert, sondern zusätzlich als eine zielgerichtete, also intentionale Aktivität, bei der die Auseinandersetzung mit schriftsprachlichen Äußerungen einer Absicht folgt (Alexander et al., 2012; Magliano et al., 2017). Wer liest – so wie Sie jetzt diesen Text – hat eine Verwertungsabsicht, etwa etwas zum Thema digitales Lesen zu erfahren. Lesen ist zudem Kommunikation zwischen produzierenden- und rezipierendenseitigen Personen(-gruppen), bei der die Texte nicht reine Bedeutungsträger sind, sondern einen Entstehungs- und Rezeptionskontext haben, welcher die Bedeutungskonstruktion mitsteuert (Alexander et al., 2012; Frankel et al., 2017; Rouet et al., 2017). Es macht einen Unterschied, ob Sie diesen Text in Ruhe lesen, um sich der Thematik zu widmen, oder ob sie einen Begriff nachschlagen oder sich zunächst orientieren wollen. Dies wiederum führt zu einer weiteren Eigenheit des Lesens:

Es handelt sich um einen Verbund verschiedener beteiligter Prozesse, darunter nicht nur die kognitiven Prozesse der Informationsverarbeitung, sondern auch motivationale und soziale bzw. vom Kontext beeinflusste Prozesse. Das heißt, dass Lesen nicht nur als mehrdimensional angenommen wird und mehr als reine Kognition darstellt, sondern auch Fähigkeiten der motivationalen Energisierung betrifft und eine Kontextsensitivität aufweist (Alexander et al., 2012). Dieses Verständnis des Lesens fasst den Begriff damit ausgesprochen weit und rückt neben den Text und den Kontext auch Merkmale der lesenden Person als wichtige Konstituenten in sein Zentrum (RAND Reading Study Group, 2002).

Leseverstehen ist ein weiterer Begriff, der in engem Verhältnis zum Lesen steht. Leseverstehen bezeichnet Prozesse bzw. Produkte der Extraktion von Bedeutung aus Informationen, die im Fall des Lesens aus schriftlichen Äußerungen stammen (McNamara & Magliano, 2009). Das Produkt Leseverstehen bezeichnet entlang eines Kontinuums verschiedene Arten von Repräsentationen schriftlicher Aussagen, die von der wortwörtlichen Repräsentation über inhaltlich-propositionale Netzwerke hin zu mentalen Modellen reichen. Diese Repräsentationen sind das Resultat verschiedentlich modellierter, im Verbund operierender und interagierender kognitiver Prozesse, wie Leserinnen und Leser Information extrahieren, anreichern, transformieren und zu einer internalen Repräsentation formieren (McNamara & Magliano, 2009).

Lesekompetenz ist ein umfassendes Konstrukt, welches das Leseverstehen beinhaltet (Müller & Richter, 2014). Doch ist das Konstrukt erheblich breiter gefasst, da es darum geht, die Informationen nicht nur zu verstehen, sondern zusätzlich zu beurteilen und anzuwenden. Kognitive Kompetenzen werden allgemein als domänenspezifische Leistungsdispositionen betrachtet, die eine Person relativ zeitstabil besitzt und in konkreten Anwendungssituationen variabel einsetzt, um den konkreten Erfordernissen optimal gerecht zu werden – im Falle des Lesens eben bezogen auf das Verstehen oder die Nutzung schriftsprachbasierter Informationen (Groeben, 2002). Die Kompetenzmodellierung in der Domäne Lesen ist von Kontroversen und Entwicklungen geprägt, die dazu geführt haben, dass es verschiedene Begriffsverständnisse gibt (Alexander & Fox, 2018). Für die Zwecke dieses Bandes soll hier eine Definition herbeigezogen werden, die nicht aus dem Kontext von Schulleistungstudien stammt. Diese weit gefasste Definition nimmt eine deutlich selbstregulatorische Perspektive ein, indem Lesekompetenz als *Fähigkeit* deklariert wird, *die Interaktion lesebezogener Wissensbestände, Überzeugungen, Fähigkeiten und Prozesse sowie eigener Ziele innerhalb gegebener soziokontextueller Bedingungen angemessen zu modulieren und abzustimmen* (Alexander et al., 2012).

Unter diesen Bestandteilen des Konstrukts Lesekompetenz sind die Überzeugungen für die Zwecke dieses Bandes hervorhebenswert. Sie werden explizit von Alexander et al. (2012) auf epistemische Überzeugungen enggeführt, also subjektive Theorien dazu, wie Wissen hergestellt wird und was als Wissen gilt. Solche

Überzeugungen gelten als wichtige Grundlage für strategische Verstehens- und Validierungsprozesse insbesondere bei Aussagen, deren Wahrheitsgehalt unklar ist oder die mit eigenen Wissensbeständen bzw. mit Aussagen aus verschiedenen Dokumenten konfligieren – also vor allem dann, wenn es Kohärenzprobleme gibt (Bråten, Britt et al., 2011; Richter & Maier, 2017). Hierbei haben sich Überzeugungen zum Wissen als generell vorteilhaft erwiesen, die das Wissen als komplex, wandelbar, sich aus mehreren externen Quellen speisend und nach Regeln begründbar konzeptualisieren (Strømsø & Kammerer, 2016). Solche Überzeugungen sind insbesondere dann wichtig, wenn Personen auf digitale Texte mit solchen Informationen stoßen, die schwer einzuschätzen sind und die die Frage aufwerfen, wie belastbar sie sind. Solche Situationen fordern eine Entscheidung ein, ob Leserinnen und Leser plausible Informationen erhalten, die sie bei der Konstruktion des Verstehens bzw. für die aufgabenbezogene Nutzung weiterverwenden sollten. Je weiter fortgeschritten die epistemischen Überzeugungen aus einer Entwicklungsperspektive sind, desto eher sind gelingende Kohärenzleistungen erwartbar (Alexander et al., 2012; Kuhn, 2020).

1.1.2 Digitales Lesen

Ein Buch zum digitalen Lesen ohne eine Klärung dessen, was das Konstrukt bezeichnet, wäre ein grober handwerklicher Fehler. Tatsächlich ist aber „digitales Lesen“ in der wissenschaftlichen Literatur ein auffällig ungeklärter Terminus – teils gleichgesetzt mit dem Lesen am Bildschirm (Clinton, 2019, S. 289). Diesem wenig geklärten Konzept wird häufig über eingrenzende Annäherungen begegnet (s. z. B. Coiro, 2021), wenn dies denn überhaupt erfolgt. In ihrer verdienstvollen Überblicksarbeit haben Singer und Alexander (2017) den Mangel an Geklärtheit besonders eindrucksvoll gezeigt. Sie haben in drei Dutzend kriterienbasiert ausgewählter Studien aus fast zwei Jahrzehnten empirischer Forschung unter anderem die definitorische Klarheit und Explizitheit des sie interessierenden Konstrukts des Leseverstehens in digitalen Kontexten analysiert. Das ernüchternde Ergebnis: In lediglich fünf der Studien gab es Definitionen, nur in zwei Studien war die Begrifflichkeit klar bestimmt. Das bedeutet auch: Fünf Sechstel der Studien beschrieben nicht ausreichend präzise, welche Variable sie maßen.

Solche Ungenauigkeiten gefährden die zweifelsfreie Verständigung, sodass dem Begriff eine Unschärfe bzw. Beliebigkeit anhaftet, die nicht wünschenswert ist. Aus diesem Grund bedarf es einer Begriffsbestimmung. Dieser Band stützt sich auf die Arbeiten zu den „new literacies“ und nutzt eine in der Scientific Community breit etablierte Definition, die das Fähigkeitskonzept der „new literacies“ konturiert und sich explizit als Erweiterung von Lesekompetenz als Teil von „literacy“ allgemein versteht:

„The new literacies of the Internet and other ICTs include the skills, strategies, and dispositions necessary to successfully use and adapt to the rapidly changing information and communication technologies and contexts that continuously emerge in our world and influence all areas of our personal and professional lives. These new literacies allow us to use the Internet and other ICTs to identify important questions, locate information, critically evaluate the usefulness of that information, synthesize information to answer those questions, and then communicate the answers to others.“ (Leu et al., 2004, S. 1572)

Betrachtet man die Definition oben, so fällt auf: Das Wort „Lesen“ bzw. „read“ kommt darin nicht explizit vor. Das ist der Tatsache geschuldet, dass mit den „new literacies“ eine nötige Ergänzung gemeint ist, welche die traditionelle Lese- und Schreibkompetenz innerhalb einer historischen Konfiguration flankiert. Das zeigt sich deutlich im ersten Teil der Definition, welche die Anpassung an das sich stark verändernde Setting der digitalitätsbasierten Kommunikation hervorhebt. Der zweite Teil der Definition geht dann allgemein auf fünf Fähigkeiten ein, die zu Beginn der Definition als Fähigkeiten (Skills), Strategien und Dispositionen deklariert werden. Es handelt sich um a) das Identifizieren wichtiger Fragen, b) das Finden von Informationen, c) die kritische Beurteilung jener gefundenen Informationen, d) die Synthese von Informationen zur Beantwortung der unter a) genannten Fragen sowie e) die Kommunikation des Ergebnisses an andere. Mit dem letztgenannten Punkt öffnen sich die „new literacies“ dezidiert anderen sprachlichen Kompetenzen wie dem Schreiben.

Diese Definition rückt – auch im Einklang mit anderen Definitionen des kompetenten Lesens allgemein (Alexander et al., 2012) – die Fähigkeiten, Strategien und Dispositionen für den erfolgreichen, zielbezogenen Umgang mit digitalen Informationstechnologien in ihr Zentrum und betont die Fluidität einer sich konstant wandelnden, ubiquitären Durchdringung des Lebens mit solchen Technologien. Demnach ist das, was die „new literacies“ bezeichnen, kein starres Gefüge, und es ersetzt auch nicht das bisherige Verständnis von Lesekompetenz, sondern erweitert es, indem es den technologischen Wandel mit seinen Anforderungen als treibende Kraft benennt.

Eine Frage, die sich beim Thema digitales Lesen stellt, ist, ob es sich um eine kategorisch neue Form des Lesens handelt oder um eine Erweiterung bestehender Konzepte des Lesens. Zugespitzt lautet die Frage: Was ist das Genuine, bzw. besser gefragt: das geänderte Anforderungsprofil des digitalen Lesens? Man kann diese Frage auf mindestens zweierlei Arten beantworten:

- *Mediale und gesellschaftliche Veränderungen:* Zum einen gibt es unbestritten technische Veränderungen, die Phänomene wie das Scrollen, das Klicken auf Hyperlinks, das Nutzen von Suchmaschinen, die Einbindung von multimodalen und -medialen Inhalten in kommunikative Angebote oder auch eine Form von Partizipation, Produktion und unmittelbarer Anschlusskommunikation hervorbringen (Leu et al., 2013). Hinzu kommt eine große Menge an Informationen mit teils ungeklärter Herkunft, auf die Personen in digitalen Medien stoßen und mit denen sie sinnvoll verfahren müssen (Alexander et al., 2012; Bråten et al., 2020; Goldman, 2015). Diese Veränderungen betreffen das Medienmenü und die Lesegegenstände im weiten Sinne. Hinzu kommen gesellschaftliche Veränderungen auf der Ebene des Mediensystems mit neuen Anbietern und Distributoren von Inhalten, deren Geschäftsmodelle nach anderen Regeln funktionieren als in der Zeit vor dem Siegeszug des Internets (Kozyreva et al., 2020).
- *Individuelle (kognitive) Faktoren:* Zum anderen – und mit dem technologischen Wandel aufs Engste verquickt – gibt es Prozesse, Praktiken und Strategien, die sich als günstig bzw. schlicht erforderlich erwiesen haben, um beim Lesen und Nutzen digitaler Dokumente erfolgreich zu sein (Cho & Afflerbach, 2017; Cho, Afflerbach & Han, 2018; Magliano et al., 2017; Salmerón, Strømsø et al., 2018; Wylie et al., 2018). Nicht immer glücken diese Prozesse, was sich an den schädlichen Effekten von Des- und Fehlinformationen zeigt und auf die Frage verweist, wie Menschen Informationen verarbeiten und vor allem im Gedächtnis speichern, was digitale Medien teils gezielt ausnutzen bzw. ausnutzbar machen (Kozyreva et al., 2020). Diese Erfordernisse zielen auf die Nutzung von kognitiven Ressourcen in der lesenden Person selbst und zum Teil auf sich verändernde Rollen der Leserinnen und Leser vor allem in puncto der Kohärenzleistungen ab (Philipp, 2023, 2024b).

Diese sich verändernden Anforderungen an das Lesen – eher im Sinne einer Lesemenü-Ergänzung denn als Substitution traditioneller Lesemedien und -anlässe – werfen die Frage danach auf, was Spezifika des digitalen Lesens sind. Technische Veränderungen sind das eine, doch es scheinen die sich ändernden Rollen und Anforderungen an die Leserinnen und Leser zu sein, die besonders wichtig wirken. Denn es sind zunehmend die Leserinnen und Leser selbst, die Informationen gleichsam zu kuratieren haben, um mit digitalen Texten sinnvoll zu verfahren.

1.2 Ein Verbund von Prozessen und Fähigkeiten digitalen Lesens – Versuch einer (vor-)strukturierenden Synopse

Altes vs. neues, analoges vs. digitales Lesen gegeneinander auszuspielen – zumal unter einer kulturoptimistischen oder -pessimistischen Perspektive –, erscheint wenig zielführend, wenn es darum geht, das Phänomen des digitalen Lesens zu fassen und zu verstehen. Dafür gibt es Bemühungen, auch aus theoretischer Warte synoptische Zugänge zu wählen (s. etwa die Darstellung und Zusammenfassung von mehreren Modellen bei Bråten et al., 2020). In der genannten Zusammenschau von Bråten et al. (2020) wird deutlich, was relativ markant und symptomatisch für den gesamten Gegenstandsbereich ist: Eine eigenständige, umfassende (Meta-)Theorie digitalen Lesens liegt derzeit nicht vor – sie tut es schon beim nicht-digitalen Lesen nicht (McNamara & Magliano, 2009). Stattdessen bilden verschiedene Modellierungen eher so etwas wie Mosaiksteine verschiedener Granularität und Geltungsbereiche. Den Gegenstandsbereich leuchten sie also unterschiedlich stark aus und fokussieren auf einzelne Bereiche des Phänomens. Und noch etwas fällt auf: Diese Modelle substituieren nicht bisherige Modelle des Lesens oder verdrängen sie gar. Vielmehr setzen sie diese implizit oder explizit sogar voraus und bauen sie aus. Zudem sind einige dieser Modelle dezidiert hybrid, indem sie einen Geltungsbereich aufweisen, der die Grenze von digital und nicht-digital nicht zieht; sie sind also generisch formuliert. All diese Vorbemerkungen sind deshalb wichtig, weil sie von einer möglicherweise unfruchtbaren Dichotomie von digitalem und nicht-digitalem Lesen wegführen zu den gehaltvolleren Perspektiven auf die Funktionen, Prozesse sowie Produkte, Determinanten und Probleme beim digitalen Lesen – zusätzlich zu all dem, was über das nicht-digitale Lesen bereits bekannt ist.

1.2.1 Fünf Modelle und ihre zentralen Komponenten

Erfolgreiches Lesen – sei es analog, sei es digital, sei es hybrid – erfordert einen Verbund diverser Prozesse und Komponenten, die in verschiedenen Modellen beschrieben wurden (McNamara & Magliano, 2009). Im Falle des digitalen Lesens ist dies in diesem Detaillierungsgrad noch nicht geleistet worden. Allerdings gibt es Bemühungen, die sich bei aller Vorläufigkeit dennoch bündeln lassen. Denn die bislang vorlegten (deskriptiven) Modellierungen und Heuristiken weisen eine recht große Schnittmenge auf. Dies zeigt die ohne Anspruch auf Vollständigkeit auskommende Synopse von fünf solcher Modellierungen in Tabelle 1.

Tabelle 1: Synopse von Fähigkeiten beim digitalen Lesen

Übergeordnete Komponenten	Einzelne Modelle				
	Fähigkeiten „new literacies“	Modell des Informationsproblems im Internet	Modell zum Verstehen multipler Dokumente	Schritte bei MD-TRACE	Kernprozesse funktionaler Nutzung von Text(en)
Aufgabenmodell formieren	Wichtige Fragen identifizieren	Informationsproblem definieren	Aufgabeninterpretation	Aufgabenmodell kreieren und aktualisieren	–
Informationen und Informationsquellen recherchieren	Informationen finden	Informationen suchen	Ressourcen suchen und zusammenstellen	Informationsbedürfnisse beurteilen	Dokumentsuche und Informationen finden
Informationen und Informationsquellen evaluieren	Nützlichkeit der Informationen kritisch beurteilen	Informationen scannen	Sourcing	Relevanz der Informationseinheiten einschätzen	Informationsevaluation mit Sourcing und auf der Basis des inhaltlichen Verstehens
Informationen verstehen und integrieren	Informationen zur Beantwortung der Fragen synthetisieren	Informationen verarbeiten	Quellen analysieren und synthetisieren	• Textinhalte verarbeiten • Dokumentenmodell aufbauen und aktualisieren	• Verstehen von Inhalten einzelner Texte • Informationsintegration über Texte hinweg
Aufgabenprodukt herstellen	Antworten kommunizieren	Informationen organisieren und präsentieren	Anwendung	• Aufgabenprodukt kreieren und aktualisieren • Aufgabenprodukt in puncto Aufgabenziel überprüfen	–
Selbstregulation	–	Regulation	–	–	–

(Darstellung von links nach rechts gehend basierend auf Leu et al., 2004, S. 1572; Walraven et al., 2008, S. 625; Goldman et al., 2010, S. 265; Rouet & Britt, 2011, S. 31–43; Rouet & Potocki, 2018, S. 419; – = in der jeweiligen Definition nicht explizit enthalten)

Die Quellen dieser Synopse sind die oben genannten Fähigkeiten der „new literacies“ (Leu et al., 2004), das Modell des Informationsproblemlösens im Internet (Walraven et al., 2008), das Modell zum Verstehen multipler Dokumente (Goldman et al., 2010), das Modell „Multiple-Document Task-Based Relevance

Assessment and Content Extraction“ (MD-TRACE; Rouet & Britt, 2011) und zentrale Kernprozesse des funktionalen Nutzens von einzelnen bzw. multiplen Texten (Rouet & Potocki, 2018). Nicht alle dieser Modelle sind genuin spezifisch für das Lesen von digitalen Dokumenten erstellt worden. Dennoch ist beispielsweise das MD-TRACE-Modell ein generisches Modell mit breitem theoretischem Geltungsanspruch (Philipp, 2018b), das in Studien sowohl mit Print- als auch mit digitalen Texten zum Einsatz kam und sich dort gleichermaßen als tragfähig erwiesen hat.

Diese Modelle als in der Regel deskriptive Darstellungen eint, dass sie einen Verbund von Schritten oder Prozess(grupp)en postulieren, der entweder taxonomisch hinsichtlich der jeweiligen Hauptfunktionen unterschieden wird oder aber in Prozessmodellen als zyklisch-sequenziell modelliert wird. Diese einzelnen Fähigkeiten werden in den einzelnen Heuristiken bzw. Modellen unterschiedlich differenziert beschrieben. Für den Moment und als erster Überblick sollen hier die Hauptkategorien resp. Komponenten genügen, die in Tabelle 1 zusammengetragen sind. Diese werden entlang übergeordneter Komponenten bezeichnet, welche im gesamten Buch eine strukturierende Funktion insofern haben, als fünf der sechs Komponenten in je einem eigenen Kapitel mit einer stets gleichen Themenentfaltung behandelt werden.

1.2.2 Die zentralen Komponenten digitalen Lesens

Wie im Teilkapitel zuvor anhand der vergleichenden Gegenüberstellung von verschiedenen einschlägigen Modellen demonstriert werden konnte, besteht erfolgreiches digitales Lesen aus einem Verbund von Komponenten. Diese Komponenten bilden ihrerseits ein Konglomerat aus einzelnen Prozessen und (lese-)strategischen Vorgehensweisen, die dabei helfen, die jeweilige übergeordnete Aufgabe in segmentierten Teilaufgaben zu bewältigen. Diese Komponenten sind in Abbildung 1 zusammengefasst und im Uhrzeigersinn radial angeordnet, was der Sequenz von idealtypischen Schritten in den Modellen geschuldet ist (z. B. Britt & Rouet, 2012). Diese Sequenz ist allerdings eher nicht allein dem Faktor Zeit geschuldet, sondern entspringt spezifischen Funktionen von Aktivitäten, die letztlich in einer temporalen Reihenfolge absolviert werden, wobei diese Aktivitäten in empirischen Studien immer wieder in verschiedenen Ausprägungen und individuellen Mischverhältnissen vorkommen und dynamischen Reihenfolgen unterliegen (Du & List, 2020; Goldman et al., 2012; Singer Trakhman et al., 2018; Strømsø et al., 2020). Die Komponente Selbstregulation steht deshalb im Zentrum, weil die Initiierung, Überwachung und Steuerung der strategischen Prozesse der Informationsverarbeitung theoretisch gut begründbar ist und in den einzelnen Modellen partiell auch bereits ex- und implizit begründet enthalten ist (Walraven et al., 2008; Britt & Rouet, 2012).



Abbildung 1: Überblick über wichtige Komponenten des digitalen Lesens mit Hinweisen zu den Kapiteln, in denen sie behandelt werden (eigene Darstellung, basierend auf den Angaben aus Tabelle 1 in Teilkap. [1.2.1](#))

Die Darstellung aus Abbildung 1 ist demnach also kein Modell, sondern hat eine darstellerische Funktion, um die Hauptkomponenten zu versammeln, die aktuell in der Leseforschung als zentral für das digitale Lesen erachtet werden und sich auch empirisch voneinander trennen lassen (Kiili, Leu, Utriainen et al., 2018). Diese Komponenten sind im Einzelnen:

- 1) *Aufgabenmodell formieren* (s. Kap. [2](#)): Leserinnen und Leser verfolgen mit dem Lesen eine Absicht, die je nach Kontext, Dokumenten und Verwertungszweck erheblich variieren kann. Die neuere Forschung versammelt unter dem Konzept des Aufgabenmodells individuelle Repräsentationen vor allem zum Ziel und zur adäquaten Zielerreichung. Diese metakognitive Steuerung der eigenen zielgerichteten Handlungen wird nicht als starr angenommen, sondern als adaptives, über Rückkopplungen anpassbares Regulativ der eigenen Vorgehensweisen, des ursprünglichen Ziels, bereits absolvierter Schritte zur Erreichung des Ziels und etwaiger Modifikationen, aber auch der Einschätzung, das Vorgehen zu beenden.
- 2) *Informationen und Informationsquellen recherchieren*: Ausgehend von einem erkannten Informationsbedürfnis zur Zielerreichung begeben sich Personen auf die Suche nach relevanten Informationen und recherchieren. Das Recherchieren ist ein Vorgang, bei dem potenzielle Informationen bzw. Informationsquellen gesucht und auf ihre Adäquanz zur Zielerreichung überprüft werden, in der Regel gibt es hierbei mehrere Zyklen. Die Funktion des Recherchierens ist damit, Informationsquellen in Datenbanken und Informa-

tionen in Dokumenten zu beschaffen, auf die dann später zurückgegriffen werden kann. Die Recherche wird bei aller Wichtigkeit aus Platzgründen in diesem Buch nicht behandelt; das ihr immanente zentrale Element, nämlich das Beurteilen von Informationen, ist jedoch Gegenstand.

- 3) *Informationen und Informationsquellen evaluieren* (s. Kap. 3): Mit dem Evaluieren ist ein Prozessverbund angesprochen, bei denen die gefundenen Informationen, Informationsquellen und Dokumente nach verschiedenen Kriterien überprüft werden. Dazu zählen Einschätzungen zur Relevanz, also zur Verwertbarkeit für eine Zielerreichung, zur Plausibilität, ergo die vermutliche Wahrhaftigkeit der Information, und zur Glaubwürdigkeit, also zur unterstellten Verlässlichkeit der Informationsquelle und des Dokuments.
- 4) *Informationen verstehen und integrieren* (s. Kap. 4): Hierunter fällt, dass Personen auf der Basis von eigenen Wissensbeständen und den Inhalten aus Texten bzw. Dokumenten Bedeutung konstruieren bzw. die gewünschte Information für die Aufgabenbearbeitung nutzen. Dies erfolgt sowohl hinsichtlich der reinen Inhalte einzelner Texte verschiedener Modalität als auch in der Kombination von (evaluierten) Metadaten sowie über mehrere Dokumente bzw. Texte hinweg.
- 5) *Aufgabenprodukt herstellen* (s. Kap. 5): Digitales Lesen als zielbezogene Aktivität kann, muss aber nicht in jedem Fall in ein Produkt münden, seien es eigene Texte, mündliche Präsentationen oder weitere Anschlusshandlungen, denen das Lesen dient. Das Aufgabenprodukt, welches bereits als Zielvorstellung Teil des Aufgabenmodells ist, determiniert damit die Ausgestaltung des gesamten Prozesses und bildet zudem eine eigenständige Komponente, bei dem das Finden, Evaluieren und Verstehen eine dienende und vorbereitende Funktion haben.
- 6) *Selbstregulation* (s. Kap. 6): Wie schon erwähnt wurde, sind viele der Handlungen, die Personen beim digitalen Lesen absolvieren, zielbezogen und steuerbar, während andere Prozesse automatisiert und ohne aktive Steuerung ablaufen. Dennoch gilt das digitale Lesen als eine stark die exekutiven Funktionen beanspruchende und steuerungsbedürftige Angelegenheit, bei der die Leserinnen und Leser selbst all jene nötigen Prozesse und Handlungen orchestrieren, um sich einem Zielzustand selbstreguliert anzunähern. Jede der zuvor genannten fünf Komponenten trägt diesen Charakter in sich, sodass die Selbstregulationskomponente eine Sonderstellung einnimmt.

Dieser Gang durch die Komponenten fungierte als Advance Organizer für die restlichen Kapitel des Buches, in denen das Konstrukt des digitalen Lesens betrachtet wird. Deutlich geworden sein sollte, dass das Lesen digitaler Texte einen Verbund mehrerer Komponenten erfordert, die die Nutzung in- und externer Ressourcen betreffen und verschiedene Prozesse mit Graden an Steuerbarkeit und gezielter Einflussnahme beinhalten.

2 Komponente 1: Aufgabenmodell formieren

Darum geht es

Der Lesealltag beinhaltet diverse, variable Lesesituationen, in denen Leserinnen und Leser auf verschiedene Texte treffen, die sie – freiwillig oder aus institutionell überformten Gründen – mit divergierenden Zwecken nutzen. Wenn man Lesen als absichtsvolle, ein mehr oder minder klar umrissenes Ziel anstrebende Aktivität konzeptualisiert, so gewinnt ein in der aktuellen Leseforschung stärker fokussiertes Konstrukt an Bedeutung: das *Aufgabenmodell*. Damit ist eine eigene mentale metakognitive Repräsentation gemeint, welche das lesebezogene Ziel beinhaltet und welche Schritte dafür nötig sind, um dieses Ziel zu erreichen. Obwohl damit dem Aufgabenmodell eine hohe Bedeutung zukommt, welche in der aktuellen Theorie immer wieder betont wird, steckt die Erforschung noch in den Kinderschuhen.

Weiter ist man hingegen bei zwei anderen Konstrukten, die enger gefasst sind als umfassende Aufgabenmodelle. Zum einen sind dies *Kohärenzstandards*, die als angestrebter zielbezogener Verstehenserfolg das Vorgehen beim Lesen steuern – je höher der Kohärenzstandard, desto mehr Aufwand betreiben die lesenden Personen mutmaßlich. Zum anderen widmet sich die Leseforschung seit langer Zeit bereits der durch Aufgabenstellungen hervorgerufenen Zuweisung von *Relevanz*, also dem instrumentellen, außerhalb des Texts liegenden Gebrauchswert von Informationen, um eine Aufgabe erfolgreich zu bewältigen (2.1).

Diese Forschung zu Effekten von Aufgaben ist besonders reichhaltig und differenziert, wird daher zum Strukturgeber für einen Forschungsüberblick zu Produkten und Prozessen des relevanzbezogenen Lesens herbeigezogen (2.2) und fungiert überdies im Rest des Kapitels als Bezugspunkt. Der Überblick unterscheidet in aktuell mindestens fünf Arten von Aufgaben, die sich danach unterteilen lassen, auf welche Textmenge und auf welche globalen oder spezifischen Ziele sich die Leseaktivitäten beziehen sollen. Bei den Globalzielen werden a) allgemeine Zwecke wie Unterhaltung bzw. Lernen, b) Lesen mit einer definierten Perspektive und c) ein Lesen für die Bearbeitung einer meist umfassenderen Aufgabe wie Argumentieren oder Zusammenfassen unterschieden. Bei den spezifischen Zielen werden zwei Gruppen unterteilt: solche, die tiefenorientierte Elaborationsfragen zum Gegenstand der Aufgabe machen, oder solche, bei denen es eher darum geht, oberflächenorientierte Fragen zu beantworten, indem einzelne Segmente des Textes im Vordergrund stehen.

Relevanzbezogenes Lesen ist als zuvorderst kognitive Aktivität abhängig davon, mit welchen kognitiven Voraussetzungen eine lesende Person auf die zu lesenden Texte trifft. Dies wird in einem eigenen Teilkapitel (2.3) behandelt. In ihm lassen sich drei Gruppen kognitiver Merkmale als bedeutend für das relevanzbezogene Lesen identifizieren: sprachliche Fähigkeiten (Wortschatz und Leseverstehen), Arbeitsgedächtniskapazität und deklaratives thematisches Vorwissen. Je nach Ausprägung erleichtern oder erschweren diese Merkmale Prozesse und deren Resultate beim relevanzbezogenen Lesen.

Die angesprochene Erschwernis manifestiert sich in den Studien, nach denen sich Problemlagen bei der Genese von Relevanzurteilen in Prozessen und Produkten rekonstruieren lassen (2.3). Zwei Problemfelder bestehen hierbei: Erstens gibt es Defizite in der Aufgabenanalyse und der Passung von Aufgabenerfordernissen beim Lesen und gewähltem Vorgehen, zweitens sind die Produkte der Informationsauswahl beim Lesen unkalibriert und teils inadäquat.

2.1 Funktionen

Lesen wird zunehmend verschiedentlich als zielbezogene Aktivität beschrieben (Alexander et al., 2012; van den Broek & Kendeou, 2022; Magliano et al., 2017; Pressley & Afflerbach, 1995). Dies ist im Kern nicht neu. Bereits im Modell von Kintsch und van Dijk (1978) wurde Textverstehen nicht nur als Re-Konstruktion des Textsinns auf der Basis zyklisch ablaufender text- und lesendenseitiger Informationsverarbeitungsprozesse angenommen, sondern auch als von leserseitigen Zielen gesteuertes Vorgehen. Wer sich in pandemischen Zeiten beispielsweise darüber informiert, ob er oder sie eine Alltagsmaske tragen sollte, wird auf Texte stoßen, die ihr oder ihm Antwort darauf geben. Dabei kann es jemandem genügen, sich nur zu vergewissern, ob das Maskentragen angezeigt ist. Dieselbe Person könnte aber auch verstehen wollen, warum das Maskentragen sinnvoll ist, wer es verordnet hat und wie lang diese Maßnahme gelten wird. Möglicherweise wird sie sich diese Fragen sogar als Leitfragen für die Lektüre überlegt haben. In diesem zweiten Szenario ist zu vermuten, dass die Person mehr Aufwand betreiben wird, indem sie sich beispielsweise mehr Zeit nimmt, Texte langsamer und genauer liest und am Ende eine vollständigere mentale Repräsentation zum Sachverhalt erworben hat. Diese beiden Beispiele illustrieren, dass es intraindividuelle Unterschiede bei Prozessen und Produkten des Lesens gibt, die mit der Verwertungsabsicht beim Lesen korrespondieren.

Wenn Lesen als absichtsvolle Aktivität theoretisch modelliert wird, was heißt das dann für die Prozesse des Lesens und deren Produkte? Eine erste Implikation, die sich daraus ergibt, ist, dass das Verstehen ganzer einzelner Texte nur eine Facette aus dem breiten Spektrum lesebezogener Aktivitäten darstellt. Wie Personen Texte lesen, hängt damit zusammen, wozu sie sie lesen und welche Zwecke sie damit verfolgen. Das bedeutet, dass Lesen einem definierten bzw. definierbaren Zustand dient, den Leserinnen und Leser mit dem Lesen anstreben und den es entsprechend als Konstituenten des Leseverstehens zu berücksichtigen gilt (Britt et al., 2018; Rouet, 2006; Winne & Hadwin, 1998). Es ist demnach erwartbar, dass Prozesse und Produkte des Lesens je nach angestrebtem Ziel divergieren.

Eine zweite, sich daraus direkt ableitende Implikation besteht darin, dass dies theoretisch erklärbar gemacht werden muss, insbesondere wie Ziele das Lesen beeinflussen (McNamara & Magliano, 2009). In der aktuellen Leseforschung wird dies unter verschiedenen Begrifflichkeiten diskutiert, darunter dominant unter „Relevanz“ (s. für einen Überblick: McCrudden & Schraw, 2007, sowie McCrudden et al., 2011a), „Kohärenzstandards“ (Oudega & van den Broek, 2018) und jüngst den „Aufgabenmodellen“ (Rouet et al., 2017, s. auch schon Winne & Hadwin, 1998, und davor bereits Flavell, 1979). Diese Konstrukte haben eine variierende Flughöhe, und sie zielen auf Unterschiedliches ab. Gleichwohl weisen sie eine Schnittmenge und Komplementarität auf. Eine Schnittmenge besteht darin, dass die im- oder explizite, die Leseprozesse dynamisch steuernde und ihrerseits über Rückkopplungsprozesse veränderliche Zielvorstellung zum Zweck der Leseaktivität behandelt und als Regulativ anerkannt wird. Komplementarität besteht darin, worauf die einzelnen Konstrukte beruhen und was sie beschreiben:

- Das „Aufgabenmodell“ (vorgestellt im Teilkapitel [2.1.1](#)) ist im Vergleich das umfassendste Konstrukt, weil es als übergeordnete Steuereinheit für komplexe Aufgabenbearbeitungen fungiert. Demgegenüber sind die beiden anderen Konstrukte erheblich spezifischer.
- Die „Kohärenzstandards“ ([2.1.2](#)) etwa haben ihren Bezugspunkt bereits im konkreten Verstehensprozess, da sie als Entscheidungsgrundlage dafür fungieren, ob Leserinnen und Leser ihr Leseverstehen für ausreichend halten oder ob sie zu einer stärker schlussfolgernden und den Gesamtzusammenhang des Textes forcierenden Variante des Lesens übergehen sollten (van den Broek et al., 1995).
- Die „Relevanz“ ([2.1.3](#)) schließlich bezieht sich auf die Ebene einzelner Informationen und deren instrumenteller Verwertbarkeit für ein definiertes Ziel.

Demnach sind die drei hier versammelten und im Folgenden zu beschreibenden metakognitiven Konstrukte auf verschiedenen Ebenen zu verorten und erfüllen verschiedene Funktionen. Sie liegen damit quer zu allgemeinen Theorien des Leseverstehens bzw. helfen dabei, deren blinde Flecken zu erkennen und nötige Erweiterungen anzustoßen.

2.1.1 Kontext- und Aufgabenmodelle: die Rahmung und die Klärung der Leseaktivität

In jüngerer Zeit sind Tendenzen in der Leseforschung erkennbar, die der subjektiven Verarbeitung und Repräsentation von Zielen des Lesens breiten Raum geben. Diese Tendenzen betten sich ein in allgemeine Strömungen, die Einflussfaktoren für Leistungssituationen besser zu beschreiben und ihr Wirken zu erklären (s. z. B. Broekkamp & van Hout-Wolters, 2007). Diese Theoriewerke unterscheiden sich, was sich besonders deutlich an der Evolution von zwei aufeinander aufbauenden Modellen konturieren lässt. Das erste Modell, das TRACE- und dessen Nachfolger, das MD-TRACE-Modell ((Multiple-Document) Task-Based Relevance Assessment and Content Extraction; Rouet, 2006; Rouet & Britt, 2011), fokussierten auf das Aufgabenmodell. Der Nachfolger mit erheblichen Erweiterungen, das RESOLV-Modell (*Reading as Problem Solving*; Britt, Durik & Rouet, 2022; Britt et al., 2018; Rouet et al., 2017) ergänzt eine weitere Komponente, das Kontextmodell. Hinzu kommen folgenreiche Veränderungen in puncto der Modellierung dynamischer Leseverstehensprozesse, auf die aus Platzgründen an dieser Stelle nicht sinnvoll eingegangen werden kann. Entscheidender ist, dass RESOLV das Kontextmodell ergänzt, weil es dabei hilft, ein präzises Ziel zu formieren, das den Kontext der Leseaktivität systematischer berücksichtigt, als es noch bei (MD-)TRACE der Fall war.

Im Zentrum dessen, was Kontext- und Aufgabenmodell ermöglichen sollen, steht die Zielerreichung. *Ziele* lassen sich allgemein definieren als kognitive Repräsentation eines angestrebten Zustands, welcher den Schlusspunkt eines Problemraums bildet (Britt et al., 2018, S. 202). Das bedeutet für das Lesen, dass Personen eine mehr oder minder geklärte Vorstellung davon haben, wozu das bevorstehende Lesen dienen soll, sei es die Entspannung beim Lesen eines Krimis, sei es das Lernen über Lesetheorien, sei es das Finden einer benötigten Information. Anders gesagt: Leserinnen und Leser nutzen das Lesen für einen bestimmten innerhalb oder außerhalb der unmittelbaren Aktivität des Lesens liegenden Zweck, welcher seinerseits das Lesen steuert.

Welche Ziele die lesende Person verfolgt, betonen sowohl (MD-)TRACE als auch und in erheblich größerem Ausmaß RESOLV als kontextuell überformt. In RESOLV wird deshalb separat jeweils ein Kontext- und ein Aufgabenmodell als Teil von internen Ressourcen modelliert; beide Typen der Modelle haben in ihrem Verbund eine jeweils komplementäre Funktion (s. für eine Kontrastierung die Tabelle 2).

Tabelle 2: Überblick über Hauptunterschiede zwischen Kontext- und Aufgabenmodell

Vergleichsdimension	Kontextmodell	Aufgabenmodell
Inhalt	• jegliche verfügbaren Informationen über den physischen und sozialen Lesekontext: Ort, Zeit, Personen, externe Ressourcen etc. • globale Aufgabenstellungen und -anforderungen	• interpretierte Aufgabenstellungen und -anforderungen • Übersetzung von Aufgabenstellungen und -anforderungen in Ziele und Pläne (inklusive Strategien), die das Leseverhalten steuern • lokale Unterziele und Strategien
Informationsquellen, welche für die Bildung des Modells nötig sind	• primäre Quelle: Repräsentation der Aufgabe selbst • sekundäre Quellen: weitere Elemente aus dem physischen bzw. sozialen Kontext	• Interpretation des Kontextmodells (insbesondere der Aufgabenstellung innerhalb einer Situation) • relevantes Schemawissen
Zeitliche Stabilität	• relativ stabil • aktualisiert aufgrund von Sackgasen und Hindernissen	• hochdynamisch • routinierte Erneuerung/Anpassung während des Fortschritts in der Aufgabenbearbeitung
Kognitives Abrufen und Verarbeiten	• beinhaltet Extraktion und Verarbeiten der situationsbezogenen Informationen • Verwendung von Mustererkennung, um Kontextschemata zu aktivieren	• Interpretation von explizit und implizit gegebenen Aufgabenstellungen • Verwendung von aktivierten Aufgabenschemata und interpretierten Standards/Kriterien

(Quelle der Darstellung gemäß: Britt et al., 2018, S. 53, 201 und 207, und Rouet et al., 2017, S. 208 – übernommen aus Philipp, 2018a, S. 84)

Unter dem *Kontextmodell* versteht man die individuellen mentalen Repräsentationen einer Aufgabe innerhalb ihres Kontextes, wobei diese Repräsentation die Grundlage für das Aufgabenmodell bildet. Das Kontextmodell speist sich aus verschiedenen Quellen, darunter der konkreten Aufgabe, vor allem aber überdauernden Erfahrungen mit Aufgabentypen. Darunter fallen die namensgebenden Kontextinformationen wie zu Institution, Person, möglichen Hindernissen und förderlichen Faktoren sowie internale Merkmale wie Selbsteinschätzungen. Es handelt sich also um interpretierte, teils als Schema im Langzeitgedächtnis gespeicherte Wahrnehmungen des Kontexts und dessen für die Aufgabenbearbeitung nötigen Merkmale. Diese Interpretationen des Kontextes werden nicht immer vollumfänglich bewusst und gesteuert im Sinne vollständiger Analysen durchgeführt, sondern sie können auch unvollständig und unscharf sein (Britt et al., 2018; Rouet et al., 2017). Wer beispielsweise gelernt hat, worauf eine Inhaltsangabe eines gelesenen erzählerischen Textes in der Schule ausgerichtet ist und diese Art des Umgangs mit Texten erfolgreich angewendet hat, wird mit dieser Vorstellung ggf. in der Hochschule andere Erfahrungen machen, wenn es darum geht, wissenschaftliche Fachliteratur aufbereitend zu verarbeiten. Allein die mangelnde

Vertrautheit mit den Anforderungen und den institutionellen Usancen dürfte dazu beitragen, dass dieses Kontextmodell neu zu bilden ist – auch und gerade, weil die Anforderungen an die lesende Person implizit sind.

Das *Aufgabenmodell* ist die spezifisch auf eine einzelne Aufgabe bezogene Repräsentation und beinhaltet die vom Leser bzw. der Leserin durch Interpretationen und Schlussfolgerungen angereicherten Vorstellungen sowohl zum Ziel einer Aufgabe als auch zu den dafür nötigen Handlungsschritten (Britt et al., 2018) – zum Beispiel, dass es günstig ist für das schriftliche Zusammenfassen, wenn man den Text mehrfach liest, das Wichtigste anstreicht, sich ggf. Notizen macht und das Gelesene in eigenen Worten wiedergibt.

Damit stellt das Aufgabenmodell eine metakognitive Steuerungseinheit dar, welche aber nicht rigide das Vorgehen steuert, sondern über Rückkopplungen im Laufe der Aufgabenbearbeitung veränderbar ist (Rouet et al., 2017). Kernstück des Aufgabenmodells sind drei Bestandteile: a) das Ziel des Lesens, b) die damit verbundenen Handlungspläne (verstanden als Unterteilungen der Aufgaben in Teilaufgaben und adäquater Strategien für deren Bearbeitungen im Sinne einer Ziel-Mittel-Analyse) und schließlich als motivationale Komponente c) Wertkognitionen, die für den Grad der Mobilisierung entscheidend sind (Britt et al., 2018). Die Handlungspläne erscheinen deshalb hervorhebenswert, weil sie – basierend auf den kontextuell extrahierten und vom Kontextmodell überformten Interpretationen – auch vorwissensbasierte Elemente und Anreicherungen bis hin zu veritablen Modifikationen enthalten können. Dabei sind die Wissensbestände zu Strategien und die Fähigkeiten zur Selbstregulation aufs Engste mit dem Aufgabenmodell verbunden, denn sie dienen dazu, sowohl vor der Aufgabenbearbeitung als auch während der Aufgabenbearbeitung eine zielführende, adäquate Vorgehensweise aufgrund des eigenen Lern- und Lesestrategiepertoires zu wählen und aufrechtzuerhalten (Rouet et al., 2017).

Wie kommt es zu den genannten Rückkopplungen, und wie funktioniert das Aufgabenmodell im Rahmen der Aufgabenbearbeitung? Zwei Arten von Entscheidungen sind für das Aufgabenmodell bedeutsam, und diese beiden Entscheidungsarten sind komplementär (Britt et al., 2018). Die erste Art hängt mit der inhaltlichen Einschätzung der Informationen zusammen (Routineentscheidungen), die zweite mit der Modifikationsnotwendigkeit des Aufgabenmodells als Ganzes (Nicht-Routineentscheidungen). Die *Routineentscheidungen* betreffen drei informationsbezogene Fragen, die für das Verhalten beim Lesen als metakognitive Entscheidungshilfen fungieren: 1) Werden externe Informationen benötigt? 2) Ist die aktuelle Information relevant für das Ziel (s. dazu Teilkap. [2.1.3](#))? 3) Wurde das Leseziel erreicht? Die *Nicht-Routineentscheidungen* kommen dann zum Tragen, wenn die lesende Person Probleme dabei hat, die lesebezogenen Ziele der Routineentscheidungen zu erreichen. Gründe dafür können unklare Ziele sein, ins Leere laufende Handlungen oder wenn eine Entscheidung zwischen Handlungsoptionen nötig ist, und diese Optionen können sich sowohl

auf die konkreten Vorgehensweisen beziehen als auch auf die Ziele selbst. Damit unterscheiden sich die Routineentscheidungen von ihrem Pendant, den Nicht-Routineentscheidungen, dahingehend, dass erstgenannte auf der inhaltlichen Objektebene des Lesens zu verorten sind, während letztgenannte sich eher auf der Metaebene der Regulation des eigenen Vorgehens lokalisieren lassen (Nelson & Narens, 1990; de Bruin et al., 2020). Mit diesen Arten von Entscheidungen, die sich aus den lesendenseitigen, metakognitiven Einschätzungen von Handlungen ergeben, können die Leserinnen und Leser ihr Handeln steuern, wobei die Grundlage dafür ein hinreichend bestimmtes Leseziel ist.

Resümierend lässt sich also sagen: Das Aufgabenmodell, welches in einer Reihe verwandter, aufeinander aufbauender lesetheoretischer Modelle eine prominente Funktion zugewiesen bekommen hat, dient als dynamische metakognitive Steuereinheit dazu, Ziele beim Lesen zu klären und passende Strategien der Zielerreichung zu wählen. Das Aufgabenmodell wird in neueren theoretischen Modellen als eine aufgabenbezogene Repräsentation vorgeschlagen, welche von übergeordneten Kontextrepräsentationen (Kontextmodell) präformiert ist. Zur Wahrheit in puncto Aufgabenmodell gehört auch, dass es hierbei zu erheblichen Überlappungen zu der Aufgabenanalyse als Teil des selbstregulierten Lernens kommt (Karoly, 1993). Deshalb ist das Aufgabenmodell kein trennscharfes Konstrukt, sondern bildet vielmehr einen Bestandteil verschiedener Modelle, in denen zielgerichtetes und metakognitiv gesteuertes Handeln zum Zweck des Lernens Gegenstand ist (Panadero, 2017). Es ist allerdings – wie das Ergebnis von Verstehensprozessen bei Texten – eine eigene Form der mentalen Repräsentation, nur dass seinen Bezugspunkt nicht der gelesene Text bzw. das gelesene Dokument darstellt, sondern die Aufgabenstellung.

2.1.2 Kohärenzstandards: den Aufwand für den Kohärenzaufbau kalibrieren

Kohärenzstandards sind ein kognitionspsychologisch grundiertes Konstrukt, das allgemein gesprochen den lesezielbezogenen angestrebten Verstehenserfolg bezeichnet und hierbei die primär lesestrategisch erfolgenden Schlussfolgerungen (Inferenzen) betrifft (Oudega & van den Broek, 2018). Van den Broek et al. (2011, S. 124f.) definieren Kohärenzstandards als *die von der lesenden Person angestrebte Art und Stärke von Kohärenz, die sie beim Lesen aufrechterhalten will*. Kohärenzstandards bestehen aus einem Set allgemeiner im- und expliziter Kriterien, die Leserinnen und Leser auf einen spezifischen Leseanlass für einen angestrebten Verstehenserfolg adaptieren – das Leseverstehen ist damit die Zielgröße. Solche Kohärenzstandards variieren interindividuell und zwischen

Situationen, sie beeinflussen das Zusammenspiel von automatischen und strategischen Leseprozessen und werden ihrerseits von Faktoren wie Eigenschaften der lesenden Person, der zu lesenden Texte und der Aufgabe beeinflusst.

Mit der namensgebenden Kohärenz zielen Kohärenzstandards auf die aktive Herstellung von Sinnzusammenhängen, wobei diese Konstruktion von logischen Sinnzusammenhängen im Zentrum des Leseverstehens steht (Richter & Schnotz, 2018). Dabei spielen sowohl textseitige Informationen als auch Vorwissensbestände in ihrer Interaktion eine entscheidende Rolle, ob und wie Verstehensprozesse glücken. Als Herzstück hierbei gelten die Inferenzen, also Schlussfolgerungen, die textuelle und vorwissensbasierte Informationen verbinden und dabei helfen, Beziehungen zwischen diesen Informationen zu klären (van den Broek et al., 2011; Kendeou et al., 2016). Die Leseverstehensforschung illustriert, dass das Inferieren als Prozess zum einen minimalistisch und eher automatisiert erfolgen kann, indem eine lesende Person auf eher lokaler Ebene Informationen im Arbeitsgedächtnis aktiviert hält. Dies ist beim schnellen Lesen, bei geringem Hintergrundwissen zum Textthema bzw. global inkohärenten Texten der Fall. Demgegenüber postuliert – zum anderen – eine konstruktionistische Perspektive, dass Personen mit hohem Vorwissen, zum Zweck des globalen Verständnisses und mit ausreichend Zeit eher dazu neigen, mehr und vor allem auf globale Kohärenzbildung abzielende Schlussfolgerungen zu ziehen (Richter & Schnotz, 2018). Damit verbunden ist die Annahme, dass die Qualität und die Quantität der Inferenzen mit dem (tiefen) Leseverstehen in positivem Verhältnis stehen (s. Teilkap. [4.1.1.2](#)).

Die Kohärenzstandards beeinflussen, ob Personen sich mit einem minimalistischen Verstehen zufrieden geben oder mehr Aufwand im Sinne eines kognitiven Engagements betreiben, um ein Leseverstehen im Sinne der konstruktionistischen Perspektive zu erzielen. Wenn die Kohärenzstandards hoch sind, ist demnach zu vermuten, dass die lesende Person bereit ist, mehr Energie zu investieren, und lesestrategische Vorgehensweisen wählt, die dabei helfen, das Verhältnis von Informationen so zu klären, dass das Leseverstehen dem erwünschten Zustand entspricht. Dies führt bereits über zu einigen Merkmalen von Kohärenzstandards, die van den Broek et al. (2011) benennen und die auf den allgemeinen Nenner gebracht mit dem dynamischen Charakter der Standards zu tun haben:

- 1) *Strategische Prozesse werden dann aktiviert, wenn die automatischen Prozesse die Kohärenzstandards nicht bedienen.* Dieses Charakteristikum impliziert, dass die automatischen Prozesse des Lesens eine Art Grundeinstellung bilden. Erst wenn sie versagen und nicht mehr ein Mindestmaß an Verstehen ermöglichen, kommen die strategischen, kognitiv kostenintensiveren Prozesse zum Einsatz. Hierunter fällt auch, dass Standards nicht immer Maximalstandards sind, welche ein hohes Maß an vorwissensbasierter Strategieanwendung

erfordern und nach sich ziehen, sondern sich der Leseerfolg auch bei niederschweligen Lesezielen einstellen kann, etwa wenn man eine Narration zur Ablenkung liest statt für eine philologische Analyse.

- 2) *Kohärenzstandards können sowohl im- als auch explizit sein.* Nicht alle Lesesaktivitäten setzen voraus, sich über seine Leseziele vollumfänglich bewusst zu sein und sie aktiv zu kalibrieren und festzusetzen – dass sie also explizit sind. Das bedeutet, dass sie implizit wirken können, bis sie verletzt werden, erst dann werden sie von der lesenden Person gewissermaßen in ihrer Störung erkannt und aktiv bemüht, um lesestrategische Aktivitäten zu initiieren. Dieser Grundgedanke ist in neueren Theorien zum Leseverstehen aufgegriffen worden, indem dort ein „Kohärenzschwellenwert“ angenommen wird, der fortlaufend als Eichungsmaßstab für das gegenwärtige Leseverstehen fungiert (O'Brien & Cook, 2016a). Damit wäre ein Kohärenzstandard nicht nur variabel im Grad der Expliztheit, sondern auch beeinflussbar und nicht-statisch in puncto zeitlicher Variabilität. Beide Phänomene eint, dass sich Kohärenzstandards als nicht-fixierte Zielvorgaben erweisen. So kann ein Text beispielsweise zunächst uninteressant wirken, was einer motivierten Anwendung von Strategien entgegensteht, aber im Verlauf des Lesens eine motivationale Stimulanz entfalten, die dazu führt, dass Personen langsamer lesen und mehr verstehen wollen.
- 3) *Kohärenzstandards betreffen verschiedene Ausprägungen und Arten von angestrebter Kohärenz.* In der Leseverstehensforschung werden Inferenzen verschiedentlich klassifiziert (Graesser et al., 1994; van den Broek et al., 2015), um damit sowohl verschiedenen Funktionen – etwa die logische oder zeitliche Kausalität schlussfolgern – gerecht zu werden als auch schwierigkeitsgenerierende Merkmale zu konturieren, etwa wenn sehr implizite Aussagen auf der Basis weit verstreuter Informationen erschlossen werden müssen. Damit haben Kohärenzstandards gemäß Oudega und van den Broek (2018) sowohl eine quantitative als auch eine qualitative Facette. Quantitativ ist demnach, wie viele Inferenzen gebildet werden, um zu einer kohärenten Repräsentation zu gelangen. Qualitativ ist hingegen, dass es verschiedene Arten von Inferenzen gibt, die man bilden kann, um etwa räumliche, temporale oder kausale Relationen zu schlussfolgern. Ein Beispiel für qualitative Unterschiede ist, wenn man bei einem Rezept die zeitliche Reihenfolge versteht, aber nicht darauf abzielt, die kausale Begründung zu verstehen. Es ist demnach für den Erfolg beim Zubereiten einer Speise wichtiger, als lesende und das Gelesene anwendende Person zu verstehen, wann man etwas macht statt warum man etwas macht.
- 4) *Kohärenzstandards speisen sich aus einem Zusammenspiel von lesender Person, Text und Kontext.* Die mehrfach angesprochene Variabilität der Kohärenzstandards ist ein Produkt mehrerer Faktoren, die dynamisch interagieren. So kann ein und derselbe Text vorliegen und ein und dieselbe Person ihn

rezipieren, aber er kann bei unterschiedlichen Aufträgen (ihn zum Vergnügen oder zum Lernen lesen) unterschiedlich gelesen, verarbeitet und verstanden werden. Dieses dynamische Zusammenspiel wird an anderem Ort – nämlich im Teilkapitel [2.3](#) – vertiefend behandelt, wenn die Determinanten des Aufgabenmodells behandelt werden, zu denen Kohärenzstandards als Teil von Lesezielen gezählt werden können.

Kohärenzstandards sind damit ein dynamisches Element des Leseverstehens, welches erklärt, in welchem Maß Personen strategisches Inferieren betreiben, um variable Ziele zu erreichen, die text-, kontext- und personensensibel sind. Damit helfen die Kohärenzstandards, den aus Sicht der lesenden Person angemessenen Aufwand beim Lesen metakognitiv zu kalibrieren.

2.1.3 Relevanz: spezifische Verwertungsabsichten an Informationen herantragen

Wenn Personen nach Informationen in Texten suchen, dann tun sie es oftmals mit der Absicht, mit dieser Information ein Problem zu lösen, eine offene Frage zu klären etc. – mithin Dinge, für die das genaue Verstehen ganzer Texte gar nicht effektiv sein muss bzw. welche die kommunikative Absicht der Texte ignorieren. Anders gesagt: Die vollständige Lektüre mitsamt ihrer lokalen und globalen Kohärenzentfaltung muss nicht zielführend sein für das, wofür die lesende Person den Text in Gebrauch nimmt und wofür sie Relevanz zuweist. Relevanz wird verstanden als *instrumenteller Wert, den eine Information hat, um ein von der lesenden Person verfolgtes Ziel damit zu erreichen* (McCrudden et al., 2011a, S. 2). Entscheidend an dieser Definition ist, dass Relevanz sich aus der Passung von gesuchter Information und übergeordnetem, von außen an den Text herangetragenem Ziel ergibt, d. h. Relevanz ist ein Zusammenspiel aus textseitig angebotener Information und ihrer nicht-textseitig gesteuerten Verwertbarkeit und Nützlichkeit.

Die Bedeutung des Terminus „Relevanz“ lässt sich demnach gut herausarbeiten, wenn man sie derjenigen der „Wichtigkeit“ gegenüberstellt. Wichtigkeit bezieht sich auf die propositionale Position von Informationen innerhalb einer sprachlich kodierten Hierarchie von Aussagen innerhalb eines Textes. Wichtigkeit betrifft damit die Unverzichtbarkeit von Inhalten für ein Gesamtverständnis von Texten, ist also ein Merkmal der textimmanenten Makrostruktur und zu bildender Kohärenz. Die Zuweisung von Unverzichtbarkeit von „Relevanz“ basiert auf textexternen Merkmalen, die – übersetzt in mit dem Lesen verbundene Ziele – als Beurteilungsgrundlage fungieren (McCrudden, 2018; McCrudden & Schraw, 2007). Entsprechend können Informationen mit dem Status Details – also aus der Sicht der Wichtigkeit vernachlässigbare Informationen – wie eine einzelne

Nebenwirkung in einem Medikamentenbeipackzettel aus der Warte der Relevanz für eine Person mit einem gewissen Risiko für diese Nebenwirkung hochrelevant sein. Das leitet über zu der Frage, woran sich Relevanz bemisst.

Relevanz mit ihrer eingeschätzten Nützlichkeit von Informationen ergibt sich gemäß dem zielfokussierten Modell der Relevanz (McCrudden et al., 2011a) als eine *Interaktion von externalen, kontextuellen Hinweisen* wie Arbeitsaufträgen oder metakognitiven Prompts auf der einen Seite *und persönlichen Intentionen* auf der anderen Seite, die in Ziele übersetzt werden. Dies ist von eminenter Bedeutung: Ziele resultieren auch aus individuellen Merkmalen und sind nicht alleinig aus Aufgabenstellungen direkt ableitbar und erst recht nicht als rein exekutive Ausführung seitens der lesenden Person gleichsetzbar. Vielmehr beeinflussen Merkmale der lesenden Person, was sie individuell für relevante Informationen hält (was natürlich auch Fehleinschätzungen bedingen kann). Die sich aus Kontext- und individuellen Merkmalen speisende Zielvorstellung von Relevanz schlägt sich in *Relevanzstandards* nieder. Diese beschreiben eine Passung von Informationen zu definierten Kriterien, welche die zu verarbeitende Information erfüllen muss (McCrudden, 2018). Die Kriterien werden in der zentralen Exekutive, genauer: im Arbeitsgedächtnis, verarbeitet und steuern die Aufmerksamkeit der lesenden Person, die die sensorisch aufgenommenen und kognitiv verarbeiteten textuellen Informationen bereithält und diese mit ebenfalls im Arbeitsgedächtnis aktivierten Informationen aus dem Langzeitgedächtnis vergleicht (Kaakinen & Hyönä, 2008, 2011). Auf dieser Basis kommt es zur metakognitiven Einschätzung, ob die aufmerksamkeitsfokussierten Informationen der Zielerreichung dienlich sind. Die Folge ist, dass zum Ziel passende Informationen verfügbar gehalten werden, während irrelevante Informationen unterdrückt werden.

Wie konkrete Relevanzstandards als Folge von übergeordneten Zielen operationalisiert werden, ist wiederum der schon genannten *Interaktion von externalen und internalen Faktoren* geschuldet. Besonders prominent erforscht sind hierbei die Aufgabenstellungen, also kontextuelle Merkmale (McCrudden & Schraw, 2007). Diese Hinweise variieren in ihrer Spezifität (etwa: nach definierten Zielinformationen suchen oder aber komplexere Fragen zum Text beantworten) und nach Art des Ziels (wie: einen Text aus einer bestimmten Perspektive zu lesen oder in Hinblick auf eine benannte Folgeaufgabe). Sie geben also vor, was, wie und warum gelesen wird (McCrudden, 2018). Hinzu kommen weitere textseitige Merkmale. McCrudden (2018) hat beispielsweise herausgearbeitet, dass Relevanz sich im Falle des Lesens multipler Dokumente auch darüber ergibt, ob die Dokumente hinsichtlich ihrer Glaubwürdigkeit belastbar sind. Die individuellen Merkmale sind hierbei noch nicht einmal berücksichtigt, doch zeigt sich daran bereits, dass Relevanz ein vielschichtiges Konstrukt ist, welches sich aus der Interaktion von verschiedenen Faktoren speist.

Relevanz erfüllt aus Sicht der Verarbeitung von Informationen *vier wichtige Funktionen*, die für das Lesen digitaler Texte im Besondern bzw. für das Lesen im

Allgemeinen Geltung haben (McCrudden & Schraw, 2007): Erstens ist Relevanz dafür hilfreich, die Aufmerksamkeit auf spezifische Inhalte zu richten, sie hat damit eine Signalfunktion. Zweitens hilft Relevanz dabei, das eigene Vorwissen aktiv mit den zu fokussierenden Inhalten zu verbinden, um diese neuen Informationen in eine sinnvolle, dem Vorwissen entsprechende Struktur zu überführen, welche wiederum dabei hilft, sich relevante Inhalte später besser zu merken. Drittens ist zu vermuten, dass Relevanz spezifische Verarbeitungsprozesse initiiert und beeinflusst, indem die Personen relevante und irrelevante Informationen auf unterschiedliche Arten mit unterschiedlichen Kohärenzstandards prozessieren. Viertens fungiert Relevanz unterstützend dabei, eigene Kriterien für die Auswahl und Verarbeitung von Informationen zu nutzen, was teils textimmanenten Gewichungen zuwiderläuft.

2.2 Prozesse und Produkte

Die Forschung zu den Aufgabenmodellen – das ist aus den Ausführungen aus dem Teilkapitel [2.1](#) deutlich geworden – speist sich aus verschiedenen theoretischen Strömungen, die eine gewisse Nähe aufweisen, aber nicht deckungsgleich sind. Zudem haben sie eine unterschiedlich lange Historie der Erforschung, was in ungleichen Forschungsständen resultiert, da beispielsweise Kohärenzstandards und relevanzbasiertes Lesen in der Vergangenheit mehr Aufmerksamkeit erfahren haben als die Aufgabenmodelle. In diesem Teilkapitel besteht entsprechend die Aufgabe darin, diese divergierenden Forschungsstände abzubilden und zu bündeln, ohne die Differenzen zu nivellieren. Dabei ist eine erste Herausforderung für eine Forschungssynopse, dass Kohärenzstandards als Konstrukt vor allem innerhalb der Forschung zur Relevanz beim Lesen untersucht wurden und hierbei als Begründung für differenzielle (Aufgaben-)Effekte bemüht werden. Das bedeutet, dass eine genuin eigene Forschungslinie zu Kohärenzstandards nicht abgebildet werden kann, sondern sich die Forschungsbereiche zu Kohärenzstandards und Aufgabeneffekten so stark überlappen, dass sie kaum sinnvoll separat dargestellt werden können.

Die Implikation für dieses Teilkapitel besteht dementsprechend in einer zweigeteilten Ergebnisdarstellung. Der erste Teil ([2.2.1](#)) widmet sich den Kontext- und Aufgabenmodellen, wobei diese ein Desiderat sind. Deutlich umfangreicher wird hingegen die Präsentation empirischer Ergebnisse zu Aufgabeneffekten, die in der Forschung zur Relevanz beim Lesen einen prominenten Gegenstand bilden ([2.2.2](#)).

2.2.1 Kontext- und Aufgabenmodelle

Obwohl den Aufgabenmodellen in theoretischen Modellen wie (MD-)TRACE (Rouet, 2006; Rouet & Britt, 2011) oder RESOLV (Britt et al., 2018; Rouet et al., 2017) eine hohe Bedeutung zugestanden wird, sind sie bislang eher sporadisch untersucht worden (s. aber schon Flower, 1990). Das Gleiche gilt für die theoretische Ausarbeitung, die bisher eher holzschnittartig erfolgt ist und sich teils auf indirekte Hinweise in relevanzbezogenen Experimentalstudien stützt. Es gibt also derzeit ein Ungleichgewicht zwischen theoretisch hoch veranschlagter Bedeutsamkeit des Aufgabenmodells und dessen faktischer direkter Erforschung.

Dabei ist das Aufgabenmodell durchaus anschlussfähig an jene Forschung zum *konditionalen Lesestrategiewissen*, also jenem Wissen, unter welchen Bedingungen welche Strategien zur Zielerreichung dienlich sind (van Kraayenoord, 2010). Deren Ergebnisse zeigen beispielsweise, dass das Wissen über angemessene Lesestrategien die Lesefähigkeit im Längsschnitt positiv vorhersagt, wobei auch die Gegenwirkung belegt ist – es gibt also ein reziprokes Verstärkungsverhältnis (Artelt et al., 2012; Miyamoto et al., 2019; Muijselaar et al., 2017). Auch in querschnittlichen Untersuchungen wie in verschiedenen PISA-Zyklen mit dem Schwerpunkt Lesekompetenz war das konditionale Strategiewissen ein positiver Prädiktor der demonstrierten Leseleistung (Artelt et al., 2002; Artelt & Schneider, 2015). Die Gemeinsamkeit von konditionalem Strategiewissen und dem Aufgabenmodell besteht in der kontext- und zielsensitiven Einschätzung des Nutzens von Strategien. Im Falle der Erfassung des konditionalen Strategiewissens sind die Strategien als Optionen bereits vorgegeben und werden über das Erkennen von besser geeigneten Varianten gegenüber weniger nützlich wirkenden gezählt. Beim Aufgabenmodell als individueller Vorstellung müssen Personen aus dem Repertoire der eigenen und tatsächlich verfügbaren Strategien die im Kontext bestmöglichen selektieren und mit der variablen, aktuellen Anforderungssituation in Einklang bringen.

2.2.1.1 Mutmaßliche Prozesse

Das Aufgabenmodell, so wie es dessen aktuellste theoretische Modellierung gemäß RESOLV beschreibt, basiert auf dem Kontextmodell (Rouet et al., 2017). Dabei wird angenommen, dass das Kontextmodell die rahmende Basis des Aufgabenmodells ist und dieses wiederum die Leseprozesse und -ergebnisse steuert, gleichwohl gibt es über dynamische Rückkopplungsschleifen auch Modifikationen bei Aufgaben- und Kontextmodell.

Bei dem Aufbau eines *Kontextmodells* sind zwei Prozesse mutmaßlich von hoher Bedeutung. Erstens handelt es sich um die *Extraktion von aufgabenbezogenen und kontextuellen Merkmalen*, darunter die Art der Aufgabe (z. B. eine einzelne Information finden oder einen längeren Text auf der Basis von gefundenen

Informationen schreiben), die Quelle der Aufgabe (z. B. wer sie stellt: eine befreundete Person, eine Lehrperson etc.), das Publikum bzw. die Adressatinnen und Adressaten des Aufgabenprodukts, mögliche Hindernisse und günstige Faktoren (z. B. zeitliche Ressourcen, mutmaßliche Verfügbarkeit von Informationsquellen) und zu guter Letzt auch Kognitionen über die eigene Person, darunter hinsichtlich der eigenen Motivation (Britt et al., 2018). Im Gesamt sind das nicht nur viele, sondern auch sehr heterogene Informationen und Interpretationen. Deshalb wird angenommen, dass Kontextmodelle nicht immer vollständig alle Informationen – zumal bewusst zugänglich – enthalten. Hinzu kommt auf dieser Basis ein zweiter Prozess: die *Aktivierung von Schemata*, also von abstrahierten Mustern von (implizitem) Wissen, die Personen als Ressourcen dafür verwenden, die aktuellen Situationsmerkmale wiederzuerkennen und für eine sinnvolle Interpretation von Aufgabenkontexten zu verwenden, indem dadurch angemessene Strategien der Aufgabenbearbeitung aktiviert werden können (Rouet et al., 2017).

Für das *Aufgabenmodell* werden folgende kognitive und metakognitive Prozesse als zentral erachtet (Rouet et al., 2017):

- aktive *Nutzung von jeglichen Hinweisen zur Aufgabenstellung*, darunter auch solche, die eher als zeitstabile Merkmale den Kontext betreffen, vor allem aber jene, die in der Aufgabenstellung selbst explizit gegeben werden;
- *Interpretation der Aufgabe*, um eine Zielvorstellung davon zu erlangen, welcher Zielzustand erreicht werden soll;
- *Zielformulierung und -aktualisierung*, d. h. das Erkennen von Informationsbedürfnissen und Auswahl von Ressourcen zur Befriedigung der Informationsbedürfnisse, aber auch der im Leseprozess auftretenden Hindernisse, hierbei helfen in Interaktion mit den eigentlichen Leseprozessen die im Teilkapitel [2.1.1](#) skizzierten Routine- und Nicht-Routineentscheidungen dabei, die Zielformulierungen zu regulieren.

Mit Flower (1990) und Britt et al. (2018) sind noch *Kosten-Nutzen-Erwägungen* und die *eigenen Vorwissensbestände* als interne Einflussfaktoren anzuführen, die darüber entscheiden, welche Ziele aktuell erstrebenswert und erreichbar sind. Die darauf basierenden Prozesse sind allerdings noch kaum als gesamthafte Modifikatoren von Aufgabenmodellen erforscht worden. Hierin liegt also ein wichtiges, noch recht unbestelltes Feld der Leseforschung. Der Forschungslücke ist im Folgenden geschuldet, dass die Darstellung zu empirischen Ergebnissen notgedrungen ausgesprochen selektiv ausfällt.

2.2.1.2 Produkte I: Repräsentationen der Kontexte und Anforderungen des Lesens

Obwohl die Leseforschung bereits sehr früh konstatiert hat, dass Relevanz und Leseprozesse interagieren (van Dijk, 1979), ist erstaunlich wenig untersucht worden, wie der Lesekontext (jenseits der im Vergleich relativ stark untersuchten Aufgabenstellungen, s. Teilkap. [2.2.2](#)) das Lesen rahmt und vor allem beeinflusst. Für das Kontext- und das Aufgabenmodell als mentale Repräsentationen von kontextsensitiven Aufgabenzielen und adäquaten Vorgehensweisen zur (Teil-) Zielerreichung wird immer wieder in der Fachliteratur die Studie von Lorch et al. (1993) angeführt. Diese Studie ist im Kasten „Illustrierendes Beispiel 1“ dargestellt und bezieht sich vor allem auf das Kontextmodell. In dieser Studie wurde in zwei Experimenten geprüft, welche Lesestoffe mit welchem Zweck assoziiert sind und welche Eigenheiten einzelne Lesesituationen aufweisen. Wie man dem Erscheinungsdatum der Studie entnehmen kann, war das Lesen im Internet kein Gegenstand, was für eine Replikations- und Auffrischungsstudie spricht.

Illustrierendes Beispiel 1

Lesen für die (Hoch-)Schule ist nicht das Gleiche wie das selbstgewählte – eine Typologie von prototypischen Lesesituationen

A) Was war das Erkenntnisinteresse?

Die Studie interessierte sich dafür, welche prototypischen Lesesituationen im Alltag von Studierenden vorkommen und durch welche Merkmale sich diese auszeichnen (Lorch et al., 1993). Im Zentrum stand die Erstellung einer Typologie von überindividuellen Merkmalen in alltäglichen Lesesituationen.

B) Wer wurde untersucht?

Die Studie hatte verschiedene US-amerikanische Samples, die in verschiedenen Phasen der Studie befragt wurden. In Phase 1 des ersten Experiments waren dies 58 Psychologiestudierende, in Phase 2 waren es 195 Studierende. Das zweite Experiment fand mit 124 Studierenden statt.

C) Welche Instrumente kamen zum Einsatz?

Experiment 1: In Phase 1 absolvierten die Testpersonen eine offene Aufgabe: Sie sollten die typischerweise innerhalb einer Woche gelesenen Texte differenziert aufschreiben und jedem Textmaterial einen Zweck zuordnen. Dabei gab es keine Limitierung der Zwecke zu einzelnen Materialien. Auf dieser Basis wurden insgesamt 100 Lesesituationen extrahiert. Diese 100 auf Kärtchen notierten Lesesituationen wurden von 195 weiteren Studierenden dann in Phase 2 in zwei Teilgruppen sortiert: Sie sollten darüber nachdenken, wie sie in

jeder Lesesituation vorgehen und was sie beim Lesen tun. Einander ähnelnde Situationen sollten sie aufeinanderstapeln. Mittels Clusteranalysen und der Einschätzung von menschlichen Ratings ließen sich zehn Gruppen von prototypischen Situationen extrahieren.

Experiment 2: Im zweiten Experiment beurteilten die Studienteilnehmerinnen und -teilnehmer jede der vorgegebenen einzelnen, im Übrigen durchmischten Lesesituationen mittels 22 Fragen, welche kognitiven und affektiven Umgangsweisen und Merkmale wie etwa das Lesetempo oder die Vergnüglichkeit die jeweilige Lesesituation aufweist. Diese subjektiven Einschätzungen wurden über das Gesamt als Mittelwert gebündelt und sodann für die Gruppen von prototypischen Lesesituationen aus Experiment 1 über Korrelations- und Varianzanalysen ermittelt, ob sich a) Gruppen ähneln und b) die einzelnen Gruppen von prototypischen Lesesituationen Profile bildeten, indem einzelne der 22 Fragen über- oder unterdurchschnittliche Einschätzungen aufwiesen.

D) Was waren die Hauptbefunde?

Das wichtigste Ergebnis ist die Übereinstimmung, dass die Experimente zum Schluss kamen, dass sich bildungsinstitutionelles und freiwilliges Lesen deutlich unterscheiden. Das ist insbesondere im zweiten Experiment eklatant hervorgetreten, da die Mittelwerte in den 22 Fragen sich in 20 Fällen signifikant unterschieden. Diese beiden Hauptkategorien bestehen aus vier bzw. sechs Gruppen von prototypischen Lesesituationen, deren Profil in Tabelle 3 stark kondensiert dargestellt ist. Deutlich wird beim Lesen *in bildungsinstitutionellen Kontexten*, dass die Aussicht auf Überprüfung mit höherem Aufwand verquickt ist und dass die Personen ein mehr oder minder scharf umrissenes (externes) Ziel damit in Zusammenhang bringen. Im Falle des *freiwilligen Lesens* besteht mehr Divergenz, es gibt auch mehr affektive Anteile bei einzelnen Lesesituationen. Alles in allem zeichnen sich damit recht deutliche Unterschiede zwischen den beiden Hauptkategorien ab, wobei das bildungsinstitutionelle Lesen im Vergleich mehr Strategieeinsatz, zeitliche Investition und Genauigkeit erfordert.

E) Diskussion

Die Studie mit ihrem mehrstufigen Vorgehen, bei dem vor allem mit offenen Formaten gearbeitet wurde, aus dem dann im zweiten Experiment geschlossene Aufgaben entstanden sind, verdeutlicht eine Sensibilität bei bildungsnahen Erwachsenen, die als Kollektiv zehn verschiedene Lesesituationen unterscheiden bzw. unterscheidbar wirken lassen. Diese Befunde sind aus der Sicht des Kontextmodells von hohem Belang: Sie verdeutlichen nämlich, dass ein Lesen zum Wissenserwerb (sei es institutionell und dort durchaus mit Leistungsemphase) einem langsamen, kognitiv aufwändigen und lesestrategischem Vorgehen ähnelt, während andere Lesesituationen andere profilgebende Merkmale aufweisen. Personen können also differenzieren, welche Merkmale prototypische Lesesituationen ausmachen. Ob das für das digitale Lesen ähnliche Ergebnisse hätte, müsste dringend untersucht werden.

Tabelle 3: Gegenüberstellung von zwei Hauptkategorien von Leseanlässen mit ihren zehn prototypischen Situationen und deren Merkmalen

	Lesen für bildungsinstitutionelle Zwecke	Freiwilliges Lesen
Typische allgemeine Merkmale	• Lesen ist vorgegeben und dient der Überprüfung zu einem bestimmten Punkt • Lesen erfolgt langsam, ist kognitiv anspruchsvoll, affektiv eher weniger reizvoll	• Lesen erfolgt autonomer und ohne Überprüfung • schnelles Lesen mit affektiver Involvierung, weniger kritisches Lesen, aber auch weniger Ablenkung vom eigentlichen Lesen • insgesamt größere Variabilität als in der bildungsinstitutionellen Gruppe
Prototypische Lesesituationen und ihre Merkmale	• <i>Prüfungsvorbereitung</i>: gut spezifizierter Lesezweck mit unmittelbarer Überprüfung; langsames Lesen, Selbsttestungen, Nutzung von ordnenden Strategien • <i>Lesen für Lernen und schriftliche Produkte</i>: gut spezifizierter Lesezweck mit unmittelbarer Überprüfung; hohe Analyseanteile, wenig Ablenkung, tendenziell affektive Involvierung, tendenziell mehr kritische Analyse • <i>vorgängige Lektüre für Unterricht/Veranstaltungen</i>: globaler Lesezweck ohne unmittelbare Überprüfung; schnelles Lesen mit geringer Fokussierung auf Details • <i>Lesen, um etwas zu lernen</i>: globaler Lesezweck ohne unmittelbare Überprüfung	• <i>Lesen für Anwendungen des Gelesenen</i>: Lesen dient Informationssammlung für eine spezifische Anwendung des Gelesenen; langsames, aufmerksames Lesen mit hohem kognitivem Anspruch • <i>suchendes Lesen</i>: Lesen dient der Informationsfindung bei global klarer Absicht, aber noch ungeklärter Information, die benötigt wird; schnelles Lesen, das wenig auf Verstehen abzielt • <i>Lesen zum Wissenserwerb</i>: Wissensaneignung in eigenem Interessengebiet als Absicht; langsames Lesen mit hohem kognitiven Anforderungsgrad • <i>Lesen zur intellektuellen Herausforderung</i>: Leseabsicht ist auf intellektuelles, nicht auf affektives Involvement ausgelegt und dient der gesellschaftlichen Partizipation durch Rezeption von Fachliteratur und Qualitätsmedien • <i>affektiv stimulierendes Lesen</i>: Leseabsicht besteht in der affektiven Involvierung und Entspannung; Lesen erfolgt mit hoher Konzentration und geringer Ablenkung vom Lesen • <i>leichte Lektüre</i>: Lesen erfolgt zur Entspannung und für positive Affekte; Lesen ist vergnüglich, weniger involviert und erfolgt eher oberflächlich

(eigene, modifizierte und vereinfachte Darstellung, basierend auf Lorch et al., 1993, S. 249)

Selbst wenn der Kenntnisstand in puncto Kontextmodelle bestenfalls als fragmentarisch zu bezeichnen ist und insbesondere aktuelle, sich auf verschiedene Personengruppen beziehende Studien eklatant fehlen (eine Ausnahme bildet die Studie von Schoor et al., 2023), so weist die einzige an dieser Stelle herbeigezogene Untersuchung von Lorch et al. (1993) auf zwei wichtige Aspekte hin:

- Personen differenzieren deutlich, welche Kontexte mit welchen Lesematerialien und welchen Zielen verknüpft sind. Erwartbar ist, dass zwischenzeitlich noch mehr solche kontextsensitiven Ziele und Dokumente hinzugekommen sind.
- Bildungsaffine Leserinnen und Leser können zudem mehrdimensional Anforderungen unterscheiden und typische Eigenheiten des Lesens einschätzen. Dabei wurde deutlich, dass Lesen in bildungsinstitutionellen Kontexten in sich homogener und kognitiv anspruchsvoller wirkt und sich die wahrgenommenen Anforderungen mutmaßlich auch in höheren Kohärenzstandards niederschlagen dürften.

Damit offeriert diese Studie eine Anschlussfähigkeit für die Aufgabenmodelle, wie sie theoretisch postuliert wird: Prototypische Merkmale von Lesesituationen in freizeitlichen oder institutionellen Zusammenhängen legen es nahe, dass die konkreten, aktuellen Lesesituationen überformt mit Kontextschemata wahrgenommen werden und die als typisch erkannten Eigenheiten einen Rahmen für die spezifische Ausgestaltung bilden. In diesem Rahmen sind die Aufgabenmodelle Spezifikationen, die sensitiver für die aktuellen Erfordernisse sind, aber auch individuelle Ziele deutlicher aufgreifen.

2.2.1.3 Produkte II: Aufgabenverständnisse und -modelle

In der empirischen Forschung lassen sich zwei Forschungstraditionen entlang ihrer Erkenntnisinteressen voneinander unterscheiden. Erstens gibt es Studien, die sich verschiedentlich der Frage widmen, welches Aufgabenverständnis Personen haben, also was sie denken, was das Ziel einer Aufgabe ist. Das Aufgabenmodell als zweites, globaleres Konstrukt zielt darauf ab, dass auch die nötigen Handlungen erfasst werden. So fragten beispielsweise Schoor et al. (2021) nach dem Lösen von Aufgaben die Probandinnen und Probanden, wonach die Aufgabe verlangte und welche Schritte man zu absolvieren habe, um die Aufgabe korrekt zu lösen. Greene et al. (2012) stellten ihren Probandinnen und Probanden vor und nach dem Lösen einer Aufgabenstellung die Frage, was jene glaubten, was die Aufgabe verlangte. Die freien Antworten wurden dann nach oberflächen- und tiefenorientierten Merkmalen sowie dem Ressourceneinsatz mehrdimensional kodiert.

Das *Aufgabenverständnis* ist bislang eher weniger als Hauptgegenstand in Studien untersucht worden. Dies fällt unter anderem daran auf, dass die empirischen Belege für das Aufgabenmodell von Rouet et al. (2017) praktisch ausnahmslos aus Studien stammten, die sich dezidiert nicht mit dieser Hauptfrage befassen haben. Vielmehr sind brückenhypothetisch Aussagen formuliert worden, um das Aufgabenmodell zu beschreiben. Überblickt man die Forschung, wird evident, wie symptomatisch es ist, dass thematisch verwandte Fragen in Studien Gegenstand geworden sind. Aus diesen Studien, die das Aufgabenmodell indirekt behandelten

(Ackerman, 1991; Cerdán et al., 2018, 2019; Cerdán, Gilabert & Vidal-Abarca, 2013; Cheong et al., 2024; Flower, 1990; Greene, 1993, 1994; Hadwin et al., 2001; Higgs et al., 2023; Lenski, 1998; List, Oaxaca et al., 2021; Llorens & Cerdán, 2012; Lorch et al., 1993; McCrudden, 2011; McCrudden et al., 2010; Penrose, 1992; Plakans, 2010), lässt sich ein übergeordnetes Muster rekonstruieren. Dieses Muster besteht aus drei markanten Ergebnissen, nämlich dass

- Leserinnen und Leser teils deutlich zwischen Aufgabenanforderungen differenzieren, wobei es Unterschiede zwischen Leistungsgruppen bzw. Personen mit divergierendem Vorwissen gibt,
- (partiell) eine Kongruenz von Prozess- und Produktdaten mit der Repräsentation der Aufgabenerfordernisse besteht und
- dass elaboriertere und korrektere mentale Repräsentationen zu dem, was die Aufgabe erfordert, mit besseren Leistungen korrespondieren.

Die bisherigen Studien, die Aufgabenmodelle direkt erfassten und damit eine große Forschungslücke zu schließen beginnen (Chevrier et al., 2020; Greene et al., 2012; Lee & List, 2023b; List, Du & Wang, 2019; Luyten et al., 2001; Schoor et al., 2021), haben entweder vor einer konkreten Aufgabe relativ abstrakt (Chevrier et al., 2020; Greene et al., 2012; Lee & List, 2023b; List, Du & Wang, 2019; Luyten et al., 2001) oder nach einer Aufgabebearbeitung spezifisch (Schoor et al., 2021) das Aufgabenmodell zu erfassen versucht. Dabei kamen unter bildungsnahen Erwachsenen offene Fragen zum Einsatz, die im Nachhinein inhaltsanalytisch kodiert und in Relation zu Prozess- und/oder Leistungsdaten gesetzt wurden. Diese Studien ergaben, dass sich (ausgeprägte) lineare Zusammenhänge zwischen allgemeinem Aufgabenverständnis und spezifischer Aufgabebearbeitung nicht ohne weiteres nachweisen lassen. Gleichwohl zeigen diese Studien, dass Personen teils differenzierte Vorstellungen von den Aufgabenerfordernissen haben. Dennoch sind die Verbindungen zwischen diesen Vorstellungen und tatsächlichem Verhalten und dessen Resultat anscheinend komplexer. Dies zeigt die Studie, die im Kasten „Illustrierendes Beispiel 2“ präsentiert wird.

Illustrierendes Beispiel 2

Welches Verständnis haben junge Erwachsene von fünf Typen von Aufgaben und hat dies etwas mit ihren Textprodukten zu tun?

A) Was war das Erkenntnisinteresse?

Das Forschungsteam interessierte sich dafür, ob junge Erwachsene bei fünf verschiedenen Aufgabentypen (Zieltextsorten), die in der Sekundarstufe und im Studium typisch sind, sensitiv für die Aufgabenerfordernisse sind. Ein weiteres Erkenntnisinteresse lag darin, dieses Aufgabenverständnis mit Leistungsdaten in Form von drei Produktmerkmalen schriftlicher Argumentationen in Bezug zu setzen (List, Du & Wang, 2019).

B) Wer wurde untersucht?

Die Stichprobe bildeten 132, überwiegend weibliche, mehrheitlich europäische, im Durchschnitt 21 Jahre alte Studierende aus einer US-amerikanischen Universität im mittleren Westen der USA. Die Studierenden studierten im Gros Fächer im bildungswissenschaftlichen Bereich.

C) Welche Instrumente kamen zum Einsatz?

Die Testpersonen erhielten einen Stimulus wie diesen: „Bitte schreiben Sie eine Argumentation zu einem Thema der Entwicklungspsychologie, z. B. Natur vs. Kultur, Piaget, Wygotski.“ Hierbei variierte jeweils die Zieltextsorte (Argumentation, Essay, Zusammenfassung, Meinungsäußerung und Forschungsbericht – also insgesamt fünf, die Reihenfolge wurde randomisiert). Die Studierenden sollten dann so genau wie möglich schriftlich beantworten, was sie denken, was diese Aufgabe von ihnen verlangt. Die im Rahmen der offenen Aufgabe entstandenen Texte wurden inhaltsanalytisch kodiert, es gab knapp zwei Dutzend Kodes. Ausgewertet wurde per Zählung, wie viele Personen in ihren Äußerungen Aussagen der jeweils interessierenden Variable/Kodierung getätigt hatten. Für die neun Kategorien, bei denen sich mindestens zehn Prozent der Personen schriftlich geäußert hatten, wurde ermittelt, ob sich die fünf Aufgabentypen statistisch signifikant unterschieden.

Außerdem wurde regressionsanalytisch geprüft, ob beim Aufgabentypus des schriftlichen Argumentierens die Ausprägungen in den Aufgabenmodell-Kategorien mit zwei im Rahmen der Studie zu schreibenden Argumentationen statistisch zusammenhängen. Grundlage hierfür waren jeweils zwei inhaltlich konfligierende Texte. Die beiden studentischen Argumentationen wurden in drei Merkmalen ausgewertet. Erstens betraf dies die Qualität der Argumentationen, zweitens die Nutzung von Zitaten und drittens die Verwendung von inhaltlichen Belegen aus den beiden Bezugstexten.

D) Was waren die Hauptbefunde?

Insgesamt fiel bei den Nennungen auf, dass mit nur einer Ausnahme (Meinungsbildung bei der Zieltextsorte Meinungstext) die Kategorien nicht mehrheitsfähig waren, und nur in zwei weiteren Fällen wurde ein Anteil von 40 Prozent überschritten. Die Daten zeigen spezifische Profile bei den Aufgabenverständnissen (s. Abbildung 2). Im Fall der *Argumentationen* ist ein abzeichnender Schwerpunkt das Positionieren, beidseitiges Abwägen und Begründen. Bei *Meinungsäußerungen* dominieren das subjektive Positionieren und Äußern der eigenen Meinung. *Essays* sind die am stärksten mit dem Schreiben assoziierte Aufgabenform und überdies jene, bei der das Erklären am häufigsten angeführt wurde. Bei den *Zusammenfassungen* fällt die Prävalenz der paraphrasierenden Wiedergabe jener Hauptinhalte auf, die von den Leserinnen und Lesern wiedergegeben und zuvor gefunden werden müssen. *Forschungsberichte* weisen insofern die Konturen des wissenschaftlichen Schreibens auf, als Informationsrecherche und die adressatengerecht sowie objektiv informierende Informationstransformation vordergründig sind. Hinzu kommen die Nutzung und das Belegen von fremden Quellen.

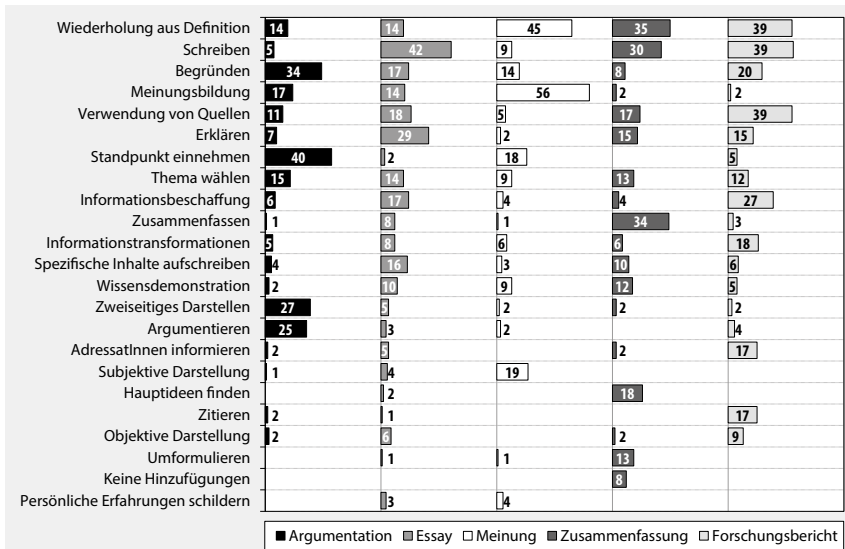


Abbildung 2: Kodierte Merkmale in 23 Kategorien zum Verständnis von fünf Aufgaben (gerundete Zahlen geben Prozente der Personen mit Nennungen an; eigene Darstellung, basierend auf den Daten aus List, Du & Wang, 2019, S. 9)

In Bezug auf die Zusammenhänge von Aufgabenverständnissen zu Argumentationen und den beiden geschriebenen Argumentationen auf der Basis zweier gelesener Texte gab es wenige (und allenfalls schwache statistische Zusammenhänge) bei der Textqualität. In den Regressionsanalysen waren die Kodierungen der Aufgabenverständnisse nicht prädiktiv für die Textqualität. Ähnliches galt für den fehlenden Zusammenhang von der Auffassung, Zitieren gehöre zum Argumentieren, und dem faktischen Zitieren in den beiden eigenen Argumentationen. Auch die Verwendung von Inhalten aus den gelesenen Texten in der eigenen Argumentation ließ sich nicht statistisch auf Teile des Aufgabenverständnisses zurückführen.

E) Diskussion

Die Studierenden benannten im Gesamt der Untersuchungsgruppe einerseits durchaus profilgebende Merkmale von fünf verschiedenen typischen Aufgaben, taten dies allerdings nicht mit großen Mehrheiten, was auch als Indikator von sehr individuellen Vorstellungen zu werten ist. Auffällig ist allerdings, wie stark eigentlich konstitutive Merkmale von Zieltextsorten, etwa das dialogische Abwägen von Argumenten bei Argumentationen oder die explizite Nennung des Zusammenfassens bei geforderten Zusammenfassungen, *nicht* in großer Mehrheit genannt wurden. Als dezidiert heiklen Befund wertet das Forschungsteam den auffälligen Mangel an Nennungen, die auf tiefes Verstehen hinweisen, etwa die Informationstransformation oder die Nutzung verschiedener Dokumente nebst sinnvollen Arten des Belegens.

Neben diesen Auffälligkeiten im deskriptiven Bereich der Befunde ist noch eine weitere anzuführen: Die mangelnde statistische Signifikanz bei der Vorhersage von Produktdaten wirft die Frage danach auf, ob sich das allgemeine Aufgabenmodell – zumal ein recht abstrakt gehaltenes, wie es in der Studie erfragt wurde – als ausreichend kontextsensitiv und valide darstellte bzw. ob die Zusammenhänge zwischen Aufgabenmodell und Textprodukt nicht-linear sind.

Wie auch beim Kontextmodell ist bei dem Aufgabenmodell zu konzedieren, dass es sich gegenwärtig nicht um einen sonderlich stark erforschten Bereich des Lesens handelt. Dabei gibt es durchaus anschlussfähige Forschungszweige wie die Metakognitionsforschung – und hier: jene zum metakognitiven prozeduralen und konditionalen Wissen über Lesestrategien – oder die mit ihr verwandte Selbstregulationsforschung (s. dazu Teilkap. [6.1](#)). Dennoch: Das Ungleichgewicht in der Forschung in Bezug auf Verstehensprozesse und mentale Repräsentationen von Texten im Vergleich zu mentalen Repräsentationen über Aufgabenerfordernisse frappt. Dies erscheint dringend korrekturbedürftig, da Modelle wie jüngst RESOLV (Rouet et al., 2017) darauf aufmerksam machen, dass Leserinnen und Leser bei der Aufgabenbearbeitung parallel mehrere Repräsentationen – zum Kontext, zur Aufgabe und zum lesend zu verarbeitenden Gegenstand – aufzubauen und simultan aufeinander abzustimmen haben.

Die bisherigen empirischen Befunde unterstreichen bei aller Vorläufigkeit, dass Personen differenzieren können, dass Aufgaben unterschiedliche Anforderungen mit sich bringen und dass ihnen dies anscheinend besser gelingt, je kompetenter sie bereits sind. Dies zeigt sich sowohl beim Verständnis der Aufgabe als auch bei umfassenderen Konstrukten wie dem Aufgabenmodell. Dennoch: Der Nachweis von direkten Zusammenhängen mit dem Verhalten und Leistungsdaten ist bislang noch nicht zufriedenstellend erbracht, was auch damit zu tun haben mag, dass lineare Zusammenhänge, wie sie vielen eingesetzten Analysemethoden innewohnen, möglicherweise nicht das Mittel der Wahl darstellen, um den Dynamiken gerecht zu werden, insbesondere der in der Theorie benannten rekursiven Veränderung des Aufgabenmodells durch den fortschreitenden Verlauf in der Aufgabenbearbeitung.

2.2.2 Effekte der Aufgabenstellungen

2.2.2.1 Eine Systematik zu den Relevanzeffekten in Leseaufgaben

Wie im Teilkapitel [2.1.3](#) dargestellt wurde, ist Relevanz eine verwertungsbasierte Zuschreibung von Nützlichkeit einzelner Informationen bzw. ganzer Informationseinheiten, welche der wahrscheinlichen Zielerreichung beim Lesen folgt. Diese Entscheidung treffen lesende Personen in Abwägungen vor dem Hintergrund eines Leseziels, das sich aus externalen und internalen Faktoren ergibt. In der Forschung sind Effekte der Relevanz vor allem und dominant aus der Perspektive untersucht worden, wie sich verschiedene Aufgabenstellungen auf primär produktbezogene Maße auswirken (Relevanzeffekte; McCrudden & Schraw, 2007). Dabei wird angenommen, dass Relevanzhinweise in zielorientiertem Verhalten münden und dieses wiederum zu Lerneffekten führt (sog. „Modell der Zielfokussierung der Relevanz“; McCrudden & Schraw, 2007; McCrudden et al., 2011a).

Die Aufgabenstellungen und ihre Effekte sind unter dem Stichwort der Relevanzinstruktionen seit geraumer Zeit Gegenstand experimenteller Studien. In aller Regel erhalten dabei Personengruppen in experimentellen Settings einen spezifischen Auftrag, welcher gezielt die Aufmerksamkeit auf solche textuellen Informationen richten soll, die der Zielerreichung dienen. Dabei haben die Experimente unterschiedliche Arten von Instruktionen gewählt und diese Stimuli auf ihre Wirkung geprüft. Die Vielzahl der Experimente macht es nötig, sie zu systematisieren.

Ihren Ausgangspunkt nimmt die Darstellung in diesem Teilkapitel in der *Taxonomie von Relevanzhinweisen*, die McCrudden und Schraw (2007) vorgelegt haben. Sie untergliedern darin zwei Gruppen von Relevanzhinweisen in Aufgabenstellungen, nämlich generische und spezifische Hinweise (s. Tabelle 4). Diese werden nochmals unterteilt in allgemeine und perspektivische Aufträge bei den generischen Aufgabenstellungen und auf spezifische Segmente bzw. auf Inferenzbildung abzielende Elaborationsfragen. Damit fokussieren diese vier Arten von Aufgabenstellungen jeweils Spezifisches. Unterschiede in diesen Gruppen bestehen zum einen darin, welche *Informationsmengen* die Leserinnen und Leser berücksichtigen müssen, um die zugehörigen Aufträge erfolgreich zu bearbeiten. Zum anderen sind auch die *Nutzungen der eigenen Vorwissensbestände* für die Präzisierung des Arbeitsauftrages und den Aufbau eines Aufgabenmodells unterschiedlich gefragt. Dies liegt in der Natur der Arbeitsaufträge begründet, die unterschiedlich informationsreich sind und daher insbesondere bei generischen Hinweisen viele Freiheiten lassen, aber auch zu vielen Entscheidungsnotwendigkeiten führen.

Die ursprünglich auf eine Vier-Gruppen-Taxonomie beschränkte Klassifikation erscheint inzwischen erweiterungsbedürftig bei den generischen Hinweisen. Bråten, Gil und Strømsø (2011) machten darauf aufmerksam, dass das erstarkende Forschungsfeld zum Leseverstehen multipler Dokumente und Texte im Allgemeinen (Britt & Rouet, 2020; Primor & Katzir, 2018; Rouet et al., 2019) und zu differenziellen Aufgabeneffekten (Philipp, 2022a) im Besonderen eine neue relevanzbezogene Subkategorie bei den generischen Hinweisen nötig wirken lassen, nämlich zu *Aufgabeneffekten* (s. mittlere Spalte in Tabelle 4). Der Hintergrund ist, dass die Notwendigkeit, intra- und intertextuelle Inferenzen zu bilden (also kohärenzstiftende Verbindungen zwischen einzelnen Informationen vorwissensbasiert herzustellen), nicht nur bei singulären Texten als essenzielle Bedingung gilt (Kendeou et al., 2016), sondern auch und gerade bei den multiplen Texten (Britt & Rouet, 2020; Rouet et al., 2019). Welche Art von Inferenzen gebildet werden, hängt damit zusammen, welche Aufgaben die Personen zu bearbeiten haben, da verschiedene Aufgaben unterschiedliche Kohärenzstandards nahelegen (Rouet et al., 2019) bzw. verschiedene Grade an Transformationsprozessen erfordern (Spivey, 1990). So lässt sich plausibel vermuten, dass beim Auftrag, eine Argumentation zu schreiben, ein höherer Grad an Transformation der Informationen stattfindet als bei Beschreibungen, weil Informationen als gültige oder ungültige Argumente in ihrem Verbund erkannt und schriftlich persuasiv genutzt werden müssen (Philipp, 2021c). Bråten, Gil und Strømsø (2011) subsumieren in ihrer Darstellung die aufgabenbezogenen Effekte unter die allgemeinen Zwecke, und zwar zum Zweck des Lernens, während McCrudden und Schraw (2007) den Kontext deutlich als Kriterium für diese Kategorie hervorheben. Es handelt sich also um eine jeweils unterschiedliche Granularität der Betrachtungsebene – bzw. im Licht von RESOLV als Fokussierung auf Kontext- bzw. Aufgabenmodelle. Darstellerisch wird dies in Tabelle 4 gelöst, indem die Aufgabeneffekte eine eigene Subkategorie bilden, was auch dem Umstand geschuldet ist, dass sich die Aufgabenaufträge als dynamisches Feld erweisen, welches einer genaueren Kartierung zu bedürfen scheint. Zudem wird in letzter Zeit moniert, dass die bisherige Forschung viele realweltliche Leseabsichten in ihrer Verquickung untereinander bislang nicht berücksichtigt (Ramsay & Sperling, 2011), sodass auch deshalb die Taxonomie in Tabelle 4 dezidiert vorläufig ist.

Tabelle 4: Taxonomie von Relevanzhinweisen in Aufgabenstellungen

Haupt-kategorie	Spezifische Hinweise		Generische Hinweise		
Subunter-gliederung (Teilkapitel)	Spezifische Segmente – Fragen nach dem „Was“ (2.2.2.6)	Elaborations-fragen – Fragen nach dem „Warum“ (2.2.2.5)	Aufgaben-aufträge* – Lesen, um definierte Aufgaben mit erhöhter Komplexität und Offenheit zu lösen (2.2.2.4)	Perspektiven-aufträge – Nutzung des Vorwissens zur Einschätzung (2.2.2.3)	Allgemeiner Zweck – Anpassung des Lese-verhaltens an Kontext (2.2.2.2)
Definition	Hinweise zur Fokussierung auf einzelne Segmente	Hinweise, die Inferenzen zur Erklärung erfordern	Hinweise, die Informationen für eine aufgaben-spezifische Transformati-on zu nutzen	Hinweise, Texte aus einer vor-gegebenen Perspektive zu lesen	Hinweise, für einen generel-len Zweck zu lesen
Beispiel	sich bei einer Beschreibung zweier Länder auf eines kon-zentrieren	sich fragen, warum bei einem Land die Postlauf-zeit lang ist	zwei Länder anhand ihrer Merkmale vergleichen	die Vorzüge und Nach-teile eines Landes als potenziellen Lebensmitt-el-punkt lesend erkennen	zum Zweck des Vergnü-gens oder des Lernens lesen

(Quelle: eigene, modifizierte Darstellung, basierend auf McCrudden & Schraw, 2007, S. 116, und McCrudden et al., 2010, S. 231; * in Originaltaxonomie so nicht enthalten)

Zusammengefasst lässt sich sagen: Wesentliche Informationen, die über die Ad-äquanz der Aufgabenbearbeitung entscheiden und in einem Aufgabenmodell repräsentiert sind, basieren auf subjektiven Interpretationen. Die fünf Arten von relevanzbezogenen Aufträgen in Tabelle 4 sind demnach mutmaßlich mit höheren Anforderungen assoziiert, je weiter man sich nach rechts in der Tabelle bewegt. Die fünf Arten von Aufträgen dienen im Folgenden zudem der Strukturierung des Forschungsüberblicks, die entsprechenden Nummern der Teilkapitel sind im Tabellenkopf angegeben. Der Gang der folgenden Darstellung orientiert sich im Folgenden von rechts nach links in der Tabelle 4.

2.2.2.2 Allgemeiner Zweck des Lesens 1: Lesen zu verschiedenen Zwecken (Unterhaltung vs. Lernen)

Wie wir lesen und wie gut wir Texte verstehen, hängt davon ab, *wozu* wir lesen bzw. lesen sollen. Diese Erkenntnis bildet bereits die Essenz jenes Zweigs der Forschung ab, welcher sich der Frage verschrieben hat, ob und welchen Effekt ein allgemeiner Lesezweck hat. Eine besonders häufig in der Forschungsliteratur

anzutreffende Unterscheidung wird danach getroffen, dass *Personen zum Vergnügen oder zu Zwecken des institutionellen Lernens lesen*. Dies spiegelt wider, was bereits beim Kontextmodell als Grobunterscheidung nachgezeichnet werden konnte (s. Teilkap. 2.2.1.2). Hier zeigte sich in der zentralen Studie von Lorch et al. (1993) eine deutliche Tendenz zu einer stärkeren Anstrengungsbereitschaft im Sinne einer tieferen inhaltlichen, lernstrategiebasierten Durchdringung der Texte beim Lesen zum Lernen. Diese Dichotomie von Lesen zum Lernen bzw. zur Unterhaltung steht im Zentrum der folgenden Ausführungen, wobei dies im Wissen darum geschieht, dass damit nur ein kleiner Ausschnitt fokussiert wird.

Produkte

Bei Produktmaßen werden in den experimentellen Studien zwei Gruppen von Variablen betrachtet und primär untersucht: Verstehensleistungen und – in geringerem Ausmaß – metakognitive Einschätzungen dieser Leistungen aus Sicht der lesenden Personen.

1) Verstehensleistungen: Leistungs- und Erinnerungstests

Die Experimentalstudien, die sich Effekten des Lesezwecks widmeten, kommen im Gesamt zu uneinheitlichen Befunden, die sich zu zwei Mustern verdichten lassen: Entweder es gab Vorsprünge in Studien für Gruppen mit dem Lesezweck des Lernens – oder aber es gab keine solchen Differenzen.

- *Muster 1: Vorsprünge in Leistungstests für Personen mit dem Zweck des Lernens.* Wer las, um zu lernen, erreichte bessere Leistungen zum Zweck des Lernens bzw. für Kursergebnisse in Form von besseren Ergebnissen in der Wissensanwendung (Scrimin et al., 2018) und Erinnerungsleistungen an gelesene Textinhalte (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; van den Broek, Lorch et al., 2001).
- *Muster 2: keine differenziellen Effekte.* Generelle Vorsprünge in den Verstehensleistungen sind aber nicht durchgängig belegt. Diverse Studien konnten keine Differenzen feststellen (Latini et al., 2019; Linderholm et al., 2008; Linderholm & Wilde, 2010; Narvaez et al., 1999; Yeari et al., 2015).

Gerade das zweite Muster wirft Fragen danach auf, unter welchen Bedingungen ein Lesezweck mit differenziellen Leistungen einhergeht. Hierbei helfen Mediationsanalysen wie jene von Latini et al. (2019), die zeigen konnten, dass mittels offener Leistungstest dann bessere Leseverstehensleistungen beim Lesen zu beobachten waren, wenn die Texte in den Antworten ausführlich, also in größerer Textmenge ausfielen – auch das kann als Indikator für ein stärkeres kognitives Engagement gewertet werden (Bråten, Brante & Strømsø, 2018). Bemerkens- und

bedenkenswert ist bei alldem: In keiner der Studien war das Lesen zum Zweck der Unterhaltung demjenigen zum Lernen in puncto Verstehensleistungen überlegen. Der Zweck des Lernens ist demnach nicht per se überlegen, sondern eher tendenziell.

2) Metakognitive Selbsteinschätzung der eigenen Leseleistung

Noch wenig in den Blick der Forschung zu Effekten der Aufgabenstellung geraten ist der Effekt auf die Genauigkeit der metakognitiven Selbsteinschätzung der eigenen Leistung („metacomprehension accuracy“, s. Teilkap. 6.4.2). Diese metakognitive Einschätzung gilt als bedeutsam für eine Verstehensüberwachung und als Grund dafür, warum Personen bei einem subjektiv zu gering empfundenen Verstehen des Texts zu selbstregulatorischen Aktivitäten der Verstehenssicherung greifen (Prinz et al., 2020a).

Bezüglich der Aufgabeneffekte sind bisher nur wenige Studien anführbar (Linderholm et al., 2008; Linderholm & Wilde, 2010). In einer der Studien überschätzten sich (erwachsene) Personen in Abhängigkeit von externen Lesezielen (Linderholm & Wilde, 2010). Die Gruppe der Personen mit dem Leseziel des Lernens war mehr davon überzeugt, korrekte Antworten im Lesetest gegeben zu haben, obwohl die zweite Experimentalgruppe mit dem Leseziel Unterhaltung vergleichbare Leistungen erbracht hatte. In der zweiten Studie unterschätzten die Personen in beiden Experimentalgruppen die erbrachte Leistung, aber es zeigte sich, dass die Personen in der Gruppe mit dem Leseauftrag zum Lernen sich etwas kalibrierter eingeschätzt hatten als die zum Zweck der Unterhaltung Lesenden (Linderholm et al., 2008).

Damit ergaben sich Schwierigkeiten in einer akkurat eingeschätzten Testleistung in beiden Studien. In der ersten Studie waren die Personen mit dem Ziel, Texte zum Lernen zu lesen, zu optimistisch, in der zweiten Studie eher zu pessimistisch. Dafür hingen in der zweiten Studie die Leistungen und die Urteile über die Leistungen ein wenig enger zusammen als bei Personen, die zum Vergnügen lesen sollten. Und doch: Es scheint den Probandinnen und Probanden – erwachsenen Personen – schwerzufallen, ein realistisches Urteil zu treffen. Dies ist vor dem Hintergrund der Akkuratheit in puncto Metaverstehen bemerkenswert, weil insbesondere erwachsene Leserinnen und Leser ihre Leseleistung in Bezug auf die tatsächlich erbrachte Testwerte recht realistisch einschätzen können (Prinz et al., 2020a).

Die Prozesse des relevanzbezogenen Lesens zur Unterhaltung bzw. zum Lernen sind in den Untersuchungen verschiedentlich operationalisiert worden. Hier werden drei Gruppen von Prozessdaten fokussiert: erstens Lesezeiten, zweitens Indikatoren, die sich aus Blickbewegungen ableiten lassen, und drittens die per Lautdenk-Protokoll gewonnenen Inferenzbildungen in situ.

1) Lesezeiten

Wenn Lesezeiten als indirekte Prozessindikatoren gemessen werden, so liegt dem eine Grundannahme zugrunde: Längeres Lesen gilt als Hinweis dafür, dass es zu einem vertieften kognitiven Engagement mit dem Inhalt kommt (Miller, 2015). In einigen Studien wurden Lesezeiten erfasst, wobei die Evidenz uneindeutig ist: Personen, die Texte zum Zweck des Lernens lesen sollten, lasen langsamer als jene, die zum Zweck des Unterhaltens lasen (Linderholm et al., 2008; Yeari et al., 2015). Dieser Effekt ist nicht durchgängig beobachtet worden (Latini et al., 2019; Narvaez et al., 1999). Diese Uneindeutigkeit verweist zuallererst darauf, dass Aufgabeneffekte zu globalen Zwecken wie Unterhaltung oder Lernen nicht zwangsläufig zu quantitativen Unterschieden führen, sondern dass es sich lohnt, mit genaueren Analysen und mit genaueren Instrumenten auszuloten, worin konkret die Unterschiede bestehen – wenn es sie denn gibt.

2) Eye-Tracking-Indikatoren

Wie schon bei den Lesezeiten handelt es sich bei Eye-Tracking-Maßen um indirekte Varianten, kognitives Engagement zu erfassen. Man geht dabei davon aus, dass erfasste Fixationen von Inhalten, die über Eye-Tracker gemessen werden, als Hinweis dafür zu werten sind, dass Personen ihre Aufmerksamkeit gezielt auf Inhalte richten und dass längere Verweildauern bzw. wiederholte Betrachtungen als Indikatoren für eine verstärkte kognitive Auseinandersetzung mit dem jeweils aktuell im Blick fokussierten Gegenstand gelten (Miller, 2015). Bezüglich der Erforschung von allgemeinen Lesezwecken wie Unterhaltung oder Lernen gibt es bisher nur wenige Studien. Die im Kasten „Illustrierendes Beispiel 3“ vorgestellte Untersuchung ist damit eine Ausnahme, aber sie ist ausgesprochen instruktiv.

Illustrierendes Beispiel 3

Worauf haben Personen ein Auge, wenn sie zu verschiedenen Zwecken lesen?

A) Was war das Erkenntnisinteresse?

Die Studie hatte zum Ziel, die Blickbewegungen in Abhängigkeit von vier verschiedenen Leseaufträgen zu erfassen und auszuwerten (Yeari et al., 2015). Dabei interessierten diverse prozessbezogene Maße (darunter auch die Lesezeit), aber auch Erinnerungsleistungen. Hierfür wurde bei den Informationen (bzw. Informationseinheiten in der Terminologie der Studie) sowohl bei Prozess- als auch bei den Produktmaßen danach differenziert, ob sie entweder zentral für das Verstehen des jeweiligen Textes waren oder aber peripher.

B) Wer wurde untersucht?

An der Studie nahmen 64 im Durchschnitt 19 Jahre alte Studierende teil, von denen 95 Prozent Frauen waren. Alle diese Personen stammten aus den Niederlanden und sprachen Niederländisch als Muttersprache. Über die Studienfächer der Probandinnen und Probanden gibt es keine weiteren Angaben.

C) Welche Instrumente kamen zum Einsatz?

Jede Person wurde zufällig einer von zwei Gruppen zugeteilt, welche jeweils zwei mal fünf kurze Sachtexte mit einem Leseziel las. Die eine Gruppe las im ersten Block fünf Texte zur Unterhaltung und im zweiten Block fünf weitere Texte mit dem Ziel, nach dem Lesen eine Präsentation vor anderen zum Thema zu halten und auf Fragen Antwort geben zu können. Die zweite Gruppe hatte zwei andere Aufträge: Sie las eine Hälfte der Texte in Vorbereitung auf einen Multiple-Choice-Test und die zweite Hälfte für einen Test mit offenen Fragen. Nach jedem Block fünf gelesener Texte lösten die Studierenden tatsächlich Aufgaben, welche dem Lesezweck entsprachen.

Während die Personen die Texte lasen, trugen sie einen Eye-Tracker, der die Blickbewegungen erfasste. Durch die vorab definierten wichtigen und peripheren Sätze in den Texten war es möglich, diverse Maße bei den *Prozessen* zu erfassen. Unter diesen Maßen waren a) Lesezeiten von Informationseinheiten, b) die erstmalige Lesezeit einer Informationseinheit, c) die Lesezeit mit wiederholtem Lesen einer Informationseinheit, d) durchschnittliche Länge von Fixationen von Informationseinheiten, e) Anzahl aller Fixationen der Informationseinheiten sowie f) Länge der Blicksprünge (Sakkaden) nach rechts innerhalb der Informationseinheiten (je kürzer diese Vorwärtssakkaden sind, desto kognitiv beanspruchter gelten die Personen). Im Rahmen der Analysen wurden vor allem die zentralen Informationseinheiten fokussiert. Als *Produktmaße* wurden Multiple-Choice-Tests verwendet, die pro Text nach einem zentralen und einem peripheren Inhalt fragten.

D) Was waren die Hauptbefunde?

Die Hauptbefunde, die im Folgenden präsentiert werden, beziehen sich nur auf den Vergleich von Personen, die zur Unterhaltung und zur Vorbereitung einer mündlichen Präsentation lasen. Die Ergebnisse betreffen damit nur einen Ausschnitt aus den reichhaltigen

empirischen Daten. Die Resultate sind gemischt. Keine Differenzen existierten bei den *Produktmaßen*, was bedeutet, dass die Personen sich ähnlich gut an die Textinhalte erinnert hatten. Bei den *Prozessmaßen* gab es heterogene Befunde, die sich hier nur auf die zentralen Informationseinheiten als Gegenstand der Analysen von Prozessdaten beziehen. Personen, die die jeweiligen Texte zur Unterhaltung lasen, neigten dazu, schneller zu lesen, und weniger dazu, Inhalte erneut längere Zeit zu lesen, als dies Personen mit dem Ziel zur Präsentation taten. Keine Differenzen bestanden in der Zeit, wie lange Personen erstmals die Informationseinheiten lasen. Einen Unterschied gab es jedoch dahingehend, dass Leserinnen und Leser Informationseinheiten beim Lesen zum Zweck der Präsentation häufiger fixierten, aber nicht zwingend länger.

Insgesamt sprechen die Resultate dafür, dass Leserinnen und Leser dann die zentralen Inhalte häufiger betrachteten (und dadurch auch insgesamt länger lasen), wenn sie den Auftrag hatten, die Inhalte später mündlich zu präsentieren. Das Lesen zum Zweck der Unterhaltung war demgegenüber kürzer und weniger von erneuten Lektüren geprägt. Dennoch manifestierten sich keine Differenzen in den Erinnerungsleistungen in den Multiple-Choice-Fragen, also in den Verstehensprodukten.

E) Diskussion

Eine Besonderheit der Studie liegt darin, dass ein und dieselben Personen im Sinne eines Intra-Personenvergleichs untersucht wurden und dass hierbei die Textblöcke mit den jeweiligen Aufträgen ausgeglichen wurden, was ein Pluspunkt des experimentellen Designs ist. Die Studie zeigt, dass sich die Personen in Prozessmaßen unterschieden, die auf eine häufigere Zuwendung zu zentralen Inhalten schließen lassen, die diese Personen häufiger erneut konsultierten und dadurch auch eine längere Zeit mit dem Lesen dieser Inhalte verbrachten. Damit ist das wichtigste Ergebnis, dass dieselben Personen ihre Leseprozesse resp. ihre Augenbewegungen systematisch variieren bzw. modulieren, auch wenn sich dies nicht in statistisch signifikanten Leistungsdifferenzen niederschlägt. Dies wird von Analysen unterstützt, nach denen die hier berichteten Variablen der Lesezeit und der Produktmaße nicht statistisch zusammenhängen. Sie bestehen also unabhängig voneinander, was die Frage aufwirft, welche Prozesse der Leserinnen und Leser bzw. welche weiteren Variablen dafür verantwortlich sind, wie gut jemand im Lesetest abschnitt.

Die Studie von Yeari et al. (2015) hat illustriert, dass sich in den Blickbewegungen Hinweise darauf finden lassen, dass es zu wiederholten und damit auch längeren Konsultationen zentraler Inhalte kommt, wenn Personen zum Zweck des Lernens Texte rezipieren. Selbst wenn sich diese nicht in besseren Testleistungen niederschlagen, so demonstrieren die quantitativ stärker ausgeprägten Zuwendungen zu zentralen Inhalten indirekt eine stärkere Sensibilität für die Aufgabenerfordernisse.

3) Inferenzbildungen

In den Studien mit lautem Denken, die für die Darstellung der Ergebnisse in diesem Teil bedeutend sind, wurden die Wortäußerungen kodiert und kategorisiert. Dadurch ließ sich feststellen, in welchen parallel zum Lesen erfassten kognitiven und metakognitiven Prozessen sich Auffälligkeiten ergaben. In einigen Studien wurden dabei die *Inferenzen* genauer betrachtet, welche man beim Lesen dazu benötigt, Kohärenz aufzubauen (Kintsch & van Dijk, 1978; s. dazu Teilkap. 4.1.2.1).

Mit dem Bilden von Inferenzen wird ein Prozess des vorwissensbasierten Verknüpfens von Informationen bezeichnet, der dazu dient, im Text fehlende Informationen bzw. implizite Verbindungen zwischen Information zu ergänzen (Kendeou et al., 2016). Diese Inferenzen werden auf unterschiedliche Arten klassifiziert, z. B. danach, ob sie sich auf die Inhalte von Texten beziehen, auf das Vorwissen erstrecken und wie die Inferenzen sich nach zeitlichen Gesichtspunkten als vor, während oder nach dem Lesen von Texten reihen lassen (Kendeou, 2015; McNamara & Magliano, 2009; van den Broek et al., 1993, 2015):

- *Text- bzw. vorwissensbasierte Inferenzen*: In diesem Fall unterscheidet sich, womit textuelle Informationen verbunden werden. Im Falle von textuellen Bezugspunkten, etwa zuvor gelesenen Informationen, ist von „Brückeninferenzen“ die Rede. Hier verbindet die Inferenz einer Brücke gleich zwei Informationen, metaphorisch gesprochen: die Brückenköpfe. Liegen die Informationen außerhalb des Textes, spricht man von „Elaborationsinferenzen“, also Anreicherungen (Elaborationen) des im Text Enthaltenen mit dem Vorwissen der lesenden Person. Als eine Art Mischform können „Erklärungsinferenzen“ gelten, bei denen der Zusammenhang zwischen kurz zuvor gelesenen Informationen über logische Relationen verbunden werden, die im Text nicht expliziert wurden.
- *Zeitlichkeitsbasierte Inferenzen*: Diese Unterscheidungsdimension betrifft eine chronologische Perspektive. So können Inferenzen sich auf noch zu lesende Textteile beziehen (z. B. was man erwartet, worum es im Folgenden gehen kann) und sie können sich auf das bereits Gelesene erstrecken.

Aus der kurzen Liste, welche nur einige Dimensionen der Unterscheidungen und hierbei wiederum nur eine Auswahl enthält, gelten vor allem die vorwissensbasierten Inferenzen als besonders wichtig. Sie helfen dabei, die Relationen zwischen textuell gegebenen Informationen mithilfe der Anwendung eigenen Wissens zu klären. Diese Inferenzen bezeichnen van den Broek et al. (2015) als *Kausalinferenzen*, also als Inferenzen, welche das spezifische Verhältnis von Informationen zueinander klären. Inwieweit solche Kausalinferenzen gebildet werden, hängt gemäß van den Broek et al. (2015) von den Kohärenzstandards in puncto der Kausalität ab, nämlich zum einen, ob es nötig ist, die jeweilige

Kausalinferenz zu bilden, und zum anderen auch davon, ob die gebildete Inferenz genügt, die Verbindung zwischen Informationen ausreichend zu klären. Je nötiger Kausalinferenzen und je schwieriger sie zu bilden sind, desto mehr lesestrategische Vorwissensnutzungen (und/oder ggf. auch eine intensivere Nutzung des Texts) sind angezeigt. Es ist demnach erwartbar, dass die Kohärenzstandards, welche Personen beim Lesen in puncto des relevanzbezogenen Lesens anlegen, zu unterschiedlichen Effekten bei den Inferenzen führen: Je höher die Standards ausfallen, d. h. je kohärenter die angestrebte Repräsentation des Gelesenen zum Zweck der adäquaten Aufgabenerfüllung sein soll, desto mehr auf tiefes Verstehen abzielende, analytische (Kausal-)Inferenzen sind erwartbar (Oudega & van den Broek, 2018).

Ob das empirisch der Fall ist, haben Studien mit variierenden Aufgabenstellungen untersucht, in denen die Testpersonen – ausnahmslos Studierende – vor allem Sachtexte zum Zwecke des Lernens vs. zur Unterhaltung lasen (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; Narvaez et al., 1999; van den Broek, Lorch et al., 2001). Die wichtige Gemeinsamkeit über die Studien hinweg: Der Arbeitsauftrag zum Lernen impliziert einen höheren Kohärenzstandard, der sich wiederum in entsprechend stärker auf die Inferenzprozesse auswirken sollte. Die empirischen Befunde bestätigen dies, wenn man sie bündelt:

- *Aufträge zum Lernen schlugen sich in tendenziell mehr Kausalinferenzen nieder.* Dieses Muster findet sich bei verschiedenen Arten von Inferenzen wieder. Darunter fallen zuvorderst die vorwissens- und textbasierten Inferenzen zu bereits Gelesenem (Linderholm & van den Broek, 2002; van den Broek, Lorch et al., 2001), aber auch die prädictierenden, weitere mögliche Textinhalte antizipierenden Aussagen (van den Broek, Lorch et al., 2001).
- *Aufträge zum Lernen führten zu mehr (paraphrasierenden) Wiederholungen von Informationen, um diese für die kognitive Verarbeitung zur Verfügung zu stellen.* Dies sind keine Kausalinferenzen, sondern Wiederholungen gelesener Inhalte. Das Wiederholen hat den Zweck, die Informationen im Arbeitsgedächtnis verfügbar zu halten und dadurch für spätere Zyklen der Informationsverarbeitung zur Verfügung zu stellen. Dieses vermehrte Wiederholen wurde in mehreren Studien beobachtet (Linderholm & van den Broek, 2002; Narvaez et al., 1999; van den Broek, Lorch et al., 2001). Das Wiederholen erfolgt nicht nur verbatim, sondern mitunter bereits verarbeitet als Paraphrase, also als sinngemäße Wiedergabe von Aussagen in eigenen Worten (Linderholm & van den Broek, 2002; van den Broek, Lorch et al., 2001). Das Wiederholen von Informationen tauchte, wo es spezifisch geprüft wurde, nicht nur bei Sachtexten allgemein auf, sondern auch in der Interaktion mit dem Leseziel, d. h. es trat insbesondere bei Sachtexten in Erscheinung, die zum Zweck des Lernens gelesen werden (Narvaez et al., 1999). Ebenfalls gab

es teils nur in Teilgruppen Differenzen, z. B. bei Personen mit einer hohen Arbeitsgedächtnisspanne (Bohn-Gettler & Kendeou, 2014).

- *Aufgaben mit dem Auftrag, um zu lernen, zogen teils mehr Evaluationen nach sich.* Unter den Evaluationen werden in Studien unterschiedliche Dinge gefasst, z. B. der Stil von Texten, das eigene Leseverstehen, aber auch die Inhalte (so allein die Beschreibung bei Narvaez et al., 1999, S. 490). Der kleinste gemeinsame Nenner ist, dass Leserinnen und Leser lesebezogene Einschätzungen treffen. Dies haben vereinzelt Experimente für kognitiv anspruchsvolle Leseanlässe nachweisen können, vor allem für expositorische Texte (Narvaez et al., 1999).
- *Aufgaben mit Auftrag in puncto Lernen führten zu weniger assoziativen Inferenzen.* Nicht nur das, was Leserinnen und Leser *mehr* machen, ist von Interesse, sondern auch das, was sie *weniger* demonstriert haben. Hier ergibt sich in einigen Studien das Bild, dass Personen, die zum Zweck der Unterhaltung lesen, mehr assoziative Äußerungen zu Protokoll gaben (Bohn-Gettler & Kendeou, 2014; Linderholm & van den Broek, 2002; van den Broek, Lorch et al., 2001). Dies wird als Indikator für ein loserer, lokales Verknüpfen gewertet denn als Bemühen, dem Textsinn kohärent gesamthaft Gestalt zu geben.

Die sich auf der Analyse diverser einzelner Studien ergebenden Muster, dass Personen beim Lesen, um zu lernen, im Vergleich zum Lesen zur Unterhaltung analytischer vorzugehen scheinen, lässt sich auch an einer einzelnen Studie verdeutlichen. Daher enthält der Kasten „Illustrierendes Beispiel 4“ die Darstellung einer solchen zentralen Untersuchung. Dieses Beispiel beschränkt sich nicht allein auf Inferenzen, sondern umfasst mehr Prozesskategorien des Lesens, was dahingehend erhellend ist, dass sich die Differenzen im Vorgehen im Gesamt der untersuchten Prozesse deutlicher herauschälen.

Illustrierendes Beispiel 4

Lesen zum Zweck des Lernens – die fokussierte lokale Kohärenzbildung als Vorgehensweise

A) Was war das Erkenntnisinteresse?

Die Studie prüfte Unterschiede in den kognitiven Prozessen beim Lesen, die sich bei Personen in Verbalprotokollen haben rekonstruieren lassen (Linderholm & van den Broek, 2002). Dadurch sollte eruiert werden, ob sich ein Lesen zur Unterhaltung von einem zur Vorbereitung auf eine Prüfungssituation unterscheidet. Dies wurde sowohl für die Prozesse als auch für die Produkte geprüft. Zusätzlich erhob man die Arbeitsgedächtnisspanne und unterschied Personen danach, ob ihre verbale Arbeitsgedächtnisleistung hoch oder niedrig war. Die Arbeitsgedächtnisleistung wurde als Kovariate durchgängig in den Analysen einbezogen.

B) Wer wurde untersucht?

An der Studie nahmen im Mittel 21-jährige Studierende der University of Minnesota teil, das Verhältnis von Frauen zu Männern war 4:1. Weitere Informationen zu den Testpersonen wurden nicht berichtet. Ausgewertet wurden die Daten von 58 Personen.

C) Welche Instrumente kamen zum Einsatz?

Die Personen wurden per Zufall einem von zwei Texten und einem von zwei Lesezwecken zugewiesen. Die Gruppe Unterhaltung sollte sich vorstellen, den jeweiligen Sachtext so zu lesen, als würde er das Interesse erregt haben, als sie durch eine Zeitschrift geblättert haben. Die Gruppe Lernen sollte sich hingegen vorstellen, den Text so zu lesen, als würde sie ihn für das Vorbereiten auf eine schriftliche Prüfung rezipieren. Ihre Vorgehensweisen wurden aufgenommen, und die Personen verbalisierten ihre Gedanken, sodass Lautdenk-Protokolle entstanden, welche transkribiert und in neun Kategorien kodiert wurden. So wurden die Prozesse parallel erfasst und später varianzanalytisch ausgewertet.

Nach dem Lesen wurden die Personen gebeten, schriftlich zu notieren, woran sie sich erinnern konnten. Diese Erzeugnisse wurden danach ausgewertet, an wie viele Informationen sich die Personen erinnern hatten. Es entstanden damit Produktmaße.

D) Was waren die Hauptbefunde?

Beginnend mit den *Produktmaßen* lässt sich konstatieren: Personen mit geringer Arbeitsgedächtnisleistung erinnerten sich in beiden Experimentalgruppen an gleich viele Inhalte. Ihre Leistungen entsprachen derjenigen Leistungen, die auch Personen mit hoher Arbeitsgedächtnisleistung beim Lesen zu Unterhaltungszwecken demonstrierten. Allerdings konnten sich Personen mit hoher Arbeitsgedächtnisleistung an rund 50 Prozent mehr Inhalte erinnern, wenn sie zu Zwecken des Lernens gelesen hatten. Ein weiterer Befund bestand darin, dass die verbale Arbeitsgedächtnisleistung stärker mit den Leistungen zum Zweck des Lernens zusammenhing als sie zum Zweck des Unterhaltungslesens korrespondierte. Bei den *Prozessen* zeigte sich besonders der Effekt sowohl des Leseziels als auch in Verbindung mit der Arbeitsgedächtnisleistung. In den neun Prozessen ergaben sich folgende Auffälligkeiten, die nach Aufgabeneffekten sortiert dargestellt werden:

- Bei den *Brückeninferenzen*, in denen Inhalte des aktuellen Satzes mit denen des vorherigen verknüpft wurden, war ein durchgängiger Vorteil zugunsten der Lern-Leserinnen und -Leser zu beobachten.
- Umgekehrt erklärten Personen sich die Inhalte des aktuell gelesenen Satzes beim Lesen, um zu lernen, weniger über ihr eigenes Vorwissen, d. h. es gab weniger *elaborative Inferenzen* beim Lernen.
- *Prädiktive Inferenzen* und *weiter entfernte Textinformationen aufgreifende Äußerungen* wiesen keine Differenzen in den Gruppen auf.
- Jedoch *paraphasierten* die lernenden Personen aktuelle Informationen stärker und sie *wiederholten* Informationen auch häufiger.
- Die zum Lernen lesenden Personen *überwachten ihr Vorgehen stärker metakognitiv*.

- *Assoziative Äußerungen* überwogen beim Lesen zur Unterhaltung, speziell in der Gruppe von Personen mit *geringerer Arbeitsgedächtnisleistung*. Ebenfalls gab es mehr *evaluative Einschätzungen* beim Unterhaltungslesen, d. h. mehr spontane Meinungsäußerungen.

Alles in allem wirkte das Vorgehen der lernenden Personen stärker auf die zum Zeitpunkt der Äußerungen aktuellen Informationen fokussiert, die sie präsent hielten und lokal kohärent verbanden.

E) Diskussion

Die Studie hat neben den an dieser Stelle nur sehr selektiv dargestellten Ergebnissen zu Arbeitsgedächtniseffekten vor allem ein Muster herausgearbeitet. Die zum Vergnügen lesenden Personen unterschieden sich von den lernenden Personen darin, dass erstgenannte assoziativer und weniger textnah zu lesen schienen. Die lernenden Personen wirkten hingegen stärker bemüht, über textnahe Inferenzen Kohärenz zu stiften. Es wirkt daher im Licht dieser Analysen so, als hätten sich die lernenden Personen den gesamten Text eher erarbeitet, was – zumindest für die Personen mit höherer Arbeitsgedächtnisleistung – mit einer deutlich erhöhten Erinnerungsleistung korrespondierte. Es sind demnach verschiedene Vorgehensweisen beobachtbar, die erst in ihrem Verbund davon zeugen, worauf die Blickbewegungen bereits erste Hinweise gegeben haben: dass beim Lernen stärker analytische Kohärenzprozesse zum Einsatz gelangen.

Zusammenfassung zu allgemeinen Relevanzeffekten des Lesezwecks auf Produkte und Prozesse des Lesens

Die Studien zu allgemeinen Zwecken des Lesens wie Unterhaltung bzw. Lernen ergeben aus der *Produktwarte* kein einheitliches Bild, was klare Unter- oder Überlegenheiten betrifft. In Leseverstehenstests waren zum Zweck des Lernens lesende Personen nur in einem Teil der Studien jenen Personen überlegen, die zum Zweck der Unterhaltung lasen. Dennoch: Der Zweck des Unterhaltungslesens erwies sich in den Experimenten in den Leistungsdaten nicht als überlegen. Mit Blick auf die Genauigkeit des eingeschätzten Leseerfolgs – einem Forschungsfeld, das allgemein zunehmend Aufmerksamkeit auf sich zieht – gestaltet sich das Bild noch unvollständiger. Dennoch scheint es Personengruppen Mühe zu bereiten, ein kalibriertes Urteil zu fällen, wie erfolgreich sie je nach Leseauftrag in ihrer Versteheleistung sind.

Etwas klarer wirkt die Forschungslage bei den *Prozessen*. Lesezeiten als zunächst unspezifisches, aber tiefere Verarbeitungen nahelegendes Konstrukt divergieren partiell in den Experimenten zugunsten des Zwecks des Lernens: Wer lernen soll, liest länger, und zwar weil – darauf verweisen erste Eye-Tracking-Befunde – zentrale Inhalte häufiger erneut betrachtet werden. Lautdenk-Protokolle ergänzen das Bild insofern, als sie weitestgehend darin übereinstimmen, dass es zu mehr Prozessen des Verfügbarhaltens von Informationen kommt, wenn

Personen lernen sollen, und dass sie diese Informationen anreichernd prüfen bzw. verknüpfen. Zudem wirkt das Vorgehen weniger assoziativ und stärker textgesteuert. Summa summarum ergeben die bisher berichteten Prozessdaten das Muster einer stärker analytischen und fokussierten Vorgehensweise, wenn Personen die gleichen Texte zum Zweck des Lernens rezipieren statt mit dem Ziel der Unterhaltung. Dies lässt sich im Lichte der Kohärenzstandards als Ausdruck variierender Kohärenzstandards je nach Aufgabe interpretieren.

2.2.2.3 Allgemeiner Zweck des Lesens 2: Lesen mit unterschiedlichen Perspektiven

Eine zweite Variante, wie sich relevanzbezogene Aufträge – und mit ihnen: die Aufgabenmodelle nebst daraufbasierenden Prozessen und Produkten – differenzieren lassen, sind Aufgaben, in denen Leserinnen und Leser aus einer Perspektive heraus Texte lesen sollen, etwa einen Text über ein Haus aus der Sicht eines Einbrechers oder aus der einer Hauskäuferin. Das heißt: Jemand soll eine breite Kategorie von Informationen hinsichtlich ihrer Relevanz erkennen, beurteilen und memorieren (Kaakinen & Hyönä, 2011; McCrudden & Schraw, 2007). Zur Wirkweise, wie Perspektiven das Leseverstehen beeinflussen, haben Kaakinen und Hyönä (2008, 2011) ihr *Modell zum perspektivengetriebenen Leseverstehen* vorgelegt. Dieses Modell besagt, dass die Perspektiveneffekte dadurch zustande kommen, dass die zentrale Exekutive als aufmerksamkeitslenkende Instanz auf zweierlei Art den Verstehensprozess beeinflusst. Bezogen auf den Input aus dem sensorischen System, also den durch das Lesen mit den Augen und kognitiv verarbeiteten sprachlichen Inhalten, ist erstens anzunehmen, dass irrelevante, also nicht zur Perspektive passende Inhalte unterdrückt werden. Dadurch gelangen selektiv perspektivenkongruente Informationen in die Aufmerksamkeit; alle anderen werden so weit wie nötig verarbeitet. Im Arbeitsgedächtnis werden diese Inhalte dann mit jenen Inhalten kombiniert, die aus dem Langzeitgedächtnis stammen und durch die Perspektive aktiviert werden, sodass es wahrscheinlicher ist, dass es zu einer kognitiven Integration von textuellen und vorwissensbasierten Informationen kommt. Daraus folgt eine doppelte Privilegierung von Informationen durch die Lenkung der Aufmerksamkeit: Perspektivenkompatible Informationen aus Texten werden zum einen fokussiert, und die dazu passenden Gedächtnisinhalte werden zum anderen gezielt aktiviert – in der Essenz werden diese beiden Informationen dadurch gezielt aufeinander bezogen, was zudem zu einem leichteren Eingang ins Langzeitgedächtnis führen dürfte.

Im Lichte dieses Modells ist es wahrscheinlich, dass sich diese Annahmen ebenfalls in Prozessen und in den Produkten des Lesens manifestieren. Die Forschung überprüft das in der Regel in experimentellen Settings wie diesen: Es werden oftmals einzelne Texte vorgelegt, in denen ein Haus, mehrere Länder oder Krankheiten beschrieben werden. Diese Texte lesen die Testpersonen

dann und werden gebeten, als imaginierte Diebe, Hausreinrichterinnen oder Hauskäufer zu lesen, die Vor- oder Nachteile eines einzelnen Landes zu prüfen oder die Eigenheiten einzelner Krankheiten zu beachten. Mit diesen Aufträgen werden dann Prozess- und Produktmaße in den Experimenten fokussiert, wobei die Analysegegenstände die sich aus der Perspektive ergebenden relevanten bzw. irrelevanten Informationen bzw. Informationseinheiten sind.

Produkte

Eine Gemeinsamkeit der Studien besteht darin, dass durch die Perspektivenvorgabe die Texte in der Regel danach unterteilt sind, welche der enthaltenen Informationen für die jeweilige Perspektiven entweder relevant oder irrelevant sind. Damit sind zugleich die beiden Analysefokusse umrissen: relevante und irrelevante Informationen. An dieser Stelle werden aus den empirischen Befunden zwei Gruppen von Resultaten präsentiert, nämlich zum einen in puncto Beurteilung der Textteile und zum anderen hinsichtlich der Erinnerungsleistungen zu relevanten Informationen nach dem Lesen.

1) Beurteilung von Textsegmenten in ihrer perspektivenbasierten Relevanz

Welche Perspektive die Personen auf ein und denselben Text beim Lesen einnehmen sollen, ist für die sogenannten Perspektiveneffekt bedeutsam: Je nach Perspektive verändern sich Leistungen und Prozesse im Umgang mit relevanten bzw. irrelevanten Informationen (Kaakinen & Hyönä, 2011). Hier lassen sich drei Befundmuster auf unterschiedlich breiter Basis aus der Forschung extrahieren:

- *Die Leserinnen und Leser stuften perspektivisch relevante Informationen als interessanter und als relevanter ein, wenn sie ihrem Lesezweck entsprachen.* Dahinter verbirgt sich das in einigen Studien gefundene Muster, dass Personen Informationen divergierend beurteilen: Was nach der eingenommenen Perspektive relevant sein sollte, beurteilen die Personen entsprechend auch als relevanter. Umgekehrt weisen sie Informationen, die sich für die eingenommene Perspektive als weniger relevant darstellen, eher den Status irrelevant zu. Es kommt also zu einer gut dokumentierten Differenzierung (Goetz et al., 1983, Experiment 1), die sich auch in Urteilen über die Interessantheit von einzelnen perspektivenkongruenten Textsegmenten niederschlägt (Ramsay & Sperling, 2010, Experiment 2; Schraw & Dennison, 1994, Experiment 2).
- *Wechselnde Perspektiven führten zu unterschiedlichen Beurteilungen identischer Informationen.* Die identischen Informationen ein und desselben Textes werden je nach Perspektive ganz anders beurteilt, d.h. diese Informationen haben einen unterschiedlichen instrumentellen Wert, also eine divergierende Relevanz. Ein besonders gutes Beispiel hierfür ist, dass die

Relevanzbeurteilungen von Sätzen je nach Perspektive nicht oder allenfalls gering miteinander zusammenhängen (Goetz et al., 1983, Experiment 1; Pichert & Anderson, 1977, Experiment 1): Wenn beispielsweise Personen einen Text lesen, in welchem ein Grundstück beschrieben wird, dann haben die Urteile von ‚Dieben‘ oder ‚Hauskäuferinnen und Hauskäufern‘ statistisch betrachtet nichts gemein. Das heißt: Personen, die die Texte aus der Warte von imaginierten Einbrechern lesen, beurteilen gänzlich andere Informationen als relevant (wie den Ort der Schmucklagerung), als Personen, die als potenzielle Käuferinnen und Käufer des Hauses auf andere Informationen Wert legen (wie den Bauzustand).

- *Personen verwendeten für eigene schriftliche Texte mehr jene Texte und Informationen, die zu ihrer Perspektive passten.* Noch wenig untersucht ist, ob sich die Personen mit spezifischen Perspektiven auf einen Sachverhalt gemäß dem Leseauftrag auch in der Nutzung von Quellen unterscheiden. Einer spanischen Studie zufolge war das Fall: In ihr nutzten die Studierenden häufiger für ihre eigene Argumentation jene Texte, welche jene Position bestätigten, die sie selbst einnehmen sollten (Cerdán, Marín & Candel, 2013).

Auf den allgemeinsten Nenner gebracht sind Perspektiveneffekte relativ deutlich nachgewiesen worden, sei es in Indikatoren direkter Urteile über kleinere Informationseinheiten und ihrer Unabhängigkeit voneinander, seien es indirekte Maße, die sich aus der Verwendung von ganzen Dokumenten in eigenen Textprodukten ergeben. All dies spricht für eine Sensibilität der Testpersonen in den Experimenten hinsichtlich der perspektivenbezogenen Relevanz.

2) Erinnerungsleistungen an Textinhalte

Zwei Schwerpunkte liegen der Forschung zu den Erinnerungen als – natürlich nur bedingter Indikator für tatsächliches umfassendes Leseverstehen – zugrunde: Einerseits eruierten diverse Studien, ob Personen besser relevante Informationen aus Texten memorierten, andererseits haben – jedoch nur wenige Studien – zusätzlich geprüft, ob die Personen innerhalb der relevanteren Informationen sensibel sind für die Hierarchie des instrumentellen Werts. Die Ergebnisse sind gebündelt wie folgt:

- *Leserinnen und Leser erinnerten sich insgesamt an mehr an die Inhalte, die der für die Aufgabenlösung einzunehmenden Perspektive entsprachen, also relevante Informationen enthielten.* Dieses Muster ist in über 20 Experimenten nachgewiesen worden, wobei sich die Erinnerungsleistungen meist auf kurzfristige Behaltensleistungen beschränkten (Anderson et al., 1983, Experiment 1; Anderson & Pichert, 1978, Experiment 1; Baillet & Keenan, 1986; Bohn-Gettler & McCrudden, 2018, Experimente 1 und 2; Burton &

Daneman, 2007; Di Vesta & Di Cintio, 1997; Goetz et al., 1983, Experiment 2; Kaakinen et al., 2001, 2002, 2003, Experiment 1; Kaakinen & Hyönä, 2005, 2008; Lee-Sammons & Whitney, 1991, Experiment 2; McCrudden et al., 2010; McCrudden & Schraw, 2010; Newsome, 1986, Experiment 1; Schraw et al., 1993, Experimente 1–3; Schraw & Dennison, 1994, Experimente 1–3). Trotz der breiten Masse an empirischen Befunden zugunsten positiver perspektivenkongruenter Erinnerungsleistungen ist auffällig, dass dies vor allem auf Studien mit Erwachsenen zutrifft: In zwei Experimenten konnten keine Effekte bei Jugendlichen der Klassenstufen 5, 6 und 8 nachgewiesen werden (Ramsay & Sperling, 2015, Experimente 1 und 2).

- *Als je relevanter Informationen aus einer Perspektive eingestuft wurden, desto mehr wurden sie von Personen memoriert.* Informationen in Texten lassen sich einerseits binär als relevant oder irrelevant klassifizieren, aber auch entlang eines Kontinuums reihen. In Studien, die ein solches Relevanzkontinuum für die Positionierung der einzelnen Informationen nutzten, ergab sich ein Zusammenhang von höherer Relevanz und besserer Erinnerung: Als je relevanter die Information von den Forschenden eingeschätzt wurden, desto eher wurde sie von den Testpersonen memoriert – dies spricht für eine sensible Enkodierung (Pichert & Anderson, 1977, Experiment 2). Das zeigt sich methodisch anders ausgewertet auch in anderen Studien (Newsome, 1986, Experiment 1).

Die Forschung hat deutlich gemacht, dass ältere Personen sich besser an jene Informationen erinnern, die aus ihrem Arbeitsauftrag heraus relevant sind, wobei dies meistens Summenwerte von erinnerten Informationseinheiten waren. Dies dürfte sich darüber erklären, dass diese Informationen eine höhere Relevanz zugewiesen bekommen haben, jedoch bleibt der exakte Wirkmechanismus trotz der doch vielen Experimente erstaunlich unterbelichtet.

Prozesse

Die Prozessdaten sind hinsichtlich der Perspektiveneffekte bisher in zwei Gruppen von Studien näher untersucht worden. Eine erste Gruppe von Experimenten erfasste die Lesezeiten, eine zweite Gruppe von Studien maß die Augenbewegungen.

1) Lesezeiten

Personen lesen jene Textteile länger, die ihrer einzunehmenden Perspektive entsprechen. Diese verlängerten Lesezeiten gelten als Indikatoren für eine verstärkte Aufmerksamkeit und konnten mehrfach nachgewiesen werden (Cerdán, Marín & Candel, 2013; Goetz et al., 1983, Experiment 2; McCrudden et al., 2010;

McCrudden & Schraw, 2010). Dabei ist die reine Unterscheidung von Experimentalgruppen mit der einen oder anderen Perspektive auf den zu lesenden Text nur ein Teil der empirischen Wahrheit. In einer bemerkenswerten Studie, in der nach dem eigentlichen Experiment die teilnehmenden Erwachsenen zu ihren Lesezielen befragt wurden, zeigten sich nicht nur eigensinnige Umwidmungen des Leseziels in einem Teil der Experimentalgruppe bzw. im Falle der Kontrollgruppe. Die Befunde unterstreichen, dass sowohl internale als auch externale Quellen die Relevanz beeinflussten und zu einer selektiven Aufmerksamkeit führten, die sich in Theorien wie jener zur Nutzung einer selektiven Aufmerksamkeitsstrategie einfügen (Reynolds, 1992).

Eine hervorzuhebende Studie enthält der Kasten „Illustrierendes Beispiel 5“. Hervorzuheben ist die Studie vor allem deshalb, weil sich nicht nur die erwartbaren längeren Lesezeiten bei relevanten Inhalten haben nachweisen lassen. Vielmehr ist das subjektive Ziel der Teilnehmenden beim Lesen in der Studie deutlich geworden: Diese haben je nach ihrer Zuschreibung von Relevanz verschiedene Textteile unterschiedlich lang gelesen.

Illustrierendes Beispiel 5

Der Eigensinn der Leserinnen und Leser: externe und individuelle Ziele im Zusammenspiel – Befunde aus Lesezeiten und Interviewdaten

A) Was war das Erkenntnisinteresse?

In der Studie bestand ein zweigeteiltes Forschungsinteresse: Einerseits wollte man mit Prozessdaten (Lesezeiten) und mit Produktmaßen (Anteil erinnelter Inhalte) Effekte von Aufgabenstellungen prüfen (McCrudden et al., 2010). Andererseits interessierte man sich mittels Leitfadeninterviews dafür, welche Ziele und Vorgehensweisen die Personen selbst zu Protokoll geben, die sie bei der Aufgabenbearbeitung bemüht hatten.

B) Wer wurde untersucht?

An dem Experiment nahmen 52 US-amerikanische Studierende aus bildungswissenschaftlichen Studiengängen teil, die im Durchschnitt 24 bis 25 Jahre alt waren. Vier von fünf der Testpersonen waren Frauen, über zwei Drittel waren europäische Amerikanerinnen und Amerikaner.

C) Welche Instrumente kamen zum Einsatz?

Das Design der Mixed-Methods-Studie sah vor, dass die Testpersonen einen digitalen Text lasen, in dem vier verschiedene abgeschiedene Länder beschrieben wurden. Zwei der Experimentalgruppen sollten die positiven und negativen Eigenschaften von jeweils einem der Länder als möglichem Lebensort gezielt filtern, die Kontrollgruppe erhielt keine solche Aufforderung, sondern sollte den Text so lesen, dass sie ihn verstand. In der Studie wurden quantitative und qualitative Zugänge kombiniert: Quantitativ erfasst wurden Lesezeiten und die Erinnerung an sämtliche Informationen – analysiert wurden

dann vor allem jene aufgeschriebenen Inhalte zu den beiden Ländern, auf die die beiden Experimentalgruppen achten sollten. Qualitativ war der zweite Teil der Studie, in dem die Hälfte der Teilnehmerinnen und Teilnehmer direkt nach dem quantitativen Teil zu ihrem Vorgehen interviewt wurde. Aus den transkribierten Leitfadeninterviews wurden inhalts-analytisch Vorgehensweisen rekonstruiert und drei Gruppen des Vorgehens rekonstruiert. Quantitative und qualitative Daten wurden trianguliert.

D) Was waren die Hauptbefunde?

Die quantitativen Daten ergaben, dass Personen Sätze länger lasen, die Inhalte zu dem Land enthielten, auf das sie achten sollten. Sie erinnerten sich zudem an mehr Inhalte aus diesen Sätzen. Die qualitativen Befunde ergaben, dass sich in den beiden Experimentalgruppen zwei Arten des Vorgehens herauskristallisierten. Die erste bestand darin, dass sich die Personen vornehmlich auf dasjenige Land konzentrierten, auf das sie je nach Aufgabenstellung achten sollten. Ihr Vorgehen spiegelte dies wider: Sätze, die zum Zielland passten, wurden als solche erkannt, länger gelesen und besser memoriert. Das Gegenteil traf auf irrelevante Sätze zu. Die zweite Vorgehensweise war anders: Hier war den Personen zwar das Ziel präsent und sie schätzten die zu lesenden Sätze auch auf der Basis der Aufgabenstellung als relevant ein. Gleichwohl hatten sie eigene Ziele wie einen Vergleich der Länder oder persönliche geografische Präferenzen. Sie fokussierten entsprechend auch Sätze länger, die mit dem extern vorgegebenen Ziel nichts direkt zu tun hatten, und sie erinnerten sich besser an deren Inhalte. Die insgesamt dritte Vorgehensweise, die vor allem in der Kontrollgruppe anzutreffen war, bestand darin, die Vertrautheit der Informationen zur Einschätzung der Relevanz für die Zielerreichung zu nutzen. Sätzen mit unvertrauten Inhalten schenken die Personen weniger Aufmerksamkeit.

E) Diskussion

Die Studie bestätigte in ihrem quantitativen Teil Bekanntes: Relevanzinstruktionen führen zu längeren Lesezeiten und besseren Behaltensleistungen jener Textteile, die zur Aufgabenstellung passten, in diesem Fall: zur Perspektive, unter der die Personen den Text lasen. Die qualitativen Daten differenzieren dies aus: In der Experimentalgruppe gab es nämlich kein uniformes Vorgehen, sondern zwei Untergruppen: Nur ein Teil der Probandinnen und Probanden ging rein zielstrebig in dem Sinne vor, ein möglichst schlankes, aufgabenkonformes Vorgehen zu wählen, das auf die rein verwertbaren Sätze und ihre Inhalte abzielte – ganz im Einklang mit den quantitativen Daten. Ein weiteres Vorgehen war aber eine Ausweitung, bei der auch irrelevante Inhalte gelesen wurden, um ein eigenständig ergänztes Ziel zu erreichen, das außerhalb der Aufgabenstellung lag. In der dritten Vorgehensweise der Kontrollgruppe war das Vorwissen die entscheidende Stellgröße: An ihm bemaßen die Testpersonen, welche Inhalte sie für lesenswert hielten, aber sie hatten kein besseres Erinnerungsvermögen in Bezug auf einzelne Länder. Damit zeigte sich in einem Teil der Experimental- und in der Kontrollgruppe ein eigensinniges Vorgehen, bei dem die internen Ressourcen der Leserinnen und Leser die Relevanzzuweisung beeinflussten.

2) Eye-Tracking-Indikatoren

Informationen, die relevant sind, werden von Leserinnen und Lesern mit Blicken bedacht. Dieses Befundmuster haben vor allem finnische Experimente wiederholt demonstriert:

- Als ein erstes Muster lässt sich festhalten, dass *Leserinnen und Leser jene Informationseinheiten länger mit ihrem Blick fixieren, die für ihr Leseziel* – in diesem Fall: Informationen, die gemäß ihrer einzunehmenden Perspektive relevant sind – *eine höhere Verwertbarkeit besitzen*. Was relevant ist, das betrachten die Personen also länger. Das trifft zum einen auf Sätze als Bezugseinheit zu (Kaakinen et al., 2002, 2003, Experiment 1; Kaakinen & Hyönä, 2005), zum anderen lässt sich derlei auch noch auf einer kleineren Einheit, nämlich der Wortebene, feststellen (Kaakinen & Hyönä, 2007).
- *Die längeren Lesezeiten relevanter Informationen ergeben sich zum einen aus längeren Verweildauern des Blicks beim ersten Lesen*. Das bedeutet: Personen fixieren bereits beim Erstkontakt solche Sätze länger, die für ihre Aufgabenstellung relevante Informationen enthalten (Kaakinen et al., 2002, 2003, Experiment 1; Kaakinen & Hyönä, 2008). Auf Wortebene ist festzustellen gewesen, dass die Verweildauern des Blicks umso länger waren, je weiter hinten die Position der Wörter in den Sätzen mit relevanten Informationen war: Die Fixationen des Blicks dauerten gegen Satzende immer länger (Kaakinen & Hyönä, 2007).
- *Relevante Informationen lesen die Personen zum anderen auch in stärkerem Maß wiederholt*. Leserinnen und Leser lesen Sätze mit relevanten Informationen nicht nur einmal, vielmehr kehren ihre Blicke zu ihnen zurück (Burton & Daneman, 2007; Kaakinen & Hyönä, 2007), und in diesem Fall fixieren sie diese Informationen ebenfalls länger als irrelevante Sätze (Kaakinen et al., 2002, 2003, Experiment 1). Längere Fixationsdauern beim erneuten Betrachten korrespondieren mit einer besseren Erinnerungsleistung (Burton & Daneman, 2007).
- *Leserinnen und Leser lassen irrelevante Satzinhalte aus*. Wenn Leserinnen und Leser auf Textteile stoßen, die für die einzunehmende Perspektive irrelevant sind, besteht die erhöhte Wahrscheinlichkeit, dass sie Satzteile auslassen. Das gilt insbesondere dann, wenn das Thema eines Textes vertrauter ist, während bei unvertrauten Themen das Auslassen weniger wahrscheinlich ist (Kaakinen & Hyönä, 2007). Auch fixieren Leserinnen und Leser Textteile weniger, sobald sie erkannt haben, dass sie keine relevanten Informationen zu enthalten versprechen (Kaakinen & Hyönä, 2014).

Die Daten zu den Fixationen lassen sich zu einem Gesamtbild verdichten: Relevante Information lesen die Personen länger und durchaus mehrfach, während sie die aus ihrer Sicht irrelevanten Informationen überspringen. Was die Personen beim Lesen denken, das verdeutlicht das nachstehende Exempel („Illustrierendes Beispiel 6“). Es ist instruktiv, weil es Muster von Eye-Tracking und verbalisierten Prozessen hervorbringt, die besser verstehen lassen, was an definierten Stellen im Text an Leseprozessen triangulativ nachweisbar ist.

Illustrierendes Beispiel 6

Was denken Personen, die Texte perspektivisch nutzen? Ergebnisse aus einer Studie mit Eye-Tracking und lautem Denken

A) Was war das Erkenntnisinteresse?

Die Studie widmete sich der Frage, inwiefern Eye-Tracking-Maße und Lautdenk-Protokolle beim Lesen aus verschiedenen Perspektiven zu korrespondierenden Ergebnissen führen (Kaakinen & Hyönä, 2005). Daneben hat die Studie vergleichend geprüft, ob die mit den Lautdenk-Protokollen erhobenen kognitiven Prozesse des Lesens je nach Aufgabenstellung anders ausgeprägt waren und ob sich auf einer aggregierten Ebene Prozessgruppen darin unterscheiden, bei welchen per Eye-Tracking erfassten Indikatoren sie besonders häufig (nicht) vorkamen.

B) Wer wurde untersucht?

Drei Dutzend finnische Studierende, davon drei Viertel Frauen, waren die Teilnehmerinnen und Teilnehmer der Studie. Sie waren im Durchschnitt 24 Jahre alt und studierten Psychologie.

C) Welche Instrumente kamen zum Einsatz?

Die Studierenden lasen einen 960 Wörter langen Text über vier Krankheiten. Die Aufgabe sah vor, dass sich die Testteilnehmerinnen und -teilnehmer auf eine Krankheit konzentrieren sollten, wobei es zwei Gruppen gab, und jede Gruppe fokussierte jeweils nur eine Krankheit: Typhus oder Trigeminusneuralgie. Jeweils zehn Sätze waren für eine der beiden Krankheiten besonders relevant bzw. irrelevant und bildeten für die Eye-Tracking-Daten die Grundlage. Dort wurden also Blickbewegungen erfasst und unterschieden, ob die Sätze erstmalig oder wiederholt betrachtet wurden und wie lang diese Fixationen dauerten. Nach dem Lesen von zehn jeweils relevanten Sätzen – entweder diejenigen, die sich auf Typhus bezogen, oder die zur Trigeminusneuralgie – erfolgten Hinweise, die zum lauten Denken animierten. In diesen Situationen wurde das Eye-Tracking nicht mehr erfasst. Eye-Tracking und lautes Denken fanden demnach zeitversetzt statt, allerdings sehr dicht aufeinander folgend. Die Lautdenk-Protokolle wurden transkribiert und kodiert, wobei sechs Verstehensprozesse fokussiert wurden.

Am Ende des Experiments – und nach dem Absolvieren einer Zwischenaufgabe – bearbeiteten die Studierenden einen Erinnerungstest, in dem sie alle Inhalte niederschrieben, an die sie sich aus dem Text erinnern konnten. Die Probandinnen und Probanden sollten sich dabei also nicht nur auf die Informationen beschränken, die alleinig die ihnen zugeteilte Krankheit betraf. Ausgewertet wurden aber nur die Inhalte der zehn (ir-)relevanten Sätze, die für die jeweilige Perspektive beim Lesen entscheidend waren. Es gab also Prozess- und Produktmaße in dieser Studie.

D) Was waren die Hauptbefunde?

Die Hauptergebnisse spiegeln bei den Produkt- und bei den Prozessmaßen, die sich allein auf die Eye-Tracking-Maße beziehen, weitestgehend Bekanntes wider. Die Erinnerungsleistungen an jene Satzinhalte mit der zugeteilten Krankheit waren besser als an die der zweiten, für die Aufgabenbearbeitung irrelevanten Krankheit. Hinzu kamen längere Blickverweildauern auf die relevanten Sätze bei der erstmaligen Betrachtung.

Die Lautdenk-Protokolle ergänzen das Bild: Von den sechs Prozessen erwies sich nur einer als überzufällig häufig: Personen, die relevante Sätze lasen, wiederholten öfter Inhalte des Gelesenen, hielten also die Informationen im Arbeitsgedächtnis präsent. Weiterführende Analysen kombinierten die Kategorien mündlicher Verbaldaten: zum einen zur Gruppe von oberflächenorientierter Verarbeitung (Paraphrasen und Assoziationen) und zum anderen zur tiefenorientierten Auseinandersetzung (Selbsterklärungen und Fragen). Dabei ergab sich dieses Muster für die Fixationsdauern: Je länger die Personen Sätze betrachteten, die sie erstmals lasen, desto häufiger neigten sie dazu, tiefenorientierte Verarbeitungsprozesse zu verbalisieren. Dieser Effekt war nicht darauf limitiert, dass er nur bei den jeweils relevanten Sätzen aufgetreten war. Anders gesagt: Längere Fixationen beim Erstkontakt gingen mit stärker auf Tiefenverstehen abzielenden Prozessen gemäß Lautdenk-Protokollen einher.

E) Diskussion

Das Experiment unterstreicht, dass längere Fixationen im Allgemeinen jenen tiefenorientierten, stark vorwissensbasierten Prozessen wie Fragen und Selbsterklärungen vorausgehen, die Personen kurz zeitverzögert nach dem erstmaligen Lesen von Informationen zu Protokoll geben. Dies ist ein Effekt gewesen, der sich in der Studie unabhängig davon hat nachweisen lassen, ob die mit längeren Blicken bedachten Informationen relevant oder weniger relevant waren. Aus methodologischer Warte lässt sich das als ein Indiz für die Validität sowohl der Eye-Tracking-Maße als auch der Verbaldaten werten.

Das Paraphrasieren wurde in der Studie dem oberflächenorientierten Verarbeiten zugeordnet, und es trat besonders bei jenen Sätzen auf, die für die lesenden Personen relevant waren. Zusammen mit den längeren Fixationsdauern dieser Sätze lässt sich der Befund so interpretieren, dass die Testteilnehmerinnen und -teilnehmer diese Informationen mit mehr Aufmerksamkeit bedachten, was sich zudem in den besseren Erinnerungsleistungen

niederschlag. Insgesamt macht die Studie also auf zwei Muster aufmerksam: auf den Zusammenhang von Fixationen und Tiefenverarbeitung einerseits und eine auf Enkodierung abzielende Aufmerksamkeitszuwendung andererseits.

Zusammenfassung zu allgemeinen Relevanzeffekten der Perspektive auf Produkte und Prozesse des Lesens

Perspektiveneffekte werden jene Folgen genannt, nach denen eine spezifisch auf den zu lesenden Text einzunehmende Warte beeinflusst, welche Informationen in ihrem Gesamt für relevant gehalten werden. Experimentalstudien, welche eine Variation vorgenommen hatten, mit welcher Perspektive die mehrheitlich erwachsenen Testpersonen ein und dieselben Texte lasen, erbrachten ein konsistentes Bild. Bezüglich der *Produkte* konnten sie auf breiter Basis belegen, dass Personen sich besser an jene Informationen erinnerten, die für sie relevant waren und die sie – wo es erforscht wurde – als relevanter einstufen. Ob sich allerdings auch darüber hinaus weitere, umfassendere Effekte auf das Gesamtverständnis von Texten einstellen, ist wenig untersucht. Die *Prozessdaten*, die sich auffällig oft auf unspezifische Maße wie Lesezeiten und Fixationsdauern beziehen, verweisen im Gesamt auf eine stärkere Zuwendung der Aufmerksamkeit auf perspektivisch relevante Informationen.

2.2.2.4 Allgemeiner Zweck des Lesens 3: Aufgabeneffekte (am Beispiel der Zieltextsorteneffekte bei der Diskurssynthese)

Aufgabeneffekte werden in der Leseforschung unter dem Stichwort „aufgabenorientiertes Lesen“ diskutiert (Vidal-Abarca et al., 2011). Damit ist gemeint, dass Leserinnen und Leser einen Text oder mehrere Dokumente lesen, um eine konkrete Aufgabe zu lösen. Das heißt: Die Aufgabe selbst ist entscheidend für die Zuweisung von Relevanz von Informationen, und die Leserin bzw. der Leser liest selektiv und so lange und in Interaktion von Aufgabenstellung und Text-/Dokumenteninhalten, bis sie bzw. er das Ziel für erreicht hält. Entsprechend wird dieses aufgabenbezogene Lesen als stark selbstreguliertes Lesen beschrieben (Bräten, Gil & Strømsø, 2011), und demgemäß dürfte hierbei wiederum das Aufgabenmodell als metakognitive Steuerungseinheit wichtig sein (s. Teilkap. [2.1.1](#) und [2.2.1.3](#)).

Während für die allgemeinen Zwecke des Lesens zur Unterhaltung vs. zum Lernen (s. Teilkap. [2.2.2.2](#)) und mit unterschiedlichen Perspektiven ([2.2.2.3](#)) methodisch zwei Vorteile bestehen – nämlich: gleiche Texte sowie gleiche Auswertungsmethode und gezielte Variation des jeweiligen Auftrags –, so trifft dies auf die Studien zu Aufgabeneffekten in dieser Form oft nicht zu. Die von McCrudden und Schraw (2007) versammelten Studien illustrierten dies bereits eindrucksvoll, da kaum seriöse studienübergreifende Muster extrahierbar waren, wenn beispielsweise die Arbeitsaufträge sehr stark divergierten. Aus diesem

Grund wird an dieser Stelle gezielt im Kasten „Illustrierendes Beispiel 7“ ein Forschungsüberblick über *Aufgabeneffekte auf Produkte* gegeben, die sich in das Forschungsfeld der „Diskurssynthese“ einreihen (s. dazu Teilkap. 5.1). Die zugrundeliegenden Experimente nutzten identische Sets multipler Dokumente und ließen die Experimentalgruppen einen schriftlichen Text über das Gelesene verfassen (sog. „materialgestütztes Schreiben“, Philipp, 2017). Diese Texte oder andere Lese- und Wissenstests wurden dazu genutzt, Aufgabeneffekte zu bestimmen. Der Aufgabeneffekt besteht in dem Fall darin, dass studienübergreifend geprüft wurde, ob die Zieltextsorte – also ob die Personen eine Zusammenfassung, eine Beschreibung oder eine Argumentation schreiben sollten – sich in differenziellen Leistungen niederschlug.

Illustrierendes Beispiel 7

Argumentieren oder informieren – welche Aufgabeneffekte hat die Zieltextsorte auf Leistungen, wenn Lernerinnen und Lerner materialgestützt schreiben?

A) Was war das Erkenntnisinteresse?

Gegenstand der Sekundäranalyse war die Frage, ob sich die Zieltextsorte – also der Auftrag, eine Zusammenfassung, eine Argumentation oder eine Beschreibung auf der Basis des materialgestützten Schreibens anzufertigen – auf Produktmaße auswirkt (Philipp, 2022a). Hierfür wurden kriterienbasiert Experimente ausgewählt, kodiert und aggregiert ausgewertet.

B) Wer wurde untersucht?

Es handelt sich um eine Sekundäranalyse von 18 kriterienbasiert selektierten Experimenten mit insgesamt 1.516 Probandinnen und Probanden. Mit einer Ausnahme wurden sämtliche Studien mit Studierenden durchgeführt, die aus westlichen Kulturkreisen stammten.

C) Welche Instrumente kamen zum Einsatz?

In den 18 Primärstudien wurden lesebezogene, geschlossene rezeptive Tests und schreibbezogene, offene expressive Tests verwendet – in allen Studien wurde mit einem identischen Set von mindestens zwei Texten operiert, die Experimentalgruppen verfassten jedoch Texte zu unterschiedlichen Zieltextsorten: Jeweils eine Experimentalgruppe schrieb nur einen Text, also beispielsweise entweder eine Zusammenfassung oder eine Argumentation. Die Gruppen absolvierten Tests, wobei teils auch die entstandenen Texte untersucht und Leistungsdaten aus ihnen gewonnen wurden.

Diese Tests mitsamt ihren ausgewerteten abhängigen Variablen waren im Gesamtkorpus sehr heterogen und wurden daher nach einem Schema kodiert und gebündelt. Dabei wurden Prozesse des Transformierens von Arten des Transformierens von Informationen unterschieden. Bei den drei *Prozessen* handelte es sich um 1) die Informationsauswahl (Welche Informationen werden abgedeckt?), 2) das Generieren von Dokumentenrelationen (Welche Informationen werden intertextuell verbunden?) und 3) die Inferenzbildungen

(Welche vorwissensbasierten Schlussfolgerungen nehmen Personen vor?). Die *Arten* der Transformationen betrafen 1) linguistische Transformationen (Fokus: sprachliche Modifikationen), 2) rhetorische Transformationen (thematische Restrukturierungen und -organisationen) sowie 3) konzeptuelle Transformationen (Nutzung und Anwendung von Informationen).

Da es sich um Studien mit Paarvergleichen – also etwa dem Vergleich von Argumentation und Zusammenfassung – handelte, wurden die Ergebnisse danach kodiert, ob die Vergleiche zugunsten einer Zieltextsorte einen statistischen signifikanten positiven Effekt im Sinne einer relativen Überlegenheit erbrachten (= positiver Effekt), eine relative Unterlegenheit der Zieltextsorte bestand (= negativer Effekt) oder keine zufallskritische Über- oder Unterlegenheit feststellbar war (= kein Effekt). Dieses Verfahren wird als *Vote-Counting* bezeichnet. Die Kodierungen wurden in rein deskriptiven Analysen als prozentualer Anteil ausgewiesen, sobald mindestens fünf Paarvergleiche vorlagen.

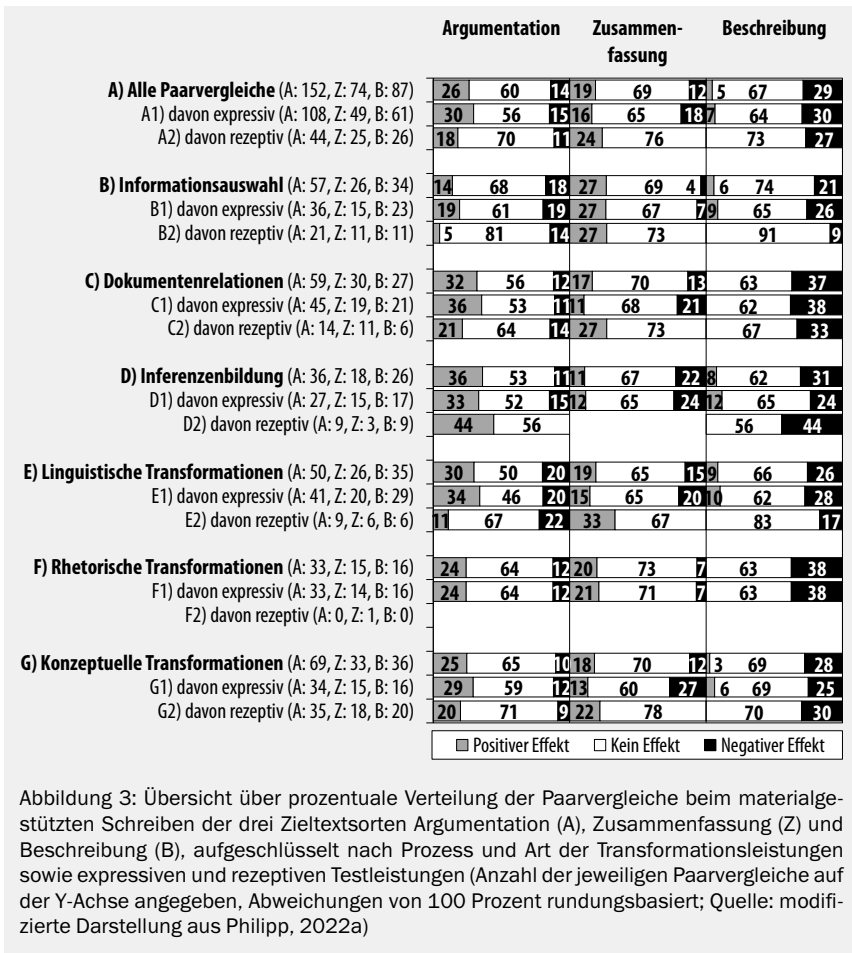
D) Was waren die Hauptbefunde?

Ausgewertet wurden die Vergleiche mit den drei am häufigsten analysierten Zieltextsorten Argumentation, Beschreibung und Zusammenfassung. Insgesamt 313 Paarvergleiche ließen sich so extrahieren und auswerten (s. Abbildung 3). Rund die Hälfte der Paarvergleiche entfiel auf die Argumentation als Zieltextsorte.

Von einer durchgängigen relativen Überlegenheit bei den Leistungen kann angesichts der Ergebnisse im Balkendiagramm in Abbildung 3 bei keiner der drei am häufigsten verglichenen Zieltextsorten die Rede sein. Auffällig ist aber, dass die *Beschreibung* besonders häufig entweder keinen Effekt oder sogar eine Unterlegenheit aufwies. Bei den *Zusammenfassungen* sind rezeptive Tests von eher positiven Effekten gekennzeichnet, bei expressiven Tests war das nur bei der Informationsauswahl (B1) und der Restrukturierung von Informationen (F1) erkennbar. Das Schreiben von *Argumentationen* ist bei expressiven Maßen eher überlegen, und zwar relativ durchgängig bei den Transformationsarten (E1, F1, G1), aber nur bei einem Teil der Transformationsprozesse (C1, D1). Vorteile in rezeptiven Maßen zeigten sich dort, wo Transformationen besonders kognitiv anspruchsvoll sind (D2, G2).

E) Diskussion

Der deskriptive Überblick mittels *Vote-Counting* illustriert, dass die Zieltextsorte differenzielle Zusammenhänge mit expressiven und rezeptiven Maßen in Paarvergleichen aufweist. Positive Effekte im Sinne von statistisch signifikanten höheren Ausprägungen in normativ wünschenswerter Richtung divergierten je nach Zieltextsorte. Die konsistentesten Überlegenheiten gab es bei Indikatoren in expressiven Tests bei Argumentationen und bei rezeptiven Leistungen bei Zusammenfassungen. Wer materialgestützt argumentiert, schreibt Texte mit höherem Grad an Transformationen von Informationen, und wer zusammenfasst, erreicht bessere Resultate in Verstehenstests. Damit wirken die beiden Zieltextsorte komplementär. Dieses Ergebnis verdient in Zukunft eine genauere meta-analytische Betrachtung.



Aufgabeneffekte wurden in diesem Teilkapitel in einem Bereich geprüft, der für das digitale Lesen besonders ertragreich – und typisch (Barzilai et al., 2018; Bräten et al., 2020; Primor & Katzir, 2018) – wirkt: dem Lesen von multiplen Dokumenten (analog oder digital), um damit eine Aufgabe zu lösen, in diesem Fall: das Schreiben eines eigenen Textes. Die Ergebnisse aus 18 Experimenten stammen aus Produktdaten. Trotz dieser Einschränkung wirken Rückschlüsse auf plausibel zugrundeliegende Transformationsprozesse möglich, welche für den Umgang mit Informationen beim materialgestützten Schreiben benötigt werden, nämlich das Auswählen, Organisieren (Dokumentenrelationen erkennen) und das inferenzbasierte Verknüpfen von Informationen (Spivey, 1990). Als Trend aus den Befunden lässt sich extrahieren, dass Vorteile sich dort abzeichnen, a) wo Informationstransformationen nötig sind und b) wo sie kognitiv besonders anspruchsvoll sind. Das mag erklären, warum wiedergebende Beschreibungen im Licht der Empirie kaum relative Vorteile erbrachten.

Die Zusammenfassungen und Argumentationen als Zieltextsorten, die aus Sicht der Schreibdidaktik wegen ihrer Transformationsbedürftigkeit der zu verarbeitenden Informationen zu kommunikativen Zwecken als lernförderlich gelten (St. Graham et al., 2020; St. Graham & Hebert, 2011; Klein & van Dijk, 2019; Philipp, 2021d), weisen Muster darin auf, bei welchen Leistungen sie zu überlegenen Effekten führen. Wer schreibend argumentiert, wird besser darin, Dokumentenrelationen schriftlich zu klären, im eigenen Text Schlussfolgerungen zu ziehen und die gelesenen Konzepte anzuwenden. Wer schriftlich zusammenfasst, kann insbesondere bei rezeptiven Tests mit höheren Leistungen in der Auswahl und kognitiven Strukturierung bessere Resultate erzielen und die Informationen auch besser im Sinne einer konzeptuellen Transformation verwenden. Dies sind Muster, die verdeutlichen, dass es sich lohnt, genauer zwischen expressiven und rezeptiven Tests, dem Konstrukt, das sie genau erfassen, und den Effekten der Zieltextsorte zu differenzieren. Vergleichende Studien mit Fokus auf die entsprechenden Prozesse, die während der Aufgabenbearbeitung von den Testpersonen absolviert werden, werden dringend benötigt.

2.2.2.5 Spezifische Hinweise zum Lesen 1: Elaborationsfragen – die Frage nach dem Warum – als tiefenorientierte Relevanzhinweise im Besonderen und tiefenorientierte Fragen im Allgemeinen

In diesem und im folgenden ([2.2.2.6](#)) Teilkapitel steht das Lesen im Fokus, bei dem Fragen die Basis des Leseauftrags bilden. Fragen dienen als metakognitive Prompts dazu, Textinhalte bzw. das Gedächtnis dahingehend zu prüfen, um eine Antwort zu geben (Andre, 1979). Cerdán, Gil und Vidal-Abarca (2011) sowie Graesser und Lehman (2011) nennen ein solches Lesen „fragengesteuertes Lesen“ und halten es wegen der im Vergleich zum zielbezogenen Lesen im Sinne der Teilkapitel [2.2.2.2–2.2.2.4](#) längeren Forschungstradition und vor allem wegen des höheren Spezifikationsgrades für ein lohnenswertes Forschungsfeld, in dem Top-down-Leseprozesse den Gegenstand bilden. Von besonderem Interesse ist in diesem weiten Forschungsfeld, a) wie viele Informationen dazu nötig sind, textbezogene Fragen zu beantworten, b) welche Verarbeitungstiefe für die Antwort nötig ist und c) welche Wissensbestände und welche Prozesse hierfür interagieren müssen (Cerdán, Gil & Vidal-Abarca, 2011; Graesser & Lehman, 2011; Hamaker, 1986). Cerdán, Gil und Vidal-Abarca (2011) bündeln diese Indikatoren, indem sie zwei Gruppen von Fragen postulieren, nämlich Fragen mit Fokus auf höhere Ebenen des Texts bzw. des Verstehens und als Gegenstück dazu solche mit Bezug auf niedrigere Ebenen. Graesser und Lehman (2011) folgen dieser Dichotomie, nennen die Fragenkategorien allerdings anders (tiefen- und oberflächenbezogene Fragen) und fügen noch eine dazwischenliegende Kategorie hinzu – sie modellieren damit also eher ein Kontinuum als eine kategoriale Dichotomie.

Folgt man der Differenzierung von oberflächenorientierten Fragen und eher auf tiefere Verstehensleistungen abzielenden Fragen, dann lassen sie sich dadurch abgrenzen, dass *tiefenorientierte Fragen* sich durch das Verstehen, Verändern und Verknüpfen von relevanten Inhalten über komplexe Inferenzen auszeichnen. Demgegenüber sind *oberflächenorientierte Fragen* eher auf singuläre Informationen bezogen, welche in nur geringem Maße oder gar nicht auf das Bilden von (Kausal-)Inferenzen angewiesen sind, um die jeweilige Frage zu beantworten (Cerdán, Gil & Vidal-Abarca, 2011).

Diese taxonomischen Überlegungen, die stellvertretend für ein gesamtes Feld der Leseforschung zu Effekten von Fragen auf Verstehensleistungen stehen (Graesser et al., 2010), können zugleich als Orientierungsrahmen helfen, die heterogenen Studien zu Relevanzeffekten von Fragen zu kartieren und zu systematisieren. Ein Zweig der Leseforschung beschäftigt sich beispielsweise mit Elaborationsfragen. Mit *Elaborationsfragen*, einem Ausdruck, der von Pressley et al. (1988) geprägt wurde, sind solche Fragen gemeint, deren Antworten das Begründen erfordern, also das Explizieren mutmaßlich kausaler Zusammenhänge. Deshalb ist es auch nicht verwunderlich, dass das Fragewort „Warum?“ häufig Bestandteil der Fragen ist, beispielsweise in der hier illustrierenden Frage „Warum sind Elaborationsfragen tiefenorientierte Fragen?“. Elaborationsfragen mit ihrem Fokus auf Kausalinferenzen können demnach als Ausschnitt der tiefenorientierten Fragen verstanden werden (Graesser et al., 2010). Für die Zwecke dieses Teilkapitels sind jedoch nicht nur Elaborationsfragen, sondern auch andere Fragen von Belang, die nicht nur den klassischen Elaborationsfragen in der Taxonomie von McCrudden und Schraw (2007) zuzuordnen sind. Dadurch soll der Blick etwas geweitet werden, da nicht nur Elaborationsfragen empirisch erforscht wurden.

Produkte

In diesem Teil des Kapitels werden zwei Arten von Produkten als Indikator von Relevanzeffekten in Studien mit tiefenorientierten Fragen präsentiert: zum einen Erinnerungsleistungen und zum anderen Resultate aus rezeptiven und expressiven Leseverstehenstests.

1) Erinnerungsleistungen

Experimentelle Studien, die den Erfolg des Lesens daran festmachen, in (meist offenen) Aufgaben möglichst viele korrekte Fakten wiederzugeben, kommen zu uneinheitlichen Befunden: Studien finden sowohl bessere Erinnerungsleistungen an Fakten aus gelesenen Texten (Boudreau et al., 1999; Halpain et al., 1985, Experimente 4 und 5; León, Moreno, Escudero, Olmos et al., 2019; Ozgungor & Guthrie, 2004; Seifert, 1993; Smith et al., 2010; Woloshyn et al., 1990, Experimente

1 und 2; Willoughby et al., 1994), allerdings auch Nulleffekte (Cerdán & Vidal-Abarca, 2008; Dornisch & Sperling, 2006; Greene et al., 1996; Kaakinen et al., 2015; McDaniel & Donnelly, 1996, Experiment 2; O'Reilly et al., 1998; Wilhite, 1985) oder sogar schlechtere Ergebnisse der Personen, die tiefenorientierte Fragen als Aufgaben lösten (Jiang & Elen, 2011).

Es gibt demnach kein klares Bild, dass sich tiefenorientierte Fragen konsistent in besseren Erinnerungen an das Gelesene niederschlagen. Erklärbar wird das darüber, dass tiefenorientierte Fragen und Erinnerungen an isolierte Fakten auf unterschiedlichen Ebenen des Lesens lokalisierbar sind (Kintsch, 2018): Tiefenorientierte Fragen mit ihrem Fokus auf Inferenzen setzen aus Sicht der Repräsentationsebenen des Lesens höher an als die an der Textoberfläche befindlichen Inhalte, die in den Tests erfasst werden. Dennoch macht sich in einem Teil der Studien bereits dort ein positiver Ertrag bemerkbar.

2) *Leseverstehenstests*

Eine zweite Gruppe von Leistungstests verschrieb sich in Experimentalstudien Verstehensleistungen höherer Ordnung, also einer tieferen kognitiven Durchdringung des Gelesenen. Diese Tests ergaben verschiedene Muster:

- *Bei den Leseverstehenstests lassen sich partiell Unterschiede darin feststellen, worauf die Lesetests abzielten: Tiefenorientierte Fragen führen zu höheren Verstehensleistungen.* Die oberflächenorientierten Fragen beantworten Personen besser als die aufgrund der erforderlichen Inferenzleistungen per se schwierigeren Fragen (Cerdán et al., 2009, 2021). Doch noch etwas ist auffällig: In der Leseforschung werden verschiedene Repräsentationsebenen von gelesenen Inhalten angenommen, darunter wortwörtliche, propositionale und ganzheitliche Repräsentationen, letztgenannte als mentales Modell bezeichnet (Kintsch, 2018). Die Ergebnisse spiegeln wider, dass Elaborations- und tiefenorientierte Fragen sich tatsächlich bei Fragen positiv auswirken, die Kognitionen auf höheren Ebenen des Verstehens voraussetzen. Erforderte der Test nämlich derartige Leistungen, etwa Inferenzen zwischen Informationen (McDaniel & Donnelly, 1996, Experiment 2; Ozgungor & Guthrie, 2004; Seifert, 1993; 1994), Inferenzen, die auf ein Situationsmodell verweisen (Cerdán et al., 2009; Seifert, 1993), Anwendung des Wissens aus Texten auf neue Situationen (Cerdán & Vidal-Abarca, 2008; Jiang & Elen, 2011) oder intertextuelles Leseverstehen (Cerdán et al., 2021), dann erbrachten die Testpersonen mit tiefenorientierten Fragen bessere Ergebnisse.
- *Personen, die tiefenorientierte Fragen beantworten, erinnern sich nicht nur besser an Textinhalte, sondern auch deren Struktur.* Wer tiefenorientiert Texte liest, ist gemäß Rouet et al. (2001, Experiment 2) besser darin, Sätze korrekt ihrem Ursprung im Text zu memorieren, indem sie randomisiert präsentierte

Sätze nach dem Lesen einem Inhaltsverzeichnis zuordneten. Auch sind Personen besser darin, Fakten übergeordneten Kategorien zuzuordnen (Woloshyn et al., 1990, Experimente 1 und 2).

- *Offene Aufgaben enthalten Spuren der Art der Frage.* In offenen Tests zeigen sich in Studien Spuren dahin gehend, dass beispielsweise Zusammenfassungen zu gelesenen Texten mehr jener Informationen enthalten, die mit der jeweiligen Frage zu tun hatten – und zwar unabhängig davon, ob das inhaltlicher Gegenstand des Tests sein sollte (Rouet et al., 2001, Experimente 1 und 2). Etwas Ähnliches ließ sich in einer anderen Studie feststellen, in der Zusammenfassungen deutlich mehr Inhalte aus Absätzen enthielten, die für die Beantwortung der Frage mit einem spezifischen Relevanzhinweis relevant waren (León, Moreno, Escudero & Kaakinen, 2019).

Im Gesamt sind demnach in mehreren Studien stärker ausgeprägte Fähigkeiten beobachtet worden, die bei Personen mit Elaborations- bzw. anderen tiefenorientierten Fragen zu stärker vernetzten und umfassenderen Repräsentationen des Gelesenen führten.

Prozesse

Die Ergebnisse aus den Experimentalstudien hinsichtlich der Prozesse lassen sich drei Bereichen zuschlagen: 1) den noch wenig untersuchten Augenbewegungen, 2) den mit Lesen einzelner Textteile und teils der Fragen zugebrachten Zeiten und 3) den aus den Navigationsverhalten ableitbaren Vorgehensweisen bei der Auswahl und kognitiven Verknüpfung von Informationen.

1) *Eye-Tracking-Indikatoren*

Eye-Tracking-Studien zu Elaborationsfragen sind noch rar. Zwei Studien aus Spanien seien herbeigezogen (León, Moreno, Escudero, Olmos et al., 2019; Moreno et al., 2021). Beide Studien eint, dass sich die Experimentalgruppen darin unterschieden, dass sie auf eine Frage, die zum Teil in den Texten im ersten Absatz dezidiert enthalten waren, entweder explizit aufmerksam gemacht worden waren (spezifischer Relevanzauftrag) oder nicht (allgemeiner Relevanzauftrag). Diese Personen fassten Texte mündlich zusammen, und diese Zusammenfassungen sollten die Elaborationsfrage beantworten. Es handelt sich also um ein spezielles Design. Im Gesamt machen die Studien darauf aufmerksam, dass Elaborationsfragen – in diesem Falle: eingebettete Fragen, die Teil des Textes waren – erst unter spezifischen Bedingungen und bei spezifischen Personengruppen mit einem spezifischen Verhalten korrespondieren:

- *Eingebettete Elaborationsfragen werden prozesswirksam bei explizitem Hinweis auf diese Fragen:* Wenn man Personen allgemein auf vorhandene Elaborationsfragen vor dem Bearbeiten der Aufgaben hinweist, fixieren die Leserinnen und Leser relevante Textteile (in diesem Fall: Absätze) häufiger und länger und sie lesen den Textteil mit der Elaborationsfrage häufiger erneut. Die längeren Lesezeiten relevanter Textteile lassen sich sowohl beim ersten Lesen als auch bei wiederholten Lektüren dieser Absätze feststellen.
- *Bedingungen dieses Effektes – spezifischer expliziter Hinweis:* Dieser Effekt war daran gebunden, dass die Elaborationsfrage im Text vorhanden war und dass die Testpersonen auf diese eingebettete Frage explizit hingewiesen worden waren. Dies zeigte insbesondere die Studie von Moreno et al. (2021).

Während in den bisherigen Studien zwei Arbeitsaufträge miteinander verglichen wurden, ist eine weitere Untersuchung (León, Moreno, Escudero & Kaakinen, 2019) für die Frage erhellend, ob sich der Erfolg beim Beantworten von Elaborationsfragen mit differenziellen Vorgehensweisen in Leistungsgruppen bei einer Aufgabe in Verbindung bringen lässt. Im nachstehenden Kasten „Illustrierendes Beispiel 8“ wird deutlich, dass das der Fall ist. In dieser Studie bearbeiteten die Personen den spezifischen Relevanzauftrag, sie wurden also explizit darauf hingewiesen, auf die eingebettete Elaborationsfrage am Ende des ersten Absatzes zu achten.

Illustrierendes Beispiel 8

Erfolg liegt im Auge des Betrachters – was in puncto Textzusammenfassungen besonders leistungsstarke Personen ins Auge fassen

A) Was war das Erkenntnisinteresse?

Im Experiment wurde der Frage nachgegangen, ob sich Leistungsgruppen im mündlichen Zusammenfassen in ihren Blickbewegungen unterschieden, wenn sie für die Beantwortung einer Elaborationsfrage Texte lasen (León, Moreno, Escudero & Kaakinen, 2019).

B) Wer wurde untersucht?

Bei den Testpersonen handelte es sich um 40 spanische Studierende des Fachs Psychologie im dritten Studienjahr. Sie waren zwischen 20 und 23 Jahre alt, 70 Prozent waren Frauen. Alle Personen sprachen Spanisch als Muttersprache.

C) Welche Instrumente kamen zum Einsatz?

Die Studierenden lasen sechs kurze Sachtexte von 200 bis 250 Wörtern Länge. Jeder Text enthielt drei Absätze. Der jeweils erste Absatz enthielt eine Elaborationsfrage wie „Aber wie konnte der Fluss so verseucht werden?“. Die beiden Folgeabsätze enthielten entweder relevante Informationen auf die Frage (in diesem Fall zwei Gründe für die Verschmutzung der Themse) bzw. irrelevante Informationen (hier: die Folgen der Flussverschmutzung).

Die Studierenden erhielten den Auftrag, dass sie nach dem Lesen den jeweils gelesenen Text mündlich zusammenfassen sollten und insbesondere die Frage aus dem Einleitungsabsatz beantworten sollten. Das Leseverhalten wurde per Eye-Tracker aufgezeichnet und dabei neben den Fixationszeiten verschiedener Absätze auch die Zahl der wiederholten Konsultationen von Absätzen erfasst.

Um die Gruppen zu bilden, wurde die gesamte Stichprobe gedrittelt; Grundlage dafür waren die Leistungen in den mündlichen Zusammenfassungen. Besonders wichtig für die Analysen war das leistungsstärkste Drittel. Diese Personengruppe stellte sich in den Analysen im Vergleich mit den anderen beiden Gruppen als deutlich anders heraus.

D) Was waren die Hauptbefunde?

Ein markantes Muster ließ sich dahingehend extrahieren, dass sich das leistungsstärkste Drittel von den anderen beiden Gruppen deutlich unterschied, während sich das Mittelfeld und das leistungsschwächste Drittel in ihren Eye-Tracking-Maßen ähnelten. Zusammengefasst gibt es fünf Merkmale des leistungsstärksten Drittels. Diese Personen lasen a) die relevanten Absätze länger, wobei sie das b) beim erstmaligen Lesen ebenso taten wie c) beim wiederholten Lesen. Außerdem kehrten leistungsstärkere Personen d) vom relevanten Absatz häufiger zum ersten Absatz zurück, welcher die Frage enthielt, und e) lasen diesen Absatz ebenfalls länger erneut.

E) Diskussion

Die Autorinnen und Autoren der Studie werten das Befundmuster, das sich über alle fünf fokussierten Prozessmaße des Eye-Trackings bei sechs gelesenen Texten erstreckte, als Indikator für eine Strategie selektiver Verarbeitung unter leistungsstarken Leserinnen und Lesern im obersten Leistungsdrittel. Dabei fällt insbesondere der Abgleich vom Ort mit der Elaborationsfrage und dem Absatz auf, der die nötigen Informationen enthält, – nebst dem Umstand, dass relevante Informationen länger fokussiert wurden. Dieses Befundmuster ist ein wichtiges Ergebnis. Ein weiteres besteht darin, dass sich keine anderen statistisch auffälligen Differenzen in den Prozessdaten der anderen beiden Personengruppen verbargen, dass sich also die in ihren später erbrachten Leistungen unterscheidenden Personengruppen in ihren per Eye-Tracking erfassten Prozessen ähnelten.

2) Bearbeitungs- und Reaktionszeiten

Verschiedene Experimentalstudien erfassten Lesezeiten – sei es bei Fragen, sei es bei den Texten, seien es Reaktionszeiten beim Bearbeiten der Leseaufträge. Die Resultate eint, dass tiefenorientierte Fragen nicht nur mit längeren Lesezeiten bei den Texten, sondern bereits bei den Fragen einhergehen:

- *Personen, die tiefenorientierte Fragen bearbeiten, lesen diese Fragen länger.* Eine längere Lesezeit gilt als Indikator für ein vertieftes Verarbeiten der Frageanforderungen. Dieses Ergebnis ist von mehreren Studien belegt worden

(Cerdán et al., 2008, Experimente 1 und 2; Cerdán et al., 2009; Rouet, 2003). Hierzu passt, dass Personen die Fragen häufiger lesen (Cerdán et al., 2008, Experimente 1 und 2; Rouet, 2003). Es wirkt damit so, als würden tiefenorientierte Fragen mehr Aufmerksamkeit auf sich ziehen, was auch Reaktionszeitexperimente verdeutlichen (Halpain et al., 1985, Experimente 3 und 4). Längeres Lesen der Fragen und häufigeres Lesen der Fragen stehen in positivem Zusammenhang mit den Erinnerungsleistungen an gelesene Inhalte (Rouet, 2003) und Transferleistungen (Jiang & Elen, 2011).

- *Die Textlesezeit ist in einigen Studien bei tiefenorientierten Fragen länger als jene bei oberflächenorientierten Fragen.* Dieses Muster wurde in mehreren Studien vorgefunden und bezieht sich hierbei auf die Zeit mit dem Lesen relevanter Inhalte (Cerdán et al., 2008, Experiment 1; Cerdán & Vidal-Abarca, 2008). Etwas kontraintuitiv mutet der Befund an, dass Personen, die tiefenorientierte Fragen beantworteten, zum Teil weniger Zeit mit dem Lesen von Texten verbracht haben, was sich in den entsprechenden Studien aber auch über die Anzahl zu bearbeitender Fragen erklären lässt (Cerdán et al., 2009). Entscheidend für die Erklärung dieses Befundes in der Studie von Cerdán et al. (2009) dürfte gewesen sein, dass die Lesezeiten sich auf globale Texte oder auf spezifische Textteile bezogen haben.
- *Wer tiefenorientierte Fragen beantworten soll, ist von dieser Aufgabe kognitiv stärker absorbiert als von Aufgaben mit oberflächenorientiertem Fokus.* Dies zeigte ein Experiment mit Reaktionszeittests, bei denen die Probandinnen und Probanden auf akustische Signale reagieren sollten. Die Reaktionszeit unterschied sich hierbei deutlich: Sie war länger bei tiefenorientierten Fragen, was als Indikator für eine höhere kognitive Beanspruchung gewertet wird (Halpain et al., 1985, Experiment 1).

3) Informationsauswahl und -verknüpfung

Einer der Vorteile, wenn man digitales Lesen erforscht, besteht unter anderem darin, dass man mithilfe der Software Daten erfassen kann, die sich direkt und non-reaktiv aus dem Nutzungs- und Navigationsverhalten der Probandinnen und Probanden ergeben. Diese wurden in den Experimenten vor allem daraufhin untersucht, ob sich Effekte bei den für die Fragen relevanten Textteilen eingestellt haben. Das war tatsächlich der Fall, wie es die hier gebündelten Resultate als Muster illustrieren:

- *Personen, die tiefenorientierte Fragen beantworten, lesen mehr Text.* Das bedeutet, sie konsultieren eine größere Textmenge, insbesondere lasen sie mehr von den Textpassagen, die für die Beantwortung von Fragen nötig waren (Cerdán et al., 2008, Experimente 1 und 2; Rouet et al., 2001, Experimente 1).

- *Es gibt verschiedene Muster bei der Lesezeit je nach Auftrag.* Dieses Resultat wurde von Cerdán und Vidal-Abarca (2008) beigesteuert. Sie zeigten, dass tiefenorientierte Fragen eher ein Verhalten auslösten, bei dem die Testpersonen Absätzen mit relevanten Informationen unterschiedlich viel Zeit widmeten: Waren die Informationen für die Antworten auf die Fragen im Text verteilt, lasen die Personen diese Absätze länger, während sie dazwischenliegende Absätze nur kurz anvisierten. Dies nennen Cerdán und Vidal-Abarca (2008) „integrierendes Verarbeiten“. Demgegenüber wandten sich Personen mit dem Suchverhalten nach einzelnen Informationen eher selektiv nur den entsprechenden relevanten Informationen zu (= „Einzeleinheiten verarbeiten“). Diese beiden Vorgehensweisen erwiesen sich zusätzlich als Prädiktoren für die Leistungen, wenn auch unterschiedlich: Ein häufigeres Verarbeiten von Einzeleinheiten ging mit höheren Nennungen von Informationen in einem Test einher, aber mit schlechteren Leistungen in einem Inferenztest. Das integrierende Verarbeiten sagte positiv das Resultat in einem Transfertest voraus. Demnach lösen tiefenorientierte Aufgaben Vorgehensweisen stärker aus, die wiederum zu besseren kognitiv anspruchsvollen Leistungen führen.
- *Wer tiefenorientierte Fragen beantwortet, neigt mehr zu einem iterativen Vorgehen.* Damit ist gemeint, dass Personen stärker dazu tendieren, zwischen dem Lesen von Fragen, dem Text und dem Antworten zu wechseln. Jiang und Elen (2011) berichten beispielsweise davon, dass Personen ohne Fragen den Text linear lasen, während solche mit tiefenorientierten Fragen mehrfache Lektüren von Textteilen vornahmen. Cerdán et al. (2008, Experiment 2; 2009) fanden ebenfalls eine deutlich höhere Anzahl von Zyklen vor, die sich aus dem Wechsel zwischen verschiedenen Aktivitäten ergeben: Es wirkt demnach so, als würden die Leserinnen und Leser ein häufigeres Aufsuchen und Verknüpfen von aufgabenbezogenen Erfordernissen und den zu verbindenden Informationen an den Tag legen. Demgegenüber zeichnen sich Personen mit dem Ziel, oberflächenorientierte Fragen zu beantworten, dadurch aus, dass sie Textteile so lange überfliegen, bis sie zu der gesuchten Informationen gelangt sind, die sie dann länger fokussieren (Rouet et al., 2001, Experimente 1 und 2). Die beiden Gruppen erfolgreicher Vorgehensweisen lassen sich bündeln als a) eine Suchen-und-Memorieren-Strategie bei oberflächenorientierten Fragen bzw. b) eine Überprüfen-und-Kombinieren-Strategie bei tiefenorientierten Fragen (s. dazu auch den Kasten „Illustrierendes Beispiel 9“).
- *Leistungsstärkere und -schwächere Personen unterscheiden sich in ihrem Vorgehen.* Die Gruppenunterschiede im Sinne, ob Personen bei den jeweiligen Aufgaben zur leistungsstärkeren oder -schwächeren Hälfte gehören, sind noch nicht häufig untersucht worden. Wo das der Fall war (Cerdán et al., 2008, Experiment 1), ergab sich das Muster, dass leistungsstarke Personen sich länger relevanten Textpassagen lesend widmeten und diese auch zu einem größeren Anteil rezipieren. Sie lasen also das Richtige und dies obendrein

länger. Dies scheint allerdings davon abzuhängen, ob die Fragen im Vorfeld gestellt wurden oder erst nach dem Lesen (Cerdán et al., 2008, Experiment 2). Jedoch zeichnen sich Muster dahingehend ab, dass leistungstärkere Personen insgesamt zielführendere Varianten bei der Konsultation der relevanten Textteile aufwiesen: Sie verbrachten weniger Zeit mit Suchen und nutzten die relevanten Textteile länger – ein sowohl effizientes wie auch effektives Vorgehen (Cerdán et al., 2008, Experimente 1 und 2). Jiang und Elen (2011) steuern einen etwas anders gelagerten Befund bei: Bei ihnen ließen sich zwei Gruppen von Personen unterscheiden. Eine Gruppe las den Text nur einmal, wechselte häufig zwischen den Absätzen und versuchte dabei partiell, Informationen aus dem Text für die Antworten zu kopieren. Die zweite und in ihren Leistungen erfolgreichere Gruppe las den gesamten Text hingegen häufiger. Da sich verschiedene Leistungsgruppen offenbar in ihrem Vorgehen unterscheiden lassen, gibt der Kasten „Illustrierendes Beispiel 9“ nochmals vertiefte Einblicke in eine der Studien.

Illustrierendes Beispiel 9

Unterschiedliche Fragen, variierende Vorgehensweisen, divergierender Erfolg – eine Rekonstruktion prototypischer funktionaler und dysfunktionaler Prozesse

A) Was war das Erkenntnisinteresse?

Das Erkenntnisinteresse bestand in dieser Studie darin, aufgrund von Navigationsverhalten, Antwortinhalten und der Analyse von für die Antwort nötigen Textteilen Aussagen über günstige und ungünstige Vorgehensweisen zu treffen (Cerdán, Gil & Vidal-Abarca, 2011). Hierfür wurde das Vorgehen bei oberflächen- und bei tiefenorientierten Fragen rekonstruiert.

B) Wer wurde untersucht?

Bei der Studie handelt es sich um eine Reanalyse von Daten aus einem Experiment mit 37 spanischen Psychologie-Studierenden (Cerdán et al., 2009). Diese waren durchschnittlich 20 Jahre alt, rund die Hälfte war männlich. Für die erneute Analyse wurden aber nicht die Daten von einzelnen Individuen oder Gruppen fokussiert, sondern insgesamt sieben Fragen aus dem Experiment ausgewählt, die besonders leicht waren (und von allen Personen gelöst worden waren) oder von niemandem.

C) Welche Instrumente kamen zum Einsatz?

Das ursprüngliche Instrumentarium bestand darin, dass in einer speziellen Software sowohl die Fragen als auch die Texte direkt von den Testpersonen ausgewählt werden. Im Fall des Textes ist jeweils nur ein Absatz klar lesbar, alle anderen Textteile erscheinen unscharf. Dadurch war es möglich, die Lesezeiten und die Reihenfolgen der gelesenen Textteile und Aktivitäten (Lesen, Antwort schreiben) zu erfassen, woraus sich wiederum über das aufgezeichnete Verhalten das lesestrategische Vorgehen partiell rekonstruieren

ließ. Der zu lesende Text war mit über 1.700 Wörtern und 31 Absätzen recht lang. Von den 31 Absätzen enthielten 16 Informationen, die man entweder bei den oberflächenorientierten Fragen oder für die tiefenorientierten Fragen zur Antwort benötigte.

Zusätzlich zu diesen Prozessvariablen wurden Produktmaße erfasst, in diesem Fall die richtigen Antworten, welche die Personen zu den oberflächen- bzw. tiefenorientierten Fragen gegeben hatten. Aus dem Fragenpool wurden vier bzw. drei Fragen mit extremen Antworttendenzen gewählt. Pro Frage wurden zwei Fälle mit vollständig korrekter bzw. vollständig inkorrektur Antwort herausgegriffen. Abgeglichen wurden dann der Antwortinhalt mit den Inhalten der für die jeweiligen Fragen relevanten Textteile und das Navigationsverhalten der Probandinnen und Probanden. Auf dieser Basis ließen sich für das Antwortverhalten zielführende und ungünstige Vorgehensweisen nachträglich beschreiben.

D) Was waren die Hauptbefunde?

Die Befunde bestehen in einer Taxonomie von drei bzw. vier Strategien bei der mehr oder weniger erfolgreichen Beantwortung von oberflächen- bzw. tiefenorientierter Fragen, die in Abbildung 4 und in Abbildung 5 dargestellt sind. Beide Varianten eint, dass nach dem Lesen der Fragen eine Suche im Text stattfindet, diese führt je nach gewähltem Weg zu differenziellen Effekten.

Ein günstiges Vorgehen im Falle der *oberflächenorientierten Fragen* besteht darin, nach der gezielten Suche nach relevanten Informationen entweder die Inhalte verbatim zu übernehmen (Copy-and-Paste-Strategie) oder mit eigenem Vorwissen anzureichern (Elaborationsstrategie auf niedriger Ebene). Beide Varianten kamen gleich häufig vor und zeichneten sich zudem dadurch aus, dass die Personen nach Beantworten der Fragen keine weiteren Suchvorgänge mehr durchführten, sondern ihre Antwort gaben, nachdem sie die Passage gefunden hatten.

Die wenig erfolgreichen Personen hatten unterschiedliche Vorgehensweisen: Eine Hälfte hatte zwar die richtige Passage gefunden, nutzte dann aber falsche Informationen, sie wählte Bedeutungsloses in puncto der gestellten Frage aus. Die andere Hälfte ging oberflächlich vor, sie schien sich auf das Wiederfinden von Wörtern aus der Frage zu stützen, wählte aber völlig irrelevante Aussagen aus.

Bei den *tiefenorientierten Fragen* führte nur eine Strategie zum Erfolg: das Aufsuchen aller relevanten Absätze und das inferenzbasierte Integrieren dieser Informationen, weshalb diese Strategie als „Prüfen und Integrieren“ bezeichnet wurde. Dass Personen an den Fragen scheiterten, hing zunächst damit zusammen, dass sie sich nicht auf alle, sondern nur auf einige relevante Absätze fixiert hatten. Dabei paraphrasierten sie in der Antwort entweder die Inhalte (in fünf Sechsteln der Fälle) oder übernahmen die Inhalte sogar wortwörtlich (ein Sechstel der Fälle).

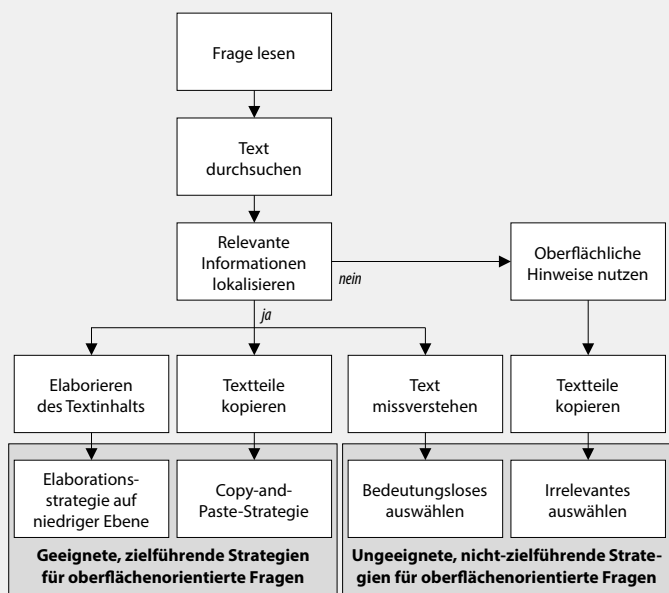


Abbildung 4: Antwortstrategien für das Beantworten von oberflächenorientierten Fragen (Quelle: modifizierte Darstellung von Cerdán, Gil & Vidal-Abarca, 2011, S. 305)

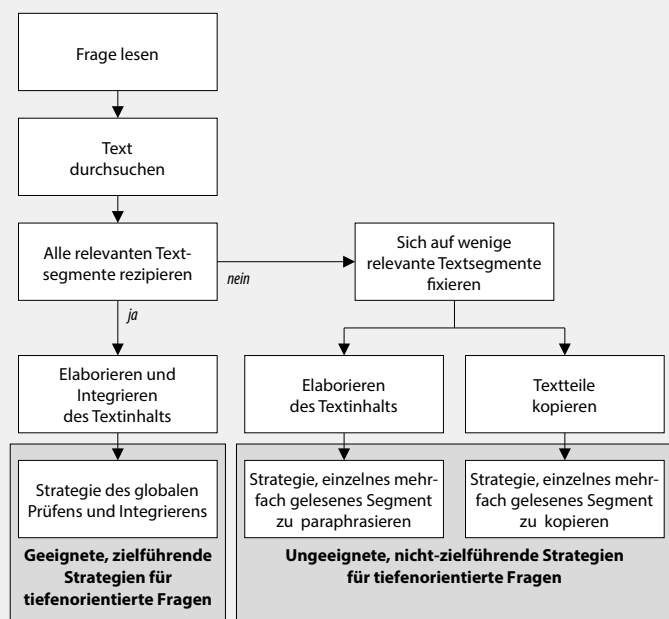


Abbildung 5: Antwortstrategien für das Beantworten von tiefenorientierten Fragen (Quelle: modifizierte Darstellung von Cerdán, Gil & Vidal-Abarca, 2011, S. 307)

E) Diskussion

Mit dem Extremgruppenvergleich und der Rekonstruktion des Vorgehens anhand des Navigationsverhaltens liegt das Verdienst der Studie vor allem darin, dass die Fehlanwendungen genauer herausgearbeitet werden konnten. Dies lässt sich so benennen: Oberflächliches Bearbeiten – wobei sich oberflächlich hierbei auf den Umgang mit dem Text in puncto Fragerfordernisse bezieht – führt zu ungenügenden Antworten. Dabei heißt oberflächlich: a) entweder Wörter aus der Frage im Text zu suchen und Antworten kopieren, b) aus tendenziell relevanten Textteilen das Falsche auszuwählen oder c) unvollständig zu lesen. Demgegenüber sind zielführende Vorgehensweisen analytisch: Sie weisen nicht nur auf die Passung, sondern – im Falle der tiefenorientierten Fragen – auch auf eine Sensibilisierung hinsichtlich der Vollständigkeit der Informationen hin.

Und auf noch etwas macht der Vergleich aufmerksam: Während das Vorgehen, das beim Beantworten von oberflächenorientierten Fragen noch als zielführend eingeschätzt werden kann, dort für den Erfolg ausreicht, zählt es bei den tiefenorientierten Fragen zum dysfunktionalen Agieren. Anders gesagt: Was bei oberflächenorientiertem Lesen noch genügt – punktuelle Übernahmen von textuellen Informationen in die eigene Antwort –, reicht beim tiefenorientierten Lesen nicht mehr.

Zusammenfassung zu spezifischen Relevanzeffekten tiefenorientierter (Elaborations-)Fragen auf Produkte und Prozesse des Lesens

Der selektive Forschungsüberblick zu Effekten von tiefenorientierten Fragen als Ausschnitt relevanzbezogener Aufgaben hat teils deutliche Folgen solcher Fragen herausgearbeitet. Diese finden sich bei den *Produktdaten* vor allem dort, wo es zu einer Passung von kognitiv zu erbringender Leistung bei der Aufgabenbearbeitung und späteren Tests kam: Leistungsindikatoren mit Fokus auf Zusammenhänge und Strukturen einzelner und teils multipler Texte waren nachweisbar. Demgegenüber ist eine Überlegenheit bei Erinnerungsleistungen an singuläre Fakten weniger deutlich ersichtlich.

Bei den *Prozessen* ist auf den allgemeinsten Nenner gebracht eine deutliche und selektive Zuwendung der Aufmerksamkeit auf relevante Textteile erkennbar. Sie zeigt sich in längeren Fixationen, mehrfachen Konsultationen und einem Abgleich von Textteilen mit Fragen, wobei auch letztgenannte mehr Zuwendung erfahren. Ein prüfendes, auf Inferenzen abzielendes und einen Abgleich mit den Frageerfordernissen vornehmendes Vorgehen ist somit ein evidentes Gesamtmuster.

2.2.2.6 Spezifische Hinweise zum Lesen 2: spezifische Segmente

Mit den Hinweisen auf spezifische Segmente sind Hinweise oder Fragen gemeint, die eine lesende Person dazu auffordern, ihre Aufmerksamkeit auf definierte und in der Regel kleinteilige Abschnitte des Textes zu richten, um

oberflächenorientierte Fragen zu beantworten – etwa: „Welche Merkmale haben die Abschnitte, auf die Personen ihre Aufmerksamkeit bei spezifischen Segmenten richten sollen?“. Dazu werden entweder vor dem Lesen des Texts Fragen präsentiert oder sie werden in den Text an die entsprechenden Stellen eingefügt (Hamilton, 1985; McCrudden & Schraw, 2007 – s. für eine frühe Metaanalyse zu Effekten solcher Fragen: Hamaker, 1986). Teils werden solche Fragen auch im Nachhinein gestellt. Hamaker (1986) hat hierbei darauf aufmerksam gemacht, dass solche Positionierungen vor oder nach dem Lesen des Textes mit teils differenziellen Effekten bei den Leistungen einhergehen, z. B. was Erinnerungsleistungen angeht, die sich auf Informationen beziehen, welche nicht direkt von den Fragen adressiert werden. Diese werden eher unterdrückt, wenn man die Fragen vorgängig erhält.

Produkte

Die Produkte des Lesens sind dominant mit Erinnerungsleistungen und nur wenig mit tieferen Leseverstehensleistungen in Tests operationalisiert worden. Dies schlägt sich in der Ergebnisdarstellung entsprechend nieder.

1) Erinnerungsleistungen

Die Erinnerungsleistungen divergieren, was sich in den hier besonders fokussierten Studien mit Fragen vor dem Lesen manifestiert: Positive Effekte treten bei einem kongruenten Verhältnis von Frage und relevanter Information auf. Das lässt sich an zwei komplementären Befundmustern illustrieren:

- *Personen, die spezifische Fragen zu den Texten erhalten hatten, erinnerten sich nach dem Lesen besser an solche Informationen, auf die sie achten sollten – allerdings mit Limitationen.* Wer seine Aufmerksamkeit auf spezifische Informationen lenken sollte, erinnert sich tatsächlich besser an diese relevanten Informationen (Britton et al., 1985; Duchastel & Brown, 1974; McCrudden et al., 2005, Experimente 1 und 2; Lapan & Reynolds, 1994; Reynolds et al., 1979; Reynolds & Anderson, 1982; studienübergreifend: Hamaker, 1986). Allerdings nimmt dieser Effekt dann ab, wenn die Anzahl der zu beachtenden Informationen stark zunimmt (Rothkopf & Billington, 1975a, 1975b). Außerdem sind Erinnerungsleistungen bei mehreren erfragten Informationen nur dann höher, wenn die relevanten Informationen dicht aufeinander folgen (Gagné & Rothkopf, 1975; Rothkopf & Koether, 1978) und wenn die eingesetzten Texte kürzer sind (Hamaker, 1986). Zudem gibt es teils nachteilige Effekte bei jüngeren Leserinnen und Lesern, die sich ohne Fragen an mehr Inhalte erinnern als mit Fragen (van den Broek, Tzeng et al., 2001).

- *Die Relevanzhinweise führen vor allem bei jenen faktualen Informationen zu erhöhten Erinnerungsleistungen, die explizit Gegenstand der Hinweise waren, aber nicht bei anderen Informationen.* Es mag trivial anmuten, aber: Der Effekt der Relevanzhinweise erstreckt sich primär auf Informationen, die von den Hinweisen inhaltlich fokussiert werden. Andere Informationen in Texten merken sich die Personen teils sogar schlechter (Duchastel & Brown, 1974; Rothkopf & Billington, 1975b; Rothkopf & Koether, 1978; van den Broek, Tzeng et al., 2001). Dieser Effekt wurde schon früh metaanalytisch von Hamaker (1986) belegt.

2) Leseverstehen

Wegen ihrer Fokussierung auf einzelne Informationen beim Verarbeiten von Texten und den dazu passenden Tests sind Leseverstehentests mit einem stärker inferenzbasierten Fokus auf Verstehensleistungen eher selten. Dennoch gibt es Hinweise, dass Personen mit dem Auftrag, sich auf einzelne Passagen zu fokussieren, tatsächlich umfassendere Verstehensleistungen erbringen, indem sie die Zusammenhänge zwischen Informationen herstellen und begründete Belege für den Zusammenhang von Inhalten geben (Lehman & Schraw, 2002, Experiment 1).

Prozesse

1) Lesezeiten

Personen verlangsamten ihr Lesen in den Textteilen, in denen die Informationen vorzufinden waren, welche für ihre jeweiligen Fragen von Relevanz waren. Sie lesen also jene Textteile länger, die Antworten enthielten, die zu ihren Fragen passen (Lapan & Reynolds, 1994; Reynolds & Anderson, 1982; Reynolds et al., 1979, 1990, Experiment 1). Teils wurden solche Effekte auch nicht gefunden, obwohl sich Personen trotzdem besser an die Inhalte erinnerten, die zu ihrem Auftrag passten (McCrudden et al., 2005, Experimente 1 und 2). Diese beiden divergierenden Befunde werden über zwei alternative Hypothesen erklärt, nach denen das Lesen relevanter Informationen entweder mehr (Increased-Effort-Hypothesis) oder nicht mehr (No-Increased-Effort-Hypothesis) kognitiven Aufwand mit sich bringt. Hierbei scheinen sich diese Effekte auch darauf zurückzuführen lassen, dass gute Leserinnen und Leser kürzere Lesezeiten aufweisen als schwächere Leserinnen und Leser (Reynolds et al., 1993).

2) Bearbeitungs- und Reaktionszeiten

Wenn Personen Texte lesen und spezifische Ziele bearbeiten sollen, wenden sie *für das Lesen allgemein eine längere Zeit* auf (Britton et al., 1985; Rothkopf & Billington, 1975a). Eine längere Lesezeit hängt mit besseren Erinnerungsleistungen zusammen (Britton et al., 1985). Studien, in denen die Reaktionszeiten erfasst wurden, verweisen auf eine stärkere Aufmerksamkeit, welche den relevanten Passagen zukommt: *Die Leserinnen und Leser reagieren langsamer in Reaktionszeitmessungen* (Britton et al., 1985; Reynolds & Anderson, 1982). Damit ergeben sich Muster einer höheren kognitiven Beanspruchung bzw. einer stärker gelenkten Aufmerksamkeit.

3) Vorgehensweisen bei der Aufgabenbearbeitung

In einer Studie (Alexander, 1986) erhielten Studierende einen langen Text und konnten auf einem Blankoblatt Notizen anfertigen und sollten danach noch einmal mündlich zu Protokoll geben, wie sie vorgegangen waren. Auf der Basis der triangulativen Analysen von schriftlichen Notizen und Verbaldaten ließ sich feststellen, dass Personen mit Hinweisen auf spezifische Textsegmente ein anderes Vorgehen wählten als jene mit einem allgemeinen Auftrag. Sie nutzten stärker lesestrategische Vorgehensweisen wie das Vorhersagen, Fragen-Stellen sowie Analogien- und Vorstellungsbildungen. Diese Strategien eint, dass die Personen das eigene Vorwissen als Ressource dafür verwenden, textuelle Informationen zu elaborieren.

Zusammenfassung zu spezifischen Hinweisen beim Lesen zur Fokussierung auf einzelne Segmente

Wenn Relevanzhinweise die Aufforderung beinhalten, sich auf einzelne Textsegmente zu fokussieren, dann geht dies bei den *Produktmaßen* grosso modo mit erhöhten Erinnerungsleistungen an die Inhalte dieser Segmente einher. Dies geschieht erwartbar zulasten anderer Informationen aus den Texten. Effekte auf Leseleistungen jenseits der Erinnerung an Inhalte sind insgesamt noch spärlich untersucht. Die gesteigerte Aufmerksamkeit, die sich durch die Hinweise ergeben, spiegeln in den groben *Prozessindikatoren* verlängerte Lese- und Reaktionszeiten wider. Zudem ließen sich – auf sehr schmaler empirischer Basis – auch stärker vorwissensbasierte Lesestrategien nachweisen.

2.3 Determinanten und interindividuelle Unterschiede

Die bisherigen Ausführungen vor allem im zweiten Teil des Teilkapitels [2.2](#) bezogen sich auf allgemeine Effekte von Aufgabenstellungen auf Prozesse und Produkte. Ungeklärt ist bis dato noch, ob es individuelle Merkmale gibt, welche die Richtung und Ausprägung von Effekten von (externen) Aufgabenstellungen beeinflussen. Die Forschung zu Relevanzeffekten hat sich primär auf individuelle Merkmale konzentriert (McCrudden & Schraw, 2007; McCrudden et al., 2011b), während sich in anderen Forschungsfeldern wie den Effekten von Fragen andere, die Fragen und das textuelle Material und nicht zuletzt die Art der Testung selbst betreffende Merkmale als potenzielle Quelle von Relevanzeffekten herauskristallisiert haben (Hamaker, 1986).

Die folgende Darstellung des Forschungsstandes konzentriert sich *rein auf individuumbezogene Merkmale*, verlässt allerdings mangels ausreichend breiter Basis an Studien die systematische und strenge Trennung von den verschiedenen Aufgabeneffekten, die im Teilkapitel [2.2.2](#) als Strukturierungshilfe genutzt wurde. Stattdessen werden übergreifende Ergebnisse präsentiert. Es handelt sich dabei um drei Bereiche, welche kognitive Merkmale der Personen betreffen und damit natürlich nur einem Ausschnitt der in- und externalen Determinanten von Leistungsunterschieden im relevanzbezogenen Lesen abbilden. Im Einzelnen sind dies: sprachliche Fähigkeiten wie Wortschatz und Lesefähigkeiten ([2.3.1](#)), Arbeitsgedächtniskapazität ([2.3.2](#)) und inhaltliches Vorwissen ([2.3.3](#)).

2.3.1 Sprachliche Fähigkeiten

2.3.1.1 Leseverstehen

Die sprachlichen Fähigkeiten der Personen sind eine Gruppe unterschiedlicher Merkmale, die in diesem Kapitel trotz aller Heterogenität gebündelt werden. Eine erste in der Forschung häufiger untersuchte Variable stellt das *Leseverstehen* dar. In den hier berichteten Studien wird dabei grundsätzlich auf zweierlei Arten dem Leseverstehen begegnet: Entweder werden nach erbrachtem Leseverstehen in separaten Tests Leistungsgruppen gebildet, deren Aufgabenergebnisse statistisch verglichen werden (so bei Cerdán, Gilabert & Vidal-Abarca, 2011; Reynolds et al., 1993; Vidal-Abarca et al., 2010, Experiment 1; Wong et al., 1982). Damit wird geprüft, ob sich verschiedene, ihrerseits einander im Lesen ähnelnde Personen-gruppen in ihren Leistungen bei den relevanzbezogenen Aufgaben unterscheiden. Alternativ dazu wird das Leseverstehen als kontinuierliche, meist metrisch erfasste Kompetenz als Kovariate berücksichtigt, um generellen statistischen Effekten des Lesens über alle Versuchspersonen hinweg nachzuspüren. Dieser Gruppe von Studien liegt die empirisch begründete Grundannahme zugrunde,

dass allgemeines Leseverstehen und Indikatoren des Erfolgs im relevanzbezogenen Lesen im Zusammenhang stehen: Wer Texte besser versteht, löst Aufgaben besser (Cerdán et al., 2019; Cerdán et al., 2021; Vidal-Abarca et al., 2010, Experiment 1). In dieser letzten Kategorie von Studien geht es darum, dass schwächere Leserinnen und Leser zwar in ihren Leistungen bei Tests zurückfallen, dass sie aber beispielsweise von Hinweisen profitieren, die ihnen die Aufgabenbearbeitung erleichtern und damit auch geringere Leistungsdifferenzen nach sich ziehen (Cerdán et al., 2019).

In der erstgenannten Gruppe von Studien, welche Gruppenvergleiche von einander in ihrem Leseverstehen ähnelnden Individuen vorgenommen haben, ergeben sich aufschlussreiche Muster. Beispielsweise zeigte sich bei Wong et al. (1982, Experiment 2), dass schwach lesende Kinder (hier: solche mit einer Lernschwierigkeit) ähnliche Leistungen in Tests erbrachten, unabhängig davon, ob sie vor der Leseaufgabe spezifische Hinweise erhalten hatten, worauf sie beim Lesen achten sollten. Die Gleichaltrigen ohne Lernschwierigkeit wiesen hingegen eine bessere Erinnerungsleistung auf, wenn sie eine relevanzbezogene Aufgabe mit spezifischem Auftrag bearbeiteten. In einer anderen Untersuchung (Cerdán, Gilabert & Vidal-Abarca, 2011) fand sich ein komplementäres Ergebnis: Bessere Leserinnen und Leser wandten sich eher von nicht-relevanten Informationen ab, während die schwächeren Leserinnen und Leser mehr Zeit darauf verwendeten, sich jenen Informationen zu widmen, die irrelevant waren für die Beantwortung von Fragen. Außerdem lokalisieren die stärkeren Leserinnen und Leser relevante Informationen für die Fragebeantwortung erfolgreicher, überwachen ihren Erfolg stärker, weil sie die Suchen nach Informationen zügiger abschließen, wenn sie die Antwort gegeben haben, und generell mehr Suchen nach Informationen durchführen (Vidal-Abarca et al., 2010, Experiment 1).

Erklärbar wird dieses Befundmuster im Licht einer anderen Studie, die den Zusammenhang von Relevanz der Informationen, erfasster Aufmerksamkeit und Erinnerungsleistung für stärkere und schwächere Leserinnen und Leser spezifischer untersuchte (Reynolds et al., 1993). Das Ergebnis war, dass nur bei den leistungsstärkeren Personen ein Zusammenhang gefunden werden konnte von längeren Lesezeiten bei den relevanten Zielinformationen – dies war unter den leistungsschwächeren Probandinnen und Probanden nicht der Fall. Dass eine solche verlängerte Lesezeit aber auch zu einer besseren Behaltensleistung führte, war anhand der Daten nicht zu belegen – was im Kern nur besagt, dass die quantitativen Zeitmaße nur ungenau Auskunft über die Qualität und Art der Prozesse erteilen, die für die Verarbeitung der Informationen den entscheidenden Unterschied ausmachen dürften (Hidi, 1995). Dennoch wird – trotz solcher uneindeutigen Befunde – vermutet, dass aufmerksameres Lesen von relevanten Informationen sich erst bei stärkeren, im Kompetenzerwerb weiter fortgeschrittenen Personen in besserem Lernen aus Texten niederschlägt – sei es auf Prozess-, sei es auf Produktebene (Reynolds, 1992).

2.3.1.2 Wortschatz

Eine weitere, zweite Variable, die in Studien als Kovariate und damit als eine potenzielle Einflussgröße erfasst wurde, ist der Wortschatz. Dieser wurde über das Wissen über Synonyme gemessen (Lapan & Reynolds, 1994; McCrudden & Schraw, 2010; Smith et al., 2010; Wilhite, 1985), und die Personen wurden nach ihren Testergebnissen in Leistungsgruppen mit hohem oder niedrigem rezeptivem Wortwissen unterteilt. Die Ergebnisse sind gebündelt: Maße wie Lesezeiten als Prozessindikatoren allein weisen entweder nicht auf Unterschiede (Lapan & Reynolds, 1994) oder auf ein schnelleres Lesen der leistungsstärkeren Personen hin (McCrudden & Schraw, 2010). Dennoch gab es in den Produktmaßen Indikatoren für eine Überlegenheit: Leistungsdifferenzen im Wortschatz zeigen sich auch in den abhängigen Produktvariablen beim relevanzbasierten Lesen. Leistungsschwächere Personen gemäß Wortschatztest erbringen schlechtere Leistungen auch in den Lese- und Erinnerungsleistungen, die in den jeweiligen Studien fokussiert wurden (Lapan & Reynolds, 1994; McCrudden & Schraw, 2010; Smith et al., 2010; Wilhite, 1985). Zum Teil ergaben sich Vorteile der Personen mit besseren Wortschatz-Resultaten bei anspruchsvolleren Fragen, während die Vergleichsgruppe mit geringerem Wortschatz verschiedene Fragearten ähnlich – nämlich ähnlich unterlegen – beantworteten (Wilhite, 1985). Das heißt: Mitunter waren die Differenzen bei besonders anspruchsvollen Fragen besonders prononciert. Dieses Muster der durchgängigen Überlegenheit wird in der Literatur als Beleg für die sogenannte „nicht-kompensatorische Hypothese“ gewertet, nach der Relevanzhinweise nicht dazu in der Lage sind, die ursprünglichen Differenzen zu kompensieren, sprich: Leistungsdifferenzen zu nivellieren imstande sind (McCrudden & Schraw, 2010). Ein solches Muster wirft die Folgefrage danach auf, inwiefern sich die Prozesse bei den leistungsstärkeren Personen unterscheiden, d. h. ob sie eine effizientere Strategie der Aufmerksamkeitslenkung verwenden, weil sie wegen eines automatisierten lexikalischen Zugriffs mehr kognitive Ressourcen übrig haben, die es ihnen im Prozess der Aufgabenbearbeitung erlauben, nicht nur die Textinhalte zu verarbeiten, sondern auch das eigene Verhalten zu steuern.

2.3.1.3 Abschlussbemerkung

Rezeptive schriftsprachliche Leistungen von Personen sind nach gegenwärtigem Kenntnisstand bedeutend für den Erfolg beim relevanzbezogenen Lesen. In diesem Teilkapitel wurden als solche Leistungsfähigkeiten Wortschatz und Leseverstehen fokussiert. Solche Fähigkeiten lassen sich als Bedingung modellieren, welche auch differenzielle Befunde in den Leistungstests erklären: Hidi (1995) legte ein Modell vor, nach dem von Personen auf zweierlei Art Aufmerksamkeit benötigt wird: erstens zum Verstehen der Texte, um zweitens auf dieser Grundlage die relevanten Informationen vertieft zu verarbeiten, um diese dann

gezielt zu lernen. Die Reihenfolge, die dieses Modell postuliert, ist entscheidend: Demnach gelingt es erst Personen mit ausreichendem Leseverstehen, Relevantes von Irrelevantem zu trennen und gezielt ihre Aufmerksamkeit auf dasjenige zu lenken, was sie lernen sollen bzw. wollen. Demgemäß wirkt diese Modellierung als eine mikrotheoretische Vertiefung des „Modells der Zielfokussierung der Relevanz“ (McCrudden et al., 2011a, s. Teilkap. 2.2.2.1).

Das Verstehen von Texten, das Hidi (1995) als Vorstufe relevanzbasierten Verarbeitens und Lernens von Textinhalten postuliert, lässt sich dabei recht gut auch auf den Wortschatz anwenden, weil der Wortschatz als wichtige Komponente des Leseverstehens gilt (s. z. B. Perfetti & Stafura, 2014) und Wissen über Lexembedeutungen empirisch stark mit Leseverstehensleistungen zusammenhängt (Ahmed & Wagner, 2020; Dong et al., 2020). Aus diesem Gedankengang heraus sind die in diesem Teilkapitel behandelten beiden schriftsprachlichen Fähigkeiten als Ressourcen zu werten, die eine effektive Verarbeitung von Informationen im Sinne einer anfänglichen Erschließung erleichtern, der dann fokussierte, relevanzbezogene Verarbeitungsprozesse und günstigenfalls gesteigerte Leistungen im Sinne von Produkten folgen.

2.3.2 Arbeitsgedächtniskapazität

Das Arbeitsgedächtnis ist eine besonders intensiv untersuchte Variable in der Forschung zu Relevanzeffekten. Im Fokus steht dabei die Kapazität des Arbeitsgedächtnisses in dem Sinne, dass es als Teil des menschlichen Gedächtnisses in seiner temporären Aufnahmekapazität von Informationseinheiten von Belang ist. Entscheidend ist also, wie viele Informationseinheiten eine Person in ihrem Arbeitsgedächtnis kurzfristig verfügbar hält, z. B. die letzten Wörter in gelesenen Sätzen („Lesespannentest“; Daneman & Hannon, 2007), und ob dies zu differenziellen Effekten bei Prozessen und Produkten des Lesens führt.

Doch warum sollte die Anzahl kurzfristig gespeicherter Informationseinheiten die Fähigkeit zum relevanzbezogenen Lesen beeinflussen? In ihrer Perspektive, das Arbeitsgedächtnis als den Ort der *kontrollierten Aufmerksamkeit* zu betrachten, offerieren Engle et al. (1999) eine Erklärung, weshalb eine höhere Kapazität des Arbeitsgedächtnisses vorteilhaft ist: Sie erlaubt eine selektivere Vorgehensweise, bei der Aufgabenziele und gelesene Inhalte präsent im Arbeitsgedächtnis bleiben. Das ist vor dem Hintergrund hervorzuheben, dass selektives, relevanzbezogenes Lesen auch verstehendes Lesen benötigt, aber darüber hinaus eine parallel bzw. zeitversetzt dazu ablaufende Entscheidung erforderlich ist, spezifischen, zum Leseziel passend wirkenden Informationen mehr Aufmerksamkeit zu widmen, um diese ins Langzeitgedächtnis zu überführen (Hidi, 1995). Engle et al. (1999) sowie Engle und Kane (2004) legen nahe, dass die Kapazität des Arbeitsgedächtnisses im Sinne der kontrollierten Aufmerksamkeit dann

wichtig wird, wenn die Aufgaben es erfordern, spezifische Gedächtnisrepräsentationen zu aktivieren oder aufrechtzuerhalten. Das gilt insbesondere dann, wenn Ablenkungen oder Interferenzen zu erwarten sind, was man als prototypische Begleiterscheinung des Lesens digitaler Texte bezeichnen kann (Wylie et al., 2018).

Zusammengefasst kann man daher Folgendes vermuten: Wenn jemand gemäß Engle et al. (1999) eine höhere Kapazität des Arbeitsgedächtnisses aufweist, so müsste sich dieser Vorteil in den Prozessen und Produkten des Lesens mit einer anspruchsvolleren Verwertungsabsicht positiv auswirken. Ein negativer Effekt wäre bei Personen mit geringer Kapazität zu erwarten.

2.3.2.1 Produkte

Bei den Produktmaßen offenbarten sich – nicht in allen Studien (z. B. Bohn-Gettler & Kendeou, 2014; Kaakinen et al., 2001, 2002, 2003) – Vorteile einer höheren Arbeitsgedächtniskapazität vor allem beim Zweck des Lernens (Linderholm & van den Broek, 2002) und beim Lesen mit Perspektiven in Form höherer Erinnerungsleistungen (Di Vesta & Di Cintio, 1997; Lee-Sammons & Whitney, 1991, Experiment 2). Ähnliches gilt allgemein für Verstehensfragen zum Gelesenen, die Personen mit hoher Arbeitsgedächtniskapazität besser beantworteten (Linderholm et al., 2008).

Neben solchen Befunden zu allgemeinen Effekten sind vertiefende Analysen von Interesse, die dabei helfen, die teils disparaten Resultate besser zu verstehen. Solche Zusatzanalysen ergaben beispielsweise, dass sich bei Personen mit geringer Arbeitsgedächtniskapazität die Anzahl relevanter und irrelevanter Informationen aus Erinnerungsprotokollen kaum unterschied, sie also weniger als Personen mit höherer Arbeitsgedächtniskapazität dazu in der Lage sind, Relevantes von Irrelevantem zu trennen (Kaakinen et al., 2001).

Dies wird gestützt von Studienbefunden (Lee-Sammons & Whitney, 1991, Experiment 2), nach denen sich Personen mit höherer Arbeitsgedächtniskapazität nicht stark nach zugewiesener Perspektive unterschieden, was ihre gleich guten Erinnerungsleistungen zeigten. Personen mit geringerer Arbeitsgedächtniskapazität waren hingegen nur in den Erinnerungsleistungen besser, die kongruent mit ihrer zugewiesenen Perspektive sind. Es wirkt also so, als wären sie weniger gut dazu in der Lage, Informationen aufmerksam zu lernen, die jenseits des originären Ziels der Leseaufgabe liegen. Zudem konnte daneben demonstriert werden, dass es Personen mit geringer Arbeitsgedächtniskapazität schwerer fällt, relevante Textinformationen wiederzugeben, wenn sie nachträglich eine andere Perspektive einnehmen sollten (Di Vesta & Di Cintio, 1997; Kaakinen et al., 2001). Dieser letztgenannte Befund lässt sich als Indiz dafür werten, dass Personen mit geringer Arbeitsgedächtniskapazität ihre limitierten Ressourcen eher dazu nutzen, ganz gezielt relevante Informationen zu suchen und diese zu memorieren, während Personen mit größerer Arbeitsgedächtniskapazität flexibler vorgehen

und Aufgabenerfordernisse und Gedächtnisnutzung besser in Einklang bringen können.

2.3.2.2 Prozesse

Bei den Prozessindikatoren zeigen sich Differenzen zwischen Gruppen mit hoher und niedriger Arbeitsgedächtniskapazität nicht durchgängig, allerdings mitunter durchaus markant. Personen mit *geringerer Arbeitsgedächtniskapazität* neigten

- zu längeren Lesezeiten insgesamt (Linderholm et al., 2008), die sich in Eye-Tracking-Studien teils über längere Fixationen von relevanten Textpassagen erklären lassen (Kaakinen et al., 2003, Experimente 1 und 2), nicht aber über die Dauer erstmaliger Fixationen (Kaakinen et al., 2002),
- zu mehr oberflächenorientierten Strategien wie dem Wiederholen von Textinhalten (Linderholm & van den Broek, 2002),
- und generell zu einer wenig dem Lesezweck angepassten Vorgehensweise, weil sich ihre Prozesse wie das Paraphrasieren oder nicht auf Kohärenz abzielenden Kognitionen kaum nach Lesezweck unterschieden (Bohn-Gettler & Kendeou, 2014).

Personen mit *hoher Arbeitsgedächtniskapazität*

- kommentieren parallel zum Lernen beim Lesen ihr Verstehen stärker metakognitiv (Linderholm & van den Broek, 2002),
- paraphasierten beim Lesen zum Lernen Gelesenes häufiger und zeigten generell beim Lesen zum Lernen weniger nicht-kohärenzbezogene Prozesse (Bohn-Gettler & Kendeou, 2014),
- fixierten relevante Informationseinheiten beim erstmaligen Lesen länger (Kaakinen et al., 2002) und sie taten es adaptiv, nämlich vor allem bei unvertrauten Themen (Kaakinen et al., 2003, Experiment 1).

Auf einen finalen und wichtigen differenzierenden Befund macht die Studie von Burton und Daneman (2007) aufmerksam. Sie fanden eine kompensatorische Interaktion von epistemischen Überzeugungen und Arbeitsgedächtniskapazität: Wenn Personen zwar eine geringe Kapazität des Arbeitsgedächtnisses aufweisen, aber gleichzeitig günstige, fortgeschrittene epistemische Überzeugungen, dann demonstrieren sie ein vorteilhaftes Leseverhalten mit längeren erneuten Betrachtungen relevanter Textteile und einer im Vergleich auch besseren Erinnerungsleistung. Das heißt, es lohnt sich, Interaktionseffekten nachzuspüren, die trotz ungünstiger kognitiver Ausgangsvoraussetzungen dabei helfen, diese Defizite auszugleichen.

2.3.2.3 Abschlussbemerkung

Die allgemeinen Befunde lassen sich bündeln, dass trotz einer gewissen Uneindeutigkeit in der Forschungslage die Personen mit einer höheren Arbeitsgedächtniskapazität ein insgesamt effektiveres Vorgehen an den Tag legen: Sie scheinen ihre Aufmerksamkeit verstärkt jenen Informationen zu widmen, die für ihr Leseziel wichtig sind. Sie steuern ihr Vorgehen anscheinend stärker und nutzen mehr Prozesse, die dem Aneignen des Inhaltes dienen, was sich partiell in besseren Leistungen niederschlägt. Dies steht in Einklang zu der Vorstellung, dass sie eine kontrollierte Aufmerksamkeit dazu nutzen, das jeweilige Leseziel erfolgreicher in Produkten und Prozessen zu erreichen.

Demgegenüber wirkt die Bearbeitung von Leseaufgaben bei Personen mit geringer Arbeitsgedächtniskapazität zum Teil ineffektiver: Sie lesen mitunter länger, konsultieren relevante Informationen erneut, die sie beim erstmaligen Lesen offenbar nicht tief genug verarbeitet haben und scheinen ihre Leseprozesse weniger stringent an der Aufgabe auszurichten. Anderen Merkmalen wie günstigen epistemischen Überzeugungen kommt hierbei eine kompensatorische Funktion zu: Wenn diese günstig ausgeprägt sind, wirkt selbst bei Personen mit geringer Arbeitsgedächtniskapazität das Vorgehen wieder zielführender in Produkt- und Prozessmaßen.

2.3.3 Thematisches Vorwissen

Das dritte und in diesem Teilkapitel letzte individuumsbezogene Merkmal ist das inhaltliche Vorwissen. Inhaltliches Vorwissen lässt sich als domänenspezifischer, also einem Wissensgebiet zuordenbarer deklarativer Langzeitgedächtnisinhalt bezeichnen, mithin als die Gesamtheit von Informationen über einen enger oder weiter gefassten Bereich (McCarthy & McNamara, 2021; Smith et al., 2021). Diese Art des Vorwissens hat in der Leseforschung nicht zuletzt darum breite Aufmerksamkeit erfahren, weil dieses Vorwissen in zentralen Modellen des Leseverstehens als Grundlage von Verstehensleistungen angenommen wird, wenngleich die Modelle zu den Mechanismen andere Wirkweisen postulieren (McNamara & Magliano, 2009).

Es wird in der Forschung zur Relevanz beim Lesen angenommen, dass ein höheres Vorwissen dabei hilft, effizienter beim Lesen vorzugehen und effektiver die Relevanz von Informationen zu beurteilen und diese Informationen auch leichter zu memorieren. Ist das Wissen hingegen geringer und damit die Basis zum Erkennen von Relevanz weniger stark ausgebaut, ist ein höherer Aufwand wahrscheinlich – nebst geringeren Leistungen im Sinne der Produkte (Kaakinen & Hyönä, 2011). Ausgehend von der Leseforschung spezifizieren McCarthy und McNamara (2021) vier trennbare, aber zusammenhängende Dimensionen

des Vorwissens: a) Menge (Anzahl von Konzepten, die für das Verstehen nötig sind), b) Akkuratheit (Korrektheit von Konzepten und ihren Verbindungen untereinander auf einem Kontinuum entlang zu-/abnehmender Akkuratheit), c) Spezifität (als Hierarchie von allgemeinem Wissen hin zu feinteiligeren Spezialkenntnissen in einzelnen Themen) und d) Kohärenz (der Grad an strukturierter Vernetztheit des Wissens). Holzschnittartig formuliert und auf die Belange des relevanzbezogenen Lesens appliziert lässt sich Folgendes annehmen: Je normativ besser (a) quantitativ mehr vorhanden, b) sachlich richtiger, c) genauer und d) gut verknüpft) das eigene inhaltliche Vorwissen als eine Grundlage für das korrekte Zuweisen von Relevanz ist, desto besser sollten die Leistungen ausfallen.

2.3.3.1 Produkte

Bei den Produktmaßen ist die Lage teils uneindeutig. Es gibt Befunde, die auf keine Effekte des Vorwissens verweisen (Kaakinen & Hyönä, 2007). Dem stehen Befunde gegenüber, die für Vorwissenseffekte sprechen. Personen mit höherem Vorwissen, die elaborative Fragen beantworteten, welche ihrerseits stark vorwissensbasiert sind, zeigten beispielsweise höhere Leistungen als Personen mit geringerem Vorwissen (Ozgunor & Guthrie, 2004; Smith et al., 2010). Mit größeren Indikatoren – nämlich dem Vergleich von Personengruppen, die Internetdokumente lesen, welche ihrer Wissensdomäne entsprechen bzw. nicht entsprechen – lassen sich ebenfalls Effekte nachweisen: Personen erinnern sich an mehr Textinhalte, die ihrer Wissensdomäne entsprechen (Rouet, 2003). In eine ähnliche Richtung gehen Studien, bei denen die unterstellte Vertrautheit des Themas variiert wurde, indem beispielsweise seltene und häufige Krankheiten behandelt wurden. Hier wiesen Personen bessere Erinnerungsleistungen bei vertrauteren Themen auf (Burton & Daneman, 2007; Kaakinen et al., 2003, Experiment 1).

2.3.3.2 Prozesse

Dass das Vorwissen als Grundlage für die Relevanzzuweisung tatsächlich im Prozess des Lesens genutzt wird, illustriert die Studie von McCrudden et al. (2010, s. „Illustrierendes Beispiel 5“ in Teilkap. [2.2.2.3](#)). In ihr zeigte sich, dass ohne spezifische Relevanzhinweise Personen der Kontrollgruppe die Relevanz über die Vertrautheit von Inhalten bestimmten. Die Personen nutzten also ihr Vorwissen als Bemessungsgrundlage und widmeten ihre Aufmerksamkeit verstärkt Informationen, zu denen sie Vorwissen hatten, und hielten diese für relevanter. Neben solche Effekte treten jene zu Gruppenvergleichen, also Ergebnisse aus Studien, die Vorwissen über den Vergleich von Personengruppen hohem bzw. geringem Vorwissen erfassten. Dies haben finnische Studien zum Beispiel über den Einsatz von Textsets operationalisiert, in denen das Thema (Gruppen von Krankheiten)

oder die Perspektive (Einbrecher vs. Innenarchitektin) den Personen mutmaßlich bekannter oder unbekannter gewesen sein dürften. Die Ergebnisse sprechen dafür, dass bei geläufigen Themen resp. Perspektiven Effekte der einzunehmenden Perspektive geringer ausfallen als bei Themen, die den Personen fremd sind. In diesem Fall lasen sie relevanten Textteile teils erheblich länger (Burton & Daneman, 2007; Kaakinen et al., 2003, Experiment 1; Kaakinen & Hyönä, 2008) und verarbeiteten relevante Sätze anders, da sie dann häufiger vom Satzende zu vorhergehenden Satzteilen blickten, was als Indiz für ein stärkeres Prüfen gewertet werden kann (Kaakinen & Hyönä, 2007).

2.3.3.3 Abschlussbemerkung

Thematisches deklaratives Vorwissen gilt als Grundlage für Verstehensprozesse, ist aber kein homogenes Konstrukt, sondern besteht aus mehreren Dimensionen. Was für die Leseverstehensforschung bereits recht systematisch aufgearbeitet wird (McCarthy & McNamara, 2021), steckt bei der Forschung zur Relevanz noch in den Kinderschuhen, was sich an der wenig systematisch wirkenden Operationalisierung des Konstrukts besonders eindrucksvoll konturieren lässt. Gleichwohl haben die hier konsultierten Studien ein übergreifendes Muster: Die Befunde verdeutlichen eine Diskrepanz von Produkt- und Prozessmaßen: Die Erinnerungsleistungen sind bei vertrauten Themen größer, während der prozessuale Aufwand geringer ausfällt. Dies spricht für eine effizientere Nutzung des Vorwissens. Bei den unvertrauten Themen sind die Leistungen in den Produkten schlechter, obwohl die Prozessmaße auf aufwändigeres Verarbeiten verweisen. Anders gesagt: Es wirkt so, als unterminiere ein zu geringes Wissen das Verstehen von Texten, die zu Zwecken des relevanzbezogenen Lesens konsultiert werden.

2.4 Probleme

Relevanz- und damit zielbezogenes Lesen ist ein aufgabenspezifisches Vorgehen, das als stark selbstregulatives Vorgehen gelten kann (Rouet et al., 2017; Vidal-Abarca et al., 2011). Als Bedingung dafür wird ein präzises Aufgabenmodell postuliert, also eine metakognitive Repräsentation von Aufgabenerfordernissen und notwendigen Handlungen zur Zielerreichung. Das Aufgabenmodell ist bereits Bestandteil umfassenderer Modellierungen, welche es als metakognitives Steuerelement beinhalten (Rouet, 2006; Rouet et al., 2017).

Das Aufgabenmodell bzw. andere damit theoretisch verwandte Konstrukte wie Kohärenzstandards und Kognitionen zur Relevanz bilden eine Basis für das selbstregulierte Vorgehen beim aufgabenbasierten Lesen. Wenn Personen selbstreguliert vorgehen sollen und damit ihre lernbezogenen Aktivitäten zielgerichtet koordinieren und ausrichten, dann sind Fehlanwendungen und Probleme

wahrscheinlich. In diesem Teilkapitel werden zwei Problemfelder skizziert, die sich empirisch andeuten und die sich vor allem in einem Missverhältnis von lesebezogener Aufgabe und der Wahl von Informationen bzw. gezeigtem Lernverhalten manifestiert haben. Dieser Forschungszweig ist noch wenig ausgebaut, sodass die Darstellung sich notwendigerweise auf einige Schlaglichter beschränken muss. Aus diesem Grund wird in diesem Teilkapitel nur ein Ausschnitt fokussiert, nämlich zum einen Studien, in denen es darum geht, die Sensibilität von Personen zu erfassen, Informationen hinsichtlich ihrer Eignung für die Zwecke der konkreten Aufgabenstellung einzuschätzen (2.4.1). Ein zweites Problemfeld betrifft das Phänomen, dass die Aufgabenanalyse bzw. ein auf ihr basierendes lernstrategisches Vorgehen ausbaufähig ist (2.4.2).

2.4.1 Problemfeld 1: Probleme bei der adäquaten Einschätzung von Informationen hinsichtlich ihrer Eignung in puncto Aufgabenrelevanz

Wie adäquat sind Personen darin, textuelle Informationen gemäß ihrem instrumentellen Wert einzuschätzen? Dieser Frage haben sich eine Handvoll von Studien verschrieben, deren Ergebnisse sich zu drei Problemen verdichten lassen.

- *Es gibt große Variationen in dem, was Personen für relevant halten, und nur einen geringen Anteil von Personen, der kalibriert genug vorgeht.* Studien, in denen Jugendliche gebeten werden, zu variierenden Aufträgen – a) interessante Textteile markieren, b) Textteile für einen bevorstehenden Test mit Hinweisen zur Lehrperson und ihren impliziten Hinweisen zur Bedeutsamkeit von Details als relevant markieren und schließlich c) vom Autor aus gedacht als wichtige und zentrale Informationen zu markieren – eine Sensibilität für geeignete Informationen zu demonstrieren, zeigen eine große Bandbreite dessen, was die Personen als aufgabenrelevant markieren (Schellings et al., 1996b). Dabei fällt insbesondere beim Auftrag des Lernens auf, dass hier die gesamte Bandbreite von Unterstreichungen – also von keiner einzigen bis hin zu allen möglichen Unterstreichungen als Markierung relevanter Textteile – beobachtet wurden. Diese beiden Extrempole verweisen darauf, dass den Testpersonen ein adäquat kalibriertes Relevanzurteil schwerfällt. Mittels einer vertiefenden Analyse ergab sich das Bild, dass nur 18 Prozent der Jugendlichen tatsächlich jene Informationen unterstrichen hatte, die für die Zwecke der jeweiligen Aufgabenbearbeitung die höhere Relevanz aufwiesen, und diese Gruppe von Jugendlichen markierte insgesamt auch selektiver und damit weniger Text als fast alle anderen Jugendlichen (Schellings et al., 1996a). Das bedeutet: Nicht einmal jede fünfte Person hatte bei den drei unterschiedlichen

Aufgaben, die sie bearbeitete, mit ihren Markierungen demonstriert, dass sie auftragssensibel Informationen relevanzbezogen ausgewählt hatte.

- *Leserinnen und Leser haben Defizite darin, zielsicher Informationen als relevant zu erkennen.* In einer Studie wurde überprüft, wie stark sich die Urteile von Biologielehrpersonen und den von ihnen unterrichteten Jugendlichen zehnter Klassen gleichen (Schellings & van Hout-Wolters, 1995). Die Jugendlichen und ihre Lehrpersonen wurden gebeten, diejenigen Textteile aus einem Lehrbuchtext zu unterstreichen, die ihrer Meinung nach wichtig waren, um sich auf einen Biologietest vorzubereiten. Einen parallelen Auftrag erhielten die Lehrpersonen, die ebenfalls jene Textteile unterstreichen sollten, die sie für einen regulären Biologietest für wichtig hielten. Ein Ergebnis bestand darin, dass nur 14 Prozent der Wortgruppen sowohl von den Jugendlichen als auch von deren Lehrpersonen unterstrichen wurden. Das heißt: Die Urteile, was für einen definierten Zweck als essenziell gilt, variieren erheblich. Wären die gewählten Informationen tatsächlich die Grundlage für das Lernen für einen Test, wäre es angesichts solcher geringen Übereinstimmungen fraglich, dass die Lernenden ihn bestehen könnten. Andere Studien in anderen Domänen wie Geschichte kommen zu ähnlichen Befunden (Broekkamp, van den Bergh et al., 2002), und sie zeigen, dass es zudem erhebliche Unterschiede zwischen einzelnen Schulklassen und ihren Lehrpersonen gibt (Broekkamp, van Hout-Wolters et al., 2002).
- *Leserinnen und Leser sind nur zu einem geringen Anteil gut darin, adäquate extratextuelle Bemessungsgrundlagen als explizite Auswahlgrundlage zu bemühen.* In der Studie von Schellings und van Hout-Wolters (1995) wurden die Jugendlichen gebeten, noch einmal schriftlich darzulegen, weshalb sie einzelne Textteile markiert hatten bzw. das nicht getan hatten. Ein Vergleich von Schülerinnen und Schülern mit hoher Ähnlichkeit bei der Einschätzung von Textteilen in Bezug auf das Urteil ihrer Lehrpersonen vs. solchen mit geringer Einschätzungsübereinstimmung ging den Begründungen für das (Nicht-)Unterstreichen nach und fand ein Muster vor. Die Jugendlichen, die stark abwichen, gaben vor allem Gründe an, die sich auf die textuellen Merkmale wie Typografie oder zentrale Inhalte bezogen. Jugendliche mit besonders hoher Korrespondenz mit dem Lehrpersonen-Urteil nannten solche Gründe eher dafür, dass sie Textteile *nicht* markiert hatten. Sie gaben stärker Begründungen an, die sich auf potenzielle Testinhalte bezogen. Dieser Gruppenvergleich erbrachte mithin, dass die Adäquanz der Beurteilung variierte und dass nur jene Personen vielversprechende Textteile selektierten, die den angemessenen Zweck des Lesens als Ursache für ihre Entscheidungen im Vergleich zu anderen Gründen präsender benennen konnten. Ansonsten ist auch in anderen Studien belegt worden, dass es Jugendlichen schwerfällt, sich von textimmanenten Merkmalen zu lösen, um textexterne Relevanz zuzuweisen (Schellings et al., 1996b; Schellings & Broekkamp, 2011).

2.4.2 Problemfeld 2: inadäquates Vorgehen bei der Aufgabenbearbeitung wegen mangelnder Aufgabenanalyse

Die Aufgabenanalyse und die sich daraus speisende sinnvolle Auswahl von zielführenden Vorgehensweisen ist in theoretischen Modellen als Bedingung glückenden selbstregulierten Lernens zentrales Element (Butler & Cartier, 2004; Panadero, 2017). Dennoch bereitet dies offenkundig Leserinnen und Lesern – zumal im Kontext des Lernens mittels Lesen, also einem Bereich des Lesens mit hohen Kohärenzstandards (s. Teilkap. [2.2.2.2](#)) – erhebliche Mühe. Dies illustrieren die folgenden zwei Ergebnislinien:

- *Die gewählten Strategien und (hohen) Situationsanforderungen passen nicht zusammen.* Es ist ein bekanntes Phänomen, dass Personen nicht dasjenige lernstrategische Verhalten an den Tag legen, welches sinnvoll und zielführend wäre (s. z. B. im Überblick bereits Garner, 1990). Dafür gibt es aktuelle Beispiele wie aus der Studie von Butler et al. (2011). Hier gaben Jugendliche der Sekundarstufe zu Protokoll, vor dem Bearbeiten lesebezogener Aufgaben wenig zu planen, und sie nannten häufig oberflächenorientierte Strategien beim Lesen wie das Beachten von Überschriften. Tiefenorientierte Strategien, die darauf abzielen, beim Lesen Informationen miteinander zu verknüpfen, kamen weniger häufig in den Nennungen der Fragebögen vor, allerdings waren sie es, die – zusammen mit einem selbstregulierten Vorgehen allgemein – mit besseren Leseleistungen korrespondierten. Dieses stellvertretende Ergebnis einer Studie spiegelt wider, dass Personen (in diesem Fall: Jugendliche der Sekundarstufe) zu wenig dazu neigen, ihr Vorgehen gemäß den Aufgaben zu planen und einen entsprechenden Aufwand zu betreiben, was aber angesichts der hohen Aufgabenerfordernisse beim Lernen mit dem Lesen erforderlich und zielführend wäre.
- *Eine sorgfältige und umfassende Aufgabenanalyse findet nicht oder nur in Ansätzen statt.* Wenn Schülerinnen und Schüler Texte zum Zweck des Lernens (oder auch anderen Verwendungsabsichten folgend) bearbeiten, ist ein wiederkehrender Befund, dass ihre Handlungen teils sofort und ausschließlich auf der Objektebene des Gegenstands stattfinden. Eine Aufgabenanalyse ist hingegen höchstens rudimentär zu beobachten, teils nicht einmal das. Dies demonstrierte beispielsweise eine Studie mit Fünftklässlern, die mehrere Texte mit dem Ziel des Lernens lasen und dabei ihr Vorgehen verbalisierten (Vandeveldt et al., 2015): Diese 10- bis 11-Jährigen nahmen über mehrere Fälle und mehrere Aufgaben hinweg keine Aufgabenanalyse vor. Ein ähnliches Bild ergab sich in einer Studie mit Zehntklässlern (Schellings & Broekkamp, 2011), also erheblich älteren Personen. In dieser Untersuchung waren Aussagen in Lautdenk-Protokollen sehr selten, die eine globale Aufgabenanalyse

indizierten, etwa das nochmalige Lesen oder das Unterstreichen von Inhalten aus Aufgaben – beides kam in maximal einem von sechs Protokollen vor. Ähnlich knapp war die Zahl der Äußerungen, die auf eine deutliche Präsenz des konkreten Arbeitsauftrags beim lokalen Vorgehen schließen lassen, in diesem Falle eine aufgabensensible Auswahl von Informationen.

2.4.3 Abschlussbemerkung

Die beiden in diesem kurzen Teilkapitel skizzierten Problemfelder spiegeln zunächst einmal wider, dass dieser Bereich in der empirischen Leseforschung noch kaum ausreichend erforscht wurde. Es klafft also eine Lücke zwischen theoretisch hoch veranschlagter Bedeutsamkeit des Aufgabenmodells und seiner empirischen Untersuchung. Doch selbst wenn das Bild fragmentarisch ist, zeigen sich Konturen: Personen neigen von sich aus zu einem geringen Anteil dazu, eine ausreichend tiefe Aufgabenanalyse vorzunehmen, die zudem handlungsleitend wird (Problemfeld 2 aus Teilkap. [2.4.2](#)). Zusätzlich besteht die deutliche Tendenz, dass es nur unzureichend glückt, Informationen treffsicher als relevant zu identifizieren, also hinsichtlich ihrer Aufgaben- und Kontextsensitivität angemessen zu unterscheiden (Problemfeld 1 aus Teilkap. [2.4.1](#)). Welche tieferen Verstehensleistungen davon betroffen sind, ist leider noch kaum erforscht.

Die beiden Problemfelder lassen sich der Prozess- und Produktperspektive relativ eindeutig zuordnen. Produktbezogene Probleme ergeben sich aus dem Problemfeld 1: Sie verweisen darauf, dass der Informationsauswahl ein gewisser arbiträrer Charakter innezuwohnen scheint. Denn die als relevant deklarierten Informationen passen entweder nicht ausreichend zum Zweck des Lesens, sie streuen interindividuell stark oder sie wurden wenig nachvollziehbar begründet. Dies lässt Probleme bei den Prozessen erwarten, und prozessbezogen war das Problemfeld 2. Ihm zufolge bestehen gravierende Defizite darin, Energie und Aufwand für die Analyse der eigentlichen Aufgabenstellung zu reservieren. Im Schulalter, für das die wenigen ausgewerteten Studien stehen, finden beobachtbare metakognitive Prozesse, die auf das Bilden von Aufgabenmodellen hinweisen würden, kaum oder gar nicht statt. Es muss demnach nicht verwundern, dass oberflächliche, dem durchaus hohen externen Anforderungscharakter der Leseaktivitäten entgegenstehende Verhaltensweisen überwiegen.

2.5 Fazit

Die erste im Rahmen dieses Buch umfangreich vorgestellte Komponente des (nicht nur) digitalen Lesens ist generisch und wird unter der Bezeichnung „*Aufgabenmodell*“ gefasst (Rouet et al., 2017). Dieses Aufgabenmodell ist eine Repräsentation des mit dem Lesen verbundenen Zielzustandes und dessen optimaler Erreichbarkeit mit dem Einsatz eigener Ressourcen. Leserinnen und Leser haben demnach ein ex- oder internal vorhandenes Ziel, das sie mit dem Lesen erreichen sollen oder wollen, und diese Zielvorstellung entscheidet als metakognitive, mindestens bewusstseinsfähige Steuerungseinheit über Einsatz und Nutzen von lesendenseitigen Ressourcen. Unter diese Ressourcen fallen Vorwissens- und Strategienutzung sowie die Bereitschaft, für den angestrebten Zielzustand das nötige Maß an Motivation und Engagement aufzubringen. Aufgabenmodelle haben damit eine klärende Funktion, was sie sehr nahe und prototypisch an die aufgabenanalytischen Aktivitäten des selbstregulierten Lernens rückt (s. Kap. 6).

Das Aufgabenmodell wird in der Theorie als ein komplexes, von einem Kontextmodell gerahmtes Konstrukt beschrieben, das zudem dynamisch ist und über Feedbackschlaufen einer Veränderung unterliegt. Doch auch das Aufgabenmodell lässt sich als ein übergeordnetes, rahmendes Konstrukt begreifen. Andere, inhaltlich verwandte Konstrukte der Leseforschung weisen einen enger fokussierten Bereich auf, der sich der Logik von Aufgabenmodell unterordnen lässt. So fungieren „*Kohärenzstandards*“ als Benchmark eines angestrebten Erfolgs beim Lesen im Sinne der angestrebten Kohärenz der mentalen Repräsentation des gelesenen Textes. Für die Bildung des Aufgabenmodells ist von Belang, dass diese Kohärenzstandards variabel (und auch nicht nur von der external definierten Aufgabe abhängig) sind (van den Broek et al., 2002). Je höher Leserinnen und Leser ihren Kohärenzstandard setzen, desto aufwändiger gestaltet sich ihr jeweiliger Leseprozess, indem sie stärker ihr Vorwissen dazu nutzen, Inferenzen zu bilden, und ein selbstreguliertes, lesestrategisches Vorgehen demonstrieren. Ist der Kohärenzstandard hingegen gering, ist erwartbar, dass die lesende Person auch mit einer weniger kohärenten Repräsentation zufrieden ist, die dem aktuell angestrebten Verstehungserfolg entspricht. Der Bezug zum Aufgabenmodell ergibt sich damit über den Zielzustand des Lesens, den beide Konstrukte als Produkt modellieren, welches sich aus der Interaktion von sowohl externalen Vorgaben und Einflussgrößen als auch internalen Merkmalen speist.

Ein weiteres Konstrukt, das zugleich am stärksten erforscht ist, ist die „*Relevanz*“. Relevanz basiert auf einer von außen an einen Text herangetragenen Intention im Sinne einer Verwertungsabsicht. Es handelt sich also um ein instrumentelles Zuweisen von Bedeutsamkeit von Informationen oder größeren Texteinheiten hinsichtlich ihrer wahrgenommenen Passung zu einem Ziel. Die bisherige Forschung hat sich darauf konzentriert, externe Zielvorgaben zu formulieren und auf eine nicht ganz unproblematische Umsetzung (im Sinne einer

1:1-Befolgung durch die Testpersonen) zu überprüfen, teils ohne ausreichende Berücksichtigung individueller Interpretationen und Repräsentationen und meist mit Studierenden als Sample in experimentellen Studien. Dabei überwiegt eine Gesamttendenz, sich auf spezifische Textteile, meist Propositionen, Sätze oder Absätze zu konzentrieren. In den variierenden Gruppen von Studien, die sich entweder global oder lokal auf Relevanz fokussieren und von denen in diesem Kapitel nur ein Ausschnitt beleuchtet wurde, ergibt sich über die fünf im Kapitel betrachteten Hinweisgruppen ein relativ homogenes Gesamtbild, das man für Prozesse und Produkte konturieren kann:

- *Prozesse:* Eye-Tracking-Studien zufolge lesen Personen relevante Informationen (in den Texten, aber auch die Hinweise) und Textteile länger und wiederholter. Irrelevante Informationen lassen sie eher aus. Das Lesen erfolgt damit selektiver, aber auch teils dynamischer, weil es zu einem Abgleich von Informationen und ihrer Passung zum Leseziel insofern kommt, als Indikatoren des Leseverhaltens auf stärkere Wechsel zwischen Aufgabe und zu ihnen passenden Textstellen hinweisen. Lautdenk-Studien ergänzen das Bild, denn sie erbrachten das Resultat, dass die relevanten Informationen im Arbeitsgedächtnis präsent gehalten und mit dem Vorwissen der Person in Verbindung gebracht werden. Für eine stärkere Beanspruchung des Arbeitsgedächtnisses sprechen Ergebnisse aus Reaktionstests: Relevante Informationen benötigen gemessen an den längeren Reaktionszeiten mehr Aufmerksamkeit. Summa summarum lassen sich die Effekte von Aufgabenstellungen zur Relevanz ergo in einer Vielzahl von Indikatoren der fokussierten Aufmerksamkeit und kognitiven Verarbeitung bei spezifischen Informationseinheiten nachweisen.
- *Produkte:* Produkte bei den Effekten von relevanzbezogenen Hinweisen sind dominant als Erinnerungsleistungen untersucht worden. Dabei lassen sich über die Studien hinweg eher oberflächenorientierte Leistungen wie die Erinnerungen an spezifische Inhalte von Leistungen trennen, die auf Inferenzen oder tiefergehenden Transformationen basieren. Je spezifischer Aufträge auf einzelne Informationen abzielen, desto konsistenter ist das Bild bei den empirischen Effekten. Je unspezifischer die Hinweise sind, desto weniger klar ist die Befundlage. Bei den Aufgabeneffekten und den Effekten von tiefenorientierten Fragen hingegen zeigen sich Vorsprünge in den konsultierten Studien eher bei den Leistungen, die das Verbinden, Verändern oder Anwenden von gelesenen Texten erfassten. Hierin deuten sich also Befundmuster an, die spezifizieren, welche Leistungen durch Relevanzhinweise gesteigert werden konnten.

Die Forschungslage zu Relevanzhinweisen und ihren Effekten auf Prozesse und Produkte macht indirekt deutlich, worum es im verwandten Zweig der Aufgabenmodelle geht: Eine (externe) Aufgabe führt zu einem Resultat, doch das dazwischenliegende Bindeglied – das Aufgabenverständnis bzw. das Aufgabenmodell

als eine eigene und mutmaßlich maßgebende interne Repräsentation –, ist weitestgehend ungeklärt. Denn gerade globale Zwecke des Lesens zum Lernen oder zur Unterhaltung weisen viele interpretationsbedürftige Ambiguitäten und empirische Lücken auf, worauf beispielsweise Ramsay und Sperling (2011) dezi- diert aufmerksam machen. Diese Unschärfe und momentane Unbestimmtheit des Aufgabenmodells in der Empirie ist umso bemerkenswerter, als beim Lesen digitaler (Internet-)Dokumente im Sinne eines Problemlöseprozesses, zumal mit offenen Aufträgen und dann, wenn es keine vorgegebenen Inhalte im Sinne a priori definierter Lesematerialien gibt, eine Zielvorstellung zwingend nötig wirkt (Brand-Gruwel et al., 2009).

Relevanzbezogenes Lesen ist in Prozessen und Produkten mehr oder minder stark damit assoziiert, welche individuellen Merkmale eine Person aufweist. Die in diesem Kapitel vorgenommene Fokussierung auf individuumsbezogene Merkmale und die ausschließliche Betrachtung kognitiver Variablen stellt selbst- redend nur einen Ausschnitt dar. Als günstig für das relevanzbezogene Lesen gelten: elaborierte sprachliche Fähigkeiten im Sinne eines breiten Wortschatzes und des Leseverstehens, eine hohe Kapazität des Arbeitsgedächtnisses sowie ein gut ausgebautes Wissen zum Thema. Alle diese kognitiven Merkmale eint, dass sie eine effiziente Verarbeitung der textuellen Informationen erleichtern. Wer hingegen lexikalische Defizite, eine geringe Lesekompetenz, eine limitierte Auf- nahmekapazität für temporär zu verarbeitende Informationen und schließlich ein geringes Vorwissen aufweist, läuft Gefahr, kognitiv aufwändigere Prozesse zu absolvieren, die aber trotzdem nicht zum ähnlichen Erfolg führen wie bei Personen, bei denen diese Merkmale besser ausgeprägt sind.

Neben solchen Determinanten für interindividuelle Unterschiede, die in va- riiertem Grad erforscht wurden, treten empirische Befunde, welche Problem- lagen konturieren. Diese Probleme beim relevanzbezogenen Lesen offenbaren sich sowohl bei den Prozessen als auch bei den Produkten. Eine geringe Aktivität der Aufgabenanalyse und ein Missverhältnis von erforderlichen und sinnvollen Vorgehensweisen und gezeigten und weniger zielführenden Handlungen stehen für die Problemlagen bei den Prozessen. Bei den Produkten gibt es Indizien für eine zu geringe Sensibilität der Relevanzurteile hinsichtlich der kontextuellen Erfordernisse, für hohe Streuungen bei der Informationsauswahl und zu guter Letzt für eine unsicher wirkende Basis für die Entscheidungen für oder gegen die Relevanzzuweisung.

Was heißt all das? Relevanzbezogenes Lesen, also ein Lesen, bei dem der Zweck der Lektüre – sei es das Finden einzelner Informationen, das Vergleichen von Aussagen oder das Verstehen von Sachverhalten in digitalen Dokumenten – sich aus Faktoren speist, die außerhalb des Textes liegen, dürfte einen typischen Anwendungsfall des digitalen Lesens darstellen. Daher erscheint es nötig, die Ziele der Leseaktivitäten hinreichend bestimmt zu haben, und hierfür bietet die Lesetheorie unterschiedliche Erklärungsansätze und Modelle an. Auch wenn viele

der Befunde in diesem Kapitel nicht genuin auf das digitale Lesen bezogen waren, stellen sie doch in ihrer Gesamtheit auch und gerade für das digitale Lesen eine Basis dar. Diese Basis hilft insofern, als die externalen Leseabsichten (darunter auch den wenig hinterfragten Normalfall des Verstehens ganzer Texte), eine wichtige Komponente bilden, da sie als metakognitive Repräsentation lesebezogener Ziele (und wie man sie erreichen kann) den Ausgangspunkt markieren, was eine Person bereits vor und zusätzlich zur Leseaktivität für sich hinreichend geklärt haben sollte. Dieses Formieren von kalibrierten Aufgabenmodellen wirkt daher sowohl theoretisch als auch empirisch als auch didaktisch als wichtiger Bereich des digitalen Lesens.

3 Komponente 2: Informationsquellen evaluieren (Sourcing)

Darum geht es

Digitales Lesen fordert einen distanzierten und provisorischen Umgang mit Informationen und Informationsquellen. Die Ubiquität von Lesegelegenheiten über das Internet und die verschiedenartigen Angebote an Informationen zwingen Personen dazu, selektiv vorzugehen. Dies erfolgt sowohl bei der Recherche als auch generell bei einem relevantbezogenen und damit instrumentellen Lesen, bei dem die Passung von Inhalt und Ziel leitend ist. Daneben gibt es weitere Bezugspunkte, welche weniger den Inhalt als vielmehr dessen Genese, Ursprung und Kontext betreffen. Diese lesendenseitigen Zuweisungen der (epistemischen) Glaubwürdigkeit werden unter dem Begriff „Sourcing“ in der Leseforschung mit einer stetig zunehmenden Zahl von Studien intensiv untersucht.

Diesem Verbund von Sourcingprozessen widmet sich dieses Kapitel, und es beginnt mit einer systematisierenden und begrifflichen Eingrenzung (3.1). Denn die Beurteilung von Informationsquellen und Informationen ist mehrdimensional und erfolgt zu unterschiedlichen Zwecken. Bei aller Divergenz fokussiert die Darstellung darauf, dass Personen mit dem Beachten, Verwenden und Beurteilen von „Metadaten“ genannten Informationen über Informationen eine Annäherung beabsichtigen, möglichst glaubwürdige Quellen zu nutzen. Darum sind Vertrauenswürdigkeit und inhaltliche bzw. fachliche Ausgewiesenheit (Expertise) wichtige Bezugspunkte für Glaubwürdigkeitseinschätzungen – auch und gerade beim Lesen von Dokumenten aus dem Internet.

Wie Personen das Sourcing betreiben, wie es sich in Prozessen manifestiert und sich in Produkten niederschlägt, ist Gegenstand eines weiteren Teilkapitels (3.2). Entlang dreier aus der Theorie stammenden Gruppen von Prozessen werden Ergebnisse aus zahlreichen Studien systematisch präsentiert und entfaltet. Die heterogene Empirie lässt sich nach Fokus im Erkenntnisinteresse und Art der Operationalisierung der Sourcing-Konstrukte trennen und aufbereiten. Dabei wird sich offenbaren, dass trotz intensiver Beforschung die Prozesse immer noch unterbelichtet sind und dass die Beurteilung der Metadaten aus Produktperspektive ein sehr dominant beforschter Gegenstand ist.

Die rege Forschungsaktivität der Leseforschung und weiterer Disziplinen bringt es mit sich, dass die Frage nach individuellen Merkmalen Interesse auf sich gezogen hat. Von diesen Merkmalen werden vier behandelt: das inhaltliche Vorwissen, die thematischen Überzeugungen, die epistemischen Überzeugungen und die Lesefähigkeiten (3.3). Zudem lassen sich Individuen aufgrund verschiedener Ausprägung in individuellen Merkmalen und behavioralen Indikatoren zu Gruppen zusammenfassen, die sich in ihren evaluativen Leistungen systematisch unterscheiden.

Eine kognitiv so anspruchsvolle Fähigkeit wie die Beurteilung von Dokumenten wirkt allein aufgrund ihrer immanenten Komplexität und hohen kognitiven Ansprüche an die lesende Person fehleranfällig. Schwierigkeiten beim Sourcing sind gut belegte Phänomene, und aus dem breiten Spektrum der Studien werden zwei Problemfelder fokussiert, nämlich Fehlleistungen in der Evaluation von Metadaten bzw. Dokumenten im Allgemeinen und solchen beim lateralen Lesen im Besonderen (3.4).

3.1 Funktionen

Gemäß der Darstellung aus dem Teilkapitel 1.2 beinhaltet digitales Lesen Recherchieren und Evaluieren gleichermaßen. In diesem Kapitel und damit auch in diesem Buch wird aus Umfangsgründen aber nur ein Teil, nämlich das Evaluieren von Informationen, behandelt. Es handelt sich im Vergleich mit dem Recherchieren nicht um identische Fähigkeiten, die darum in der Fachliteratur entsprechend als voneinander getrennte Prozesse behandelt werden (Goldman & Brand-Gruwel, 2018; Rouet et al., 2020; Rouet & Britt, 2011; Salmerón, Strømsø et al., 2018). Gleichwohl besteht eine abstrahierte Gemeinsamkeit darin, dass Personen über die Verwertbarkeit von (potenziell nutzbaren) Informationen zu befinden haben. Im Falle des Recherchierens wird vor allem die Passung von einem Informationsbedürfnis und einer Interaktion von Datenbanken bzw. Suchmaschinen fokussiert, um zu potenziell verwertbaren Resultaten zu gelangen und die Ergebnisse zu selektieren. Im Fall des Sourcings geht es darum, über die Relevanz- hinaus Glaubwürdigkeitsurteile zu fällen.

Bisherige theoretische Modelle modellieren das Recherchieren als Zusammenspiel verschiedener Komponenten (Marchionini & White, 2007) bzw. als zyklisch-rekursiven Prozess (Hoyer et al., 2022; Sutcliffe & Ennis, 1998; Vakkari, 2016), der teils modular in größere Problemlöseprozesse des Lesens eingebunden ist (Rouet & Britt, 2011). Dabei spielen nicht nur die Fragen der Relevanz in den Modellen eine Rolle (Passt das Suchergebnis für die angestrebten Zwecke?), sondern auch die der grundsätzlichen Eignung von Dokumenten aus einer epistemischen Sicht: Sind die Aussagen und Informationen als mutmaßlich wahr anzunehmen und passen darum zum angestrebten Zweck?

Die Eminenz des Sourcings beim digitalen Lesen ergibt sich vor diesem Hintergrund daraus, dass sich mit dem Lesen in digitalen Medien neue Herausforderungen stellen, die aus deskriptiver Sicht das Anforderungsprofil an Leserinnen und Leser modifizieren (und erhöhen; Magliano et al., 2017) sowie aus normativer Sicht mitunter hochproblematisch sind. Zu diesen Herausforderungen zählen erstens die Verbreitung faktisch inkorrektur Aussagen, zweitens der inflationäre Umgang mit Expertise und drittens das Phänomen postfaktischer Privilegierung von Meinungen gegenüber Fakten und Wissen (Chinn et al., 2021). Sourcing kann bei einem Teil dieser Herausforderungen helfen, indem Personen gezielt

auf Metadaten achten, um indirekte Urteile über die Inhalte zu treffen, was dabei hilft, Expertinnen und Experten leichter zu identifizieren und Wissen und Fakten mit größerer Wahrscheinlichkeit als Aussagen von Personen mit hoher Expertise zu erkennen. Sourcing hat damit eine potenziell protektive Funktion, ist aber ein voraussetzungsreicher Prozess bzw. Prozessverbund.

Sourcing wird in diesem Teilkapitel primär auf eine im Umgang mit digitalen Dokumenten zentrale Funktion hin beleuchtet: den Umgang mit epistemisch unsicheren Informationen, also mit Aussagen, deren Status als Wissen in Inhalt und Genese unklar ist und darum einer Prüfung und Evaluation seitens der lesenden Person bedarf. Die Darstellung erfolgt in mehreren aufeinander aufbauenden Teilkapiteln. Zunächst wird nach einer knappen Verortung und Systematisierung der verschiedenen Bezugspunkte evaluativer Prozesse (3.1.1) das Sourcing inhaltlich bestimmt und anhand der sogenannten eingebetteten Quellen illustriert (3.1.2). Im Anschluss wird der Frage nachgegangen, welche Dimensionen für die Zuweisung von (epistemischer) Glaubwürdigkeit bestehen (3.1.3).

3.1.1 Das Beurteilen von Informationen und Informationsquellen – ein erster kursorischer Einstieg

Während die Bedeutsamkeit des Evaluierens als Prozessverbund für das digitalen Lesen in der Forschung relativ unstrittig wirkt (Salmerón, Strømsø et al., 2018), erweist es sich bei genauerer Betrachtung als sehr globaler Prozessverbund (Philipp, 2023), der sich auf verschiedene informatorische Bezugseinheiten (Dokumente, einzelne Informationseinheiten) und -punkte (Merkmale des Inhalts und/oder der Quelle) erstreckt. Diese Heterogenität ist keineswegs neu (Pressley & Afflerbach, 1995). Es gibt und gab dementsprechend Bemühungen, die Fähigkeiten zu systematisieren (Forzani, 2020; List, 2020; Meneses, 2021). Was diese Systematisierungen eint, ist eine Differenzierung der Bezugspunkte dessen, was evaluiert wird, wobei textuelle Informationen und Metadaten einen gemeinsamen Nenner in diesen Zugängen bilden. Damit ist eine grundsätzliche Trennung von Plausibilität und Glaubwürdigkeit angesprochen:

- *Plausibilität* gilt als Beurteilung des potenziellen Wahrheitsgehalts von Aussagen bzw. der wahrgenommenen Übereinstimmung von eigenem Wissen und Aussagen, es handelt sich somit um eine vorläufige Annahme zur Richtigkeit von textuellen Inhalten (Connell & Keane, 2006; Lombardi et al., 2016).
- *Glaubwürdigkeit* bezieht sich auf eine subjektive Beurteilung der Quelle von Aussagen, die sich auf mehrere quellenbezogene Facetten wie Expertise oder Vertrauenswürdigkeit erstreckt und die Akzeptanz ihrer inhaltlichen Aussagen beeinflusst (Metzger & Flanagin, 2015; Pornpitakpan, 2004; Rieh & Danielson, 2007).

Beide Evaluationen – sowohl die inhaltliche als auch die quellenbezogene – sind damit provisorische Einschätzungen. In der jüngeren Vergangenheit wurden beide Bezugspunkte der Beurteilung in neueren Theorien und Modellen prononciert behandelt (Philipp, 2023) und in Leseleistungsstudien operationalisiert. Als ein Beispiel dafür sei hier das „Programme for the International Assessment of Adult Competencies“ (PIAAC) angeführt. In dessen theoretischem Rahmen (Rouet, Britt et al., 2021) wurden die evaluativen Leistungen im Lesen nicht nur hinsichtlich der Bezugspunkte Inhalt und Quelle differenziert, sondern noch zusätzlich in verschiedene Bemessungsgrundlagen oder auch Zielpunkte untergliedert, nämlich die Akkuratheit, die Stichhaltigkeit und die Relevanz. Dies spiegelt die Darstellung in Tabelle 5 wider. Sie dient an dieser Stelle zur Veranschaulichung, dass evaluative Prozesse kein eindimensionales Konstrukt bilden, sondern sich auf diverse Einheiten beziehen können. Die quellenbezogenen Evaluationen, die sich mehr oder minder unter dem Dach der Glaubwürdigkeitseinschätzungen subsumieren lassen (Rieh & Danielson, 2007), bilden in diesem Kapitel einen Schwerpunkt der Darstellung.

Tabelle 5: Überblick über verschiedene Bezugspunkte und Bemessungsgrundlagen von Evaluationsprozessen aus der PIAAC-Studie

Bezugspunkte der Beurteilung	Bemessungsgrundlage/Ziel der Beurteilung		
	Akkuratheit	Stichhaltigkeit	Aufgabenrelevanz
Inhalt(e)	• Plausibilität • Qualität der Belege	• Vollständigkeit der Fakten und Perspektiven • Verzerrung (Bias) bei der Erklärung oder Interpretation • Interne Konsistenz	• Beitrag für das Leseziel
Quelle(n)	• Kompetenz der Autorin/des Autors • Einstellung der Person • Editoriale Kontrolle	• Verdeckte oder offene Interessen der Autorin/des Autors	• Passung/Angemessenheit des Texttyps für die eigenen Ziele und Fähigkeiten

(Quelle: leicht modifizierte Übersetzung von Rouet, Britt et al., 2021, S. 49)

3.1.2 Begriffsklärungen

Das Sourcing ist ein Verbund verschiedener Teilprozesse mit spezifischen Funktionen, die darauf abzielen, die Quelle von Aussagen als bedeutungsvolles Supplement zu den Inhalten von Texten bzw. Textteilen zu erkennen und zu behandeln. Die inhaltliche Breite des Sourcings zwingt zu Begriffsklärungen, die nach einer Definition die Bezugspunkte – Metadaten, Quellen und Dokumente – verdeutlicht und dann auf eingebettete Quellen fokussiert.

3.1.2.1 Sourcing, Metadaten, Quellen und Dokumente – zum Zusammenhang von vier Konstrukten

Bråten et al. (2017, S. 141) definieren das Sourcing „als das Beachten, Beurteilen und Verwenden von verfügbaren oder zugänglichen Informationen über die Quellen von Dokumenten“. Sie spannen damit in ihrem ersten Teil der Definition einen Prozessverbund (Beachten, Beurteilen, Verwenden) auf, der sich auf Metadaten bezieht, wie die Informationen über die Quellen bezeichnet wurden. Die Häufung der Bezugspunkte – zugängliche Informationen/Metadaten, Quellen sowie Dokumente – im zweiten Teil der Begriffsbestimmung verdient genauere Beobachtung und Einordnung. Wie nämlich Goldman und Scardamalia (2013) stellvertretend für andere Forschende in ihrem Beitrag monieren, sind die Informationseinheiten des Sourcings aufgrund unterschiedlicher Begriffsverständnisse und -verwendungen nicht immer einheitlich in Gebrauch. An dieser Beobachtung hat sich wenig geändert, im Gegenteil hat die sich stark ausdifferenzierende Forschung die Unübersichtlichkeit noch gesteigert. Die folgende kursorische Behandlung von Metadaten, Quellen und Dokumenten soll dabei helfen, die ansteigende Korngröße und die Spezifik darzustellen, beginnend von der kleinsten Einheit hin zur größten:

- *Metadaten*: Der kleinste gemeinsame Nenner des Sourcings (oftmals in Definitionen als enge Definition bezeichnet, s. z. B. Braasch et al., 2018) besteht darin, dass Sourcing sich nicht auf inhaltliche Informationen, sondern dazu querliegende Informationen erstreckt, nämlich die Metadaten. Damit sind Informationen zur Person oder zur Organisation gemeint, welche die Aussagen produziert bzw. in Umlauf gebracht hat, ebenfalls sind der zeitliche, räumliche, mediale Kontext der Aussagen, die angestrebten Adressatinnen und Adressaten sowie das Ziel der Kommunikation mitgemeint (Braasch et al., 2018; Goldman & Scardamalia, 2013; Mason & Florit, 2018). Neuere Theorien des Leseverstehens, allen voran aus dem Kontext des Verstehens multipler Dokumente in der Tradition des „Dokumentenmodells“ (s. Teilkap. [4.1.2.3](#)), gehen davon aus, dass Metadaten eine bedeutungsmodifizierende Funktion haben, weil Inhalte mit ihren Metadaten erst in der Kombination zu sozialen Artefakten innerhalb umfassender Diskurse avancieren (Britt et al., 2013). Das gilt insbesondere für den Fall, dass inhaltliche Aussagen konfligierend ausfallen und die Metadaten dafür benötigt werden, Perspektiven der Quellen und Ursachen von Dissens über die Metadaten erklärbar zu machen und dadurch Kohärenz herzustellen (Bråten & Braasch, 2018).
- *Quellen*: Mit dem Konzept der Quelle sind in der Regel weiter gefasste Verständnisse des Sourcings tangiert. Quellen werden einerseits als Ursprünge von Aussagen samt deren Veränderung durch Weiterübermittlung in medialen Kanälen betrachtet (Bromme et al., 2018), andererseits

aber auch als abgeschlossene und abgeschlossen wirkende Einheiten von Inhalten konzeptualisiert, womit einzelne Aussagen wie Zitate einer Person bis hin zu ganzen Publikationsorganen wie Zeitungen, Websites, Forschungseinrichtungen etc. gemeint sein können (Braasch et al., 2018; Rouet, 2006). Dabei geht es weniger um die Inhalte als solche denn um die durch Metadaten charakterisierbare Quellen, sodass Rouet und Britt (2022) ein Verständnis von Quelle haben, nachdem die Quelle selbst ohne ihren Inhalt beschreibbar ist, also eine eigene Entität bildet. Hinzu kommt, dass es auch zu Gewichtungen kommt, zu Evaluationen der Quellen, wenn es darum geht, die Glaubwürdigkeit von Quellen (z. B. für indirekte Plausibilitätszuweisungen) einzuschätzen (Saux, Britt et al., 2021). Quellen sind demnach eine für das Verstehen von Texten eigene Kategorie, mittels derer eine akontextuelle Aussage zu einem kontextualisierten Inhalt avanciert.

Diese Differenzierung von verschiedenen Aussagenursprüngen bzw. von Einheiten hilft wiederum aus analytischer Sicht beim Verstehen, wie einzelne Aussagen verschiedener Quellen einzelne Facetten von Situationen beschreiben. Hier sind ebenfalls Theorien des Leseverstehens multipler Dokumente instruktiv (s. Teilkap. 4.1.2.3), denn sie beschreiben, dass Quellen und Aussagen über sogenannte „Quelle-Inhalt-Verbindungen“ miteinander verknüpft und die verbundenen Elemente, Inhalte und Quelle, von den Leserinnen und Lesern auf zwei unterschiedlichen Verstehensebenen repräsentiert werden. Während die Inhalte bzw. Aussagen als Bestandteile des integrierten mentalen Modells konzeptualisiert werden, sind quellenbezogene Repräsentationen auf der Ebene des Intertextmodells zu lokalisieren. In den sogenannten „Dokumentenknotten“ des Intertextmodells werden die Metadaten repräsentiert, wobei diese dabei helfen, kohärente Quellen-Repräsentationen aufzubauen. Dies wird wiederum dafür benötigt, das intertextuell-strukturelle Verhältnis der Quellen untereinander über „Intertextprädikate“ kognitiv zu repräsentieren, nämlich gleichsam als Gerüst, ob sich Quellen widersprechen, zitieren, beipflichten etc.

- *Dokumente*: Dokumente sind die größte der drei Einheiten, da sie deutlicher auf die Verbindung von Quelle (inkl. der sie charakterisierenden Metadaten) und dem Inhalt eingehen. Sie bezeichnen den Verbund von Inhalten und Quellen als zwei Bestandteile einer abgeschlossenen Gesamtheit (Rouet & Britt, 2022). Im Alltagsverständnis sind damit Einzeltexte (mit einem erweiterten Textbegriff) gemeint; das Konzept des Dokuments rückt aber die Kombination von Texten mit ihren Quellen in den Vordergrund, was als Verarbeitungsgrenze für das Leseverstehen wichtig ist (Philipp, 2020a). Die Dokument-als-Einheit-Annahme von Britt et al. (2013) verdeutlicht dies besonders stark: Demnach sind Dokumente als sowohl inhaltlich an einem Diskurs partizipierende Artefakte zu verstehen als auch in einem Kontext entstandene Träger von Inhalten zu betrachten, z. B. historische Primärquellen,

aber auch ganze Websites, einander zitierende und widersprechende wissenschaftliche Abhandlungen, Debatten in sozialen und konventionellen Medien etc. Die Einheit von Aussagen und ihrem Entstehungs- und Wirkzusammenhang bedingen damit einander. Dies ist wiederum ein zentraler Gedanke des Verstehens multipler Dokumente, bei dem idealerweise eine lesende Person rekonstruieren kann, welchen Anteil an der Situationsbeschreibung Dokumente haben und warum ihre jeweilige Perspektive sich im Kontext anderer Dokumente auf eine spezifische Weise darstellt.

Zusammengefasst: Metadaten helfen dabei, den Ursprung und den Kontext von Aussagen – die Quelle – zu erkennen und zu beurteilen. Diese Quellen können auf verschiedenen Ebenen vorkommen, nämlich eingebettet in Texte und/oder als externer Bestandteil von Dokumenten. Sourcing fokussiert damit auf die Quellen von Aussagen variierender Granularitätsgröße.

3.1.2.2 Eingebettete Quellen – ein erhellender Kristallisationspunkt des Sourcings

Warum ist die oben getroffene Unterscheidung von Metadaten, Quellen und Dokumenten eigentlich bedeutend? Das wird dann einsichtig, wenn man die Überlappungen zum Gegenstand der Betrachtung macht, was insbesondere bei den eingebetteten Quellen auffällig ist. Eingebettete Quellen sind innerhalb eines Textes vorkommende Metadaten, die mit einer oder mehreren Aussagen verknüpft sind, indem sie inhaltlich wiedergegeben (paraphrasiert) oder zitiert wurden (Londra & Saux, 2023). Insofern kann man Quellen auf verschiedenen Ebenen lokalisieren, nämlich zum einen auf einer eher lokalen, textinhärenten Ebene als *eingebettete Quelle(n)* und zum anderen auf einer eher globalen, meist externen und den eigentlichen Text und dessen Aussage flankierenden Ebene als *Dokumentenquelle* (Strømsø et al., 2013). Diese beiden Ebenen können analytisch getrennt werden und sich sogar widersprechen, indem eine vertrauenswürdige eingebettete Quelle in einem wenig vertrauenswürdigen Dokument vorkommt und umgekehrt (s. dazu bspw. die Experimentanordnung bei Londra & Saux, 2023). Für die lesende Person bestehen hinsichtlich des Sourcings damit verschiedene Optionen: Sie kann a) entweder die eingebettete Quelle zum Gegenstand von Sourcing machen oder b) nur die Dokumentenquelle oder aber c) die eingebettete Quelle im Kontext ihres Kontexts, nämlich dem Dokument samt der Dokumentenquelle (Bråten, Strømsø & Andreassen, 2016). Ein solches Repräsentieren und (kritisches) Nutzen von mehreren Quellen ist mutmaßlich eher Expertinnen und Experten vorbehalten (Britt et al., 2013).

Aus Sicht des Leseverstehens ist die verschachtelte Mehrebenen-Repräsentation von Quellen auf lokaler und globaler Ebene von Dokumenten ein Sonderfall dessen, was in der Multiple-Documents-Literacy-Forschung als Aufbau von

umfassenden kognitiven Strukturen für das angemessene Verstehen mehrerer Dokumente modelliert wird: das „Tag all Model“ bzw. auf Deutsch „Modell mit allen Dokumentenetiketten“ (Britt et al., 1999, s. Teilkap. [4.1.2.3](#) und dort den Abschnitt „Das Dokumentenmodell als Zielvorstellung glückender intertextueller Integrationsleistungen“). In dieser besonderen Form repräsentieren Leserinnen und Leser Aussagen gemeinsam mit ihrem Ursprung, also ihren Quellen samt Metadaten in einem „Dokumentenknotten“. Sie können damit rekonstruieren, wer was geäußert hat. Dieses Modell mit allen Dokumentenetiketten war in seinem ursprünglichen Gehalt für gesamte Dokumente mit einer Dokumentenquelle vorgesehen, aber es lässt sich in seiner Logik auch auf kleinere Einheiten wie eingebettete Quellen anwenden, wenn es darum geht, Diskurse inhaltlich zu repräsentieren, die auf verschiedenen Äußerungen verschiedener Quellen basieren. Weiter gedacht können damit eingebettete Quellen innerhalb einzelner Dokumente als Auslöser für Dokumentenknotten fungieren, mittels derer Aussagen auf einzelne Quellen innerhalb von Dokumenten zurückgeführt werden (Britt et al., 2013). Ergo ist eine potenzielle Gleichartigkeit von Verstehensprozessen bereits bei einzelnen Texten angesprochen, welche sonst als prototypische Anforderung für das Leseverstehen multipler Dokumente gelten (Philipp, 2018b, 2020a; Stadler, 2017).

Für die Modellierung von Sourcingprozessen ist vor diesem Hintergrund von Belang, wie Personen damit umgehen, wenn sie sowohl eingebettete als auch Dokumentenquellen vorfinden. Solche Studien sind dominant für Produktmaße vorgelegt und mit Studierenden als Samples durchgeführt worden, wobei die Themen der Dokumente den teilnehmenden Testpersonen in aller Regel eher unvertraut waren. Im *Prozess* des Lesens wenden die Personen ihre Aufmerksamkeit gemäß Lautdenk-Protokollen mehr den eingebetteten Quellen als den Dokumentenquellen zu, wobei sich dies vor allem auf das Erkennen von Metadaten bezieht und weniger auf die Evaluation (Strømsø et al., 2013). Diese stärkere Zuwendung zu den eingebetteten Quellen im Prozess schlägt sich aber nicht aus einer *produktbezogenen* Perspektive konsequent nieder. Denn es ergeben sich teilweise bei kognitiven Leistungen eher verstärkte Zuwendungen zu den Dokumentenquellen: Die in den Studien untersuchten Personen

- erinnerten sich besser an die Dokumentenquellen als an die eingebetteten Quellen (Bråten, Strømsø & Andreassen, 2016);
- zitierten mehr die Dokumentenquellen als eingebettete Quellen (Le Bigot & Rouet, 2007; Strømsø & Bråten, 2014; Strømsø et al., 2013) bzw. zitierten sie sie ähnlich häufig (Salmerón, Delgado & Mason, 2020; Salmerón, Gil & Bråten, 2018a);
- waren nach einem Training dazu in der Lage, vertrauenswürdige eingebettete Quellen häufiger anzuführen, die aus Dokumenten mit zuverlässigen Dokumentenquellen stammen (Londra & Saux, 2023, Experiment 2),

sie differenzierten also stärker die Passung von zitierfähigen eingebetteten Quellen;

- hatten teils Schwierigkeiten, eingebettete Quellen korrekt als Bestandteil von Dokumenten in Form von Referenzierungen in eigenen Texten wiederzugeben, sie behandelten sie mithin (unzulässig) als eigenständige Quellen (Strømsø & Bråten, 2014); zudem korrelierten die Anzahlen der Referenzierungen von Dokumentenquellen und eingebetteten Quellen sogar negativ, was bedeutet, dass Personen eine Form von Referenzierungen zulasten einer anderen tätigten (Salmerón, Gil & Bråten, 2018a);
- nutzten für Begründungen für Glaubwürdigkeitsurteile eingebettete und Dokumentenquellen in ähnlicher, nämlich im Vergleich zu anderen Beurteilungsgrundlagen ausgesprochen geringem Maß (Strømsø et al., 2017).

Es scheint also laut Grundlagenforschung eine Privilegierung von Dokumentenquellen bei den Produktmaßen bei der Benennung von Quellen zu geben. Gleichwohl verwenden Personen beide Varianten von Quellen und – so die Ergebnisse aus dem zweiten Experiment von Londra und Saux (2023) – durchaus sensibel, wenn es eine vorgängige Schulung gab. Trotz der ausbaufähigen Empirie existieren somit Hinweise darauf, dass Personen eingebettete Quellen und Dokumentenquellen zwar auf niedrigem Niveau, aber gleichwohl unterschiedlich nutzen. Das ist deshalb aus der Perspektive des Verstehens von (digitalen) Dokumenten so wichtig, weil eingebettete Quellen in ihrer Perspektive innerhalb eines Kontextes (dem Dokument) zu betrachten sind und die Rekonstruktion dieser Perspektive mittels Sourcing als einem prozessualen Bestandteil dabei hilft, umfassende mentale Repräsentationen aufzubauen (Barzilay & Weinstock, 2020). Außerdem kann die genaue Prüfung der eingebetteten Quellen als Bestandteil einer adaptiven epistemischen Leistung betrachtet werden (Chinn et al., 2021), die als ein Erfordernis im Umgang mit KI-basierten Textangeboten einzuschätzen ist (Kasneci et al., 2023). Anders gesagt: Der Umgang mit eingebetteten Quellen aus einer Sourcingperspektive, sei es bezogen auf die inhaltliche Adäquanz der Aussagen dieser Quellen, sei es bei der redlichen Wiedergabe durch andere Quellen, sei es bei der Rekonstruktion von multiperspektivischen Zugängen (durch andere Prüfprozesse, s. Philipp, 2023) erscheint als ein momentan noch wenig bearbeitetes Forschungsfeld, aber zugleich eine reallebensweltlich hochbedeutende Anforderung.

3.1.2.3 Die Multifunktionalität des Sourcings

Das Sourcing ist ein Prozessverbund des Beurteilens und Nutzens von Metadaten, welcher mehrere Funktionen erfüllt und deshalb als Regenschirm-Begriff fungiert (Scharrer & Salmerón, 2016). Diese Multifunktionalität schlägt sich bereits in der von Bråten et al. (2017) stammenden Definition des Sourcings aus

dem Teilkapitel [3.1.1](#) in der Prozesskomponente nieder, denn es sind mit dem Beachten, Beurteilen und Verwenden der Metadaten nicht nur mehrere Prozesse angesprochen, sondern vor allem bei dem Beurteilen und Verwenden prinzipiell viele Funktionen und Zwecke impliziert. Das Sourcing ist speziell bei Leseaktivitäten im Internet von hoher Bedeutung, da es jederzeit im Gesamtprozess vorkommen kann und den kontingenten Leseprozess dynamisiert (Salmerón, Strømsø et al., 2018): sei es bei der Auswahl von potenziellen Texten, sei es vor der eigentlichen Recherche bei der Wahl der geeignet wirkenden Suchmaschine, sei es dann innerhalb der Suchmaschinenergebnisse, sei es bei der Möglichkeit, eingebettete Quellen auszuwählen etc. Anders gesagt: *Sourcing ist keine diskrete Phase in einem linearen Prozess, sondern eine variabel einsetzbare Komponente in einem iterativen Problemlöseprozess und besteht auch aus empirischer Sicht aus partiell deutlich trennbaren Bestandteilen* (Braasch et al., 2023; Braasch, Bråten, Strømsø & Anmarkrud, 2014; Bråten, Salmerón & Strømsø, 2016; Cho, Han & Kucan, 2018; Dyoniziak et al., 2023; Incognito & Tarchi, 2024; Kiili et al., 2021; Kiili, Leu, Utriainen et al., 2018; Kobayashi, 2014; Paul et al., 2018, Studie 1). In diesem Problemlöseprozess erfüllt das Sourcing diverse Funktionen, wie es die exemplarische und keinen Anspruch auf Vollständigkeit hegende Übersicht aus Tabelle 6 verdeutlicht. Aus ihr geht hervor, dass Sourcing sowohl für das Finden als auch für das Auswählen von glaubwürdig wirkenden Informationen und für den Aufbau angemessenen kohärenter Repräsentationen nötig ist.

Tabelle 6: Prozesse der Online-Recherche und die ihnen inhärenten Sourcingprozesse

Prozess der Online-Recherche	Sourcingprozesse
Planung der Recherche	Nachdenken über potenzielle Informationsquellen, wer über das Thema und in welchen Publikationsorganen berichten könnte
Lokalisieren von Informationen	Verwendung potenzieller Informationsquellen zur Formulierung von Suchanfragen
Beurteilen von Informationen	Nutzung der verfügbaren Metadaten einer Suchergebnisseite, um die Nützlichkeit und Glaubwürdigkeit von Quellen zu beurteilen und dadurch eine Auswahl von Dokumenten zu treffen
Synthetisierung von Informationen aus verschiedenen Online-Dokumenten	Verwendung von Metadaten zur Bewertung der Glaubwürdigkeit von geeigneten Quellen
Kommunikation von Informationen	Aufbau einer kohärenten Darstellung des Themas durch Vergleich und Gegenüberstellung von Informationen, die von verschiedenen Quellen stammen

(Quelle: Kiili, Brante et al., 2020, S. 177, leicht modifizierte Darstellung)

Die Multifunktionalität des Sourcings, welche in Tabelle 6 angerissen ist, ist häufig Gegenstand einzelner Studien, die sich auf spezifische Teile der einzelnen Prozesse und auch Phasen beziehen und darum nur fragmentarische Einblicke erlauben. Eine Ausnahme bildet eine Untersuchung, die im Kasten „Illustrierendes Beispiel 10“ dargestellt wird.

Illustrierendes Beispiel 10

Der Nutzen des durchgängigen Sourcings vom Klären des Informationsbedürfnisses bis zur schriftlichen Antwort auf der Basis einer Internetrecherche

A) Was war das Erkenntnisinteresse?

Die Studie (Kiili et al., 2021) widmete sich der Frage, in welchem Ausmaß Sourcingaktivitäten in verschiedenen Phasen einer Internetrecherche stattfinden und wie sich diese Sourcingaktivitäten aus vorhergehenden Phasen als prädiktiv für spätere Sourcingleistungen erweisen.

B) Wer wurde untersucht?

Die Teilnehmerinnen und Teilnehmer waren 167 finnische Schülerinnen und Schüler der Sekundarstufe, die im Mittel 17 Jahre alt waren. Der Großteil des Samples bestand aus Schülerinnen (57 Prozent).

C) Welche Instrumente kamen zum Einsatz?

Die Studie erfasste mittels produktbezogener Daten das Sourcing in verschiedenen Phasen einer Problemlösung. Die Studie nutzte als Stimulus ein kontroverses Thema, nämlich das Impfen von Kindern, bei dem die Schülerinnen und Schüler auf die Frage zu einer Impfpflicht schriftlich Stellung beziehen sollten. Dafür wurde eine digitale Lernumgebung verwendet, die sequenziell den Aufgabenbearbeitungsprozess strukturierte und die insgesamt vier Sourcingindikatoren elizitierte.

- Der erste Indikator wurde über die *Spezifikation des Informationsbedürfnisses* generiert. Dazu schrieben die Testpersonen, welche Art von Informationen nötig seien, um die Frage zur Impfung zu beantworten. Die Antworten wurden auf einer dreistufigen Skala nach sourcing-spezifischen Bestandteilen (wie Nennung von geeigneten Quellen wie Institutionen oder Expertinnen bzw. Experten, aber auch Hinweise auf evaluative Einschätzungen) kodiert.
- Der zweite Indikator wurde aus den *Suchanfragen* extrahiert. Dafür wurden die Anfragen danach kodiert, in welchem Ausmaß sie bereits Metadaten beinhalteten, also ob die Jugendlichen gezielt nach Internetquellen gesucht hatten, die sie über Metadaten in der Suchmaschine eingrenzten. Je mehr solche sourcing-basierten Elemente die Anfragen enthielten, desto höher war der Punktwert.
- Der dritte Indikator betraf die *Einschätzungen der Glaubwürdigkeit* jener drei Dokumente, welche die Schülerinnen und Schüler für ihre Antwort nutzen wollten. Für jedes Dokument sollten sie in einer offenen Frage jeweils einschätzen, welche Faktoren die

Glaubwürdigkeit erhöhen bzw. minimieren. Die Antworten wurden nach Benennung von Autorinnen/Autoren, deren Intention und nach Einschätzung der Dokumentenquelle kodiert. Das Vorhandensein dieser drei Elemente wurde ausgezählt und über die drei Websites zu einem Gesamtwert addiert.

- Der vierte produktbezogene Indikator stammte aus den *schriftlichen Antworten*, die mehrdimensional kodiert wurden. Das Ausmaß von Benennung und Verwendung von Metadaten nebst Evaluationen der Quellen wurden in einem Ordinalsystem zu einem Gesamtscore verdichtet.

Neben diesen Sourcingmaßen wurden separat die Leseflüssigkeit und das Vorwissen als Kovariaten ermittelt.

D) Was waren die Hauptbefunde?

Der erste eher deskriptive Befund bestand darin, dass in allen Phasen Sourcing auftrat. Dabei erwiesen sich gemessen an der erreichbaren Maximalpunktzahl der einzelnen Sourcingmaße die Mittelwerte aber unterschiedlich ausgeprägt. Beim Informationsbedürfnis betrug der Wert umgerechnet 38 Prozent, im Sourcing der Suchanfragen waren es nur 15 Prozent. Die Evaluationsleistungen lagen ebenfalls bei 38 Prozent, die Gesamtleistung im Produkt bei 43 Prozent.

Die Studie nutzte zwei Regressionsanalysen, in denen die beiden Kovariaten einfließen. Die Regressionen behandelten zum einen das Sourcing in den evaluativen Einschätzungen der drei Quellen (dritter Indikator) und zum anderen den Gesamtscore zum Sourcing in der schriftlichen Antwort (vierter Indikator) als abhängige Variable. Als unabhängige Variablen wurden die chronologisch vorgängig gesammelten Sourcingmaße verwendet. Die Ergebnisse zeigen, dass das unmittelbar vorgängige Sourcing immer prädiktiv war. Das bedeutet: Je mehr sourcingbasierte Indikatoren beim Informationsbedürfnis vorgekommen waren (erster Indikator) und je spezifischer mit Sourcingmaßen die Suchanfragen formuliert worden waren (zweiter Indikator), desto mehr Kriterien nutzten die Jugendlichen für ihre Vertrauenswürdigkeitseinschätzungen (dritter Indikator). Für den Gesamtscore des Sourcings in der finalen Antwort (vierter Indikator) erwies sich nur die Glaubwürdigkeitseinschätzung (dritter Indikator) als positiver Prädiktor, nicht aber mehr die ersten beiden Sourcingmaße.

E) Diskussion

Sourcingindikatoren scheinen wie in einer Kettenlogik zu bestehen: Jeweils vorherige Sourcingleistungen präzisieren ihnen folgende. Damit wirkt es so, als vollziehe sich Sourcing einerseits kontinuierlich, andererseits auch spezifisch in seinen Beiträgen: Es hilft bei der Eingrenzung geeigneter Quellen, manifestiert sich in Urteilen zu den gefundenen Dokumenten und schließlich in kurzen eigenen Texten, welche das Sourcing bei einem kontroversen Thema inkludieren.

Was sich in dieser nur knappen Illustration anhand der Online-Recherchen andeutet – die Multifunktionalität des Sourcing –, ist inzwischen Gegenstand ausführlicherer Theoriearbeit in der Leseforschung geworden (Bråten et al., 2017; Philipp, 2020a), die zugleich mit einer Aufwertung des Sourcing als essenziellem Verstehensprozess einhergeht (Philipp, 2021b). Verschiedene theoretische Modelle beinhalten das Sourcing als einen Bestandteil, der verschiedenen Zwecken dient und der mit anderen Prozessen interagiert. Aus einer vergleichenden Sichtung von sieben Modellen hat Philipp (2023) drei Funktionen des Sourcing – 1) umfassendes Verstehen, 2) Vermeiden von Inkohärenz sowie 3) Überwinden verzerrter Repräsentationen – extrahiert, die im Folgenden dargestellt werden.

Funktion 1: Verstehen selbst widersprüchlicher Aussagen über Sourcingprozesse auf der Ebene von Intertextmodellen und über den Aufbau von Quelle-Inhalt-Verbindungen

Leserinnen und Leser sind oftmals mit inhaltlich widersprüchlichen Aussagen konfrontiert, die einen Kohärenzbruch und ein Hindernis für kohärente Gesamtvorstellungen darstellen. Wer sich über das Ausmaß von Impfnebenwirkungen im Internet informieren will, stößt beispielsweise auf konträre Aussagen, welche die Gefahren unterschiedlich darstellen. Das „Dokumentenmodell“ (Britt et al., 1999) hilft dabei, den inhaltlichen Dissens innerhalb der kognitiven Repräsentation aufrechtzuerhalten, ohne dass daran der Kohärenzaufbau scheitert. Denn die Positionen bzw. Inhalte der Quellen werden auf einer separaten mentalen Ebene, nämlich jener des integrierten mentalen Modells, im Arbeitsgedächtnis repräsentiert. Hinzu kommt eine zweite Ebene, für das Sourcing unmittelbar wichtig ist: die des Intertextmodells. Diese Ebene ist den Quellen bzw. den Dokumenten vorbehalten, die gemäß der Dokumentenmodell-Architektur als separate Einheiten repräsentiert und mit sogenannten Quelle-Inhalt-Verbindungen mit dem Inhalt verquickt werden. Dadurch können Leserinnen und Leser die Dokumente bzw. Quellen als Ursache für Dissens in ihrer mentalen Repräsentation aufrechterhalten (z. B. „Quelle A hält Impfnebenwirkungen für sehr gefährlich, Quelle B hält sie für sie vertretbar und risikoarm“), sie können die Quellen über sogenannte „Intertextprädikate“ per Inferenzen verbinden („Quelle A und B widersprechen sich“). Über die Quelle-Inhalt-Verbindungen kann der Dissens auf der Ebene des integrierten mentalen Modells („Impfnebenwirkungen werden unterschiedlich dargestellt“) über die Dokumente bzw. Quellen erklärt werden („Die divergierende Darstellung der Impfnebenwirkungen ergibt sich daraus, dass Dokument/Quelle A Impfnebenwirkungen für gefährlich hält, während Dokument/Quelle B als vertretbar bezeichnet“). Diese Mehrebenen-Repräsentation ist produktbezogen und wird als Ergebnis von Integrations- und Sourcingprozessen beschrieben (s. dazu ausführlicher Teilkap. [4.1.2.3](#)).

Aus einer *Prozesswarte* wird Sourcing in verschiedenen Modellen als ein bedingter, d. h. als ein Folgeprozess postuliert. Die Leserinnen und Leser nutzen das Sourcing (als Folge), wenn sie im Leseprozess einen auflösungsbedürftig wirkenden Dissens wahrnehmen (als Voraussetzung). Die einzelnen Modelle aus der Leseforschung nehmen diesen Grundmechanismus an, stellen ihn aber unterschiedlich dar. Die einfachste Variante dieses Mechanismus' wird im „Discrepancy-Induced Source Comprehension Model“ (DISCM) beschrieben (Braasch & Bräten, 2017; Braasch & Kessler, 2021). Das Modell nimmt an, dass Personen bei auflösungsbedürftigen Inkonsistenzen infolge einer Diskrepanz zwischen Aussagen einen Kohärenzschwellenwert unterschreiten, was zu kognitiver Dissonanz führt. Ein solcher Kohärenzschwellenwert ist nicht mit dem Kohärenzstandard gleichzusetzen, der einen angestrebten Zustand des Verstehens Erfolgs bezeichnet (s. Teilkap. 2.1.2). Der Kohärenzschwellenwert ist ein kontinuierlich im fortschreitenden Verstehensprozess auftretender Benchmark ausreichender Widerspruchsfreiheit, der ohne strategisches Zutun der lesenden Person über automatisch ablaufende Gedächtnisprozesse genutzt wird (O'Brien & Cook, 2016b). Personen gleichen also fortwährend beim Lesen die Widerspruchsarmut ab und unterbrechen automatisch oder gezielt bei einem Unterschreiten den aus verschiedenen wellenförmig überlappenden Teilprozessen bestehenden Leseprozess, um die Diskrepanz bzw. Dissonanz zu beheben (Singer, 2019). Um diese Dissonanz aufzulösen, wenden Personen gemäß DISCM das Sourcing an, um mindestens die Quelle-Inhalt-Verbindungen zu erzielen oder aber weitere kohärenzstiftende Prozesse mittels der Metadaten wie Inferenzen umfassendere mentale Repräsentationen wie z. B. das „Dokumentenmodell“ zu entwickeln. Das DISCM beschreibt damit ein simples Wenn-dann-Gefüge und hat über den unterschrittenen Kohärenzschwellenwert ein individuell variierendes Soll an Kohärenz als Auslöser für Sourcingprozesse. Das DISCM postuliert dabei lediglich, dass Sourcing stattfindet, aber nicht genauer, wie es zur Überwindung der Diskrepanz kommt. An dieser Stelle setzen weitere Modelle an.

Das „Conceptual Model of Accuracy Determination“ (CMAD; Afflerbach et al., 2014) und das „Content-Source Integration Model“ (CSIM; Braasch & Scharrer, 2020; Stadtler & Bromme, 2014 – ähnlich auch das Modell „Integrated Framework of Multiple Text Use“ (IFMTU) von List und Alexander, 2019, welches weite Teile des CSIM in sich integriert) fokussieren diesen Mechanismus ebenfalls. Sie behandeln das Sourcing als einen von mehreren Bestandteilen in einem gesamthaft sehr dynamischen Gesamtprozess. CMAD und CSIM beschreiben einen ähnlichen Umgang mit der Diskrepanz, die im Einklang mit den Annahmen zum Dokumentenmodell stehen: Wenn die Diskrepanz zwischen Aussagen zu groß ist, um sie aufgrund eigenen Vorwissens (oder mithilfe anderer Quellen) begründet aufzulösen, erhalten Leserinnen und Leser auf der Ebene des integrierten mentalen Modells die inhaltliche Diskrepanz aufrecht und führen sie auf den Dissens von Dokumenten-/eingebetteten Quellen zurück. Sie bilden

damit im Optimalfall ausreichend komplexe Intertextmodelle und Quelle-Inhalt-Verbindungen und repräsentieren somit nicht nur die (partiell) konfligierenden Inhalte, sondern auch vor allem die über Dokumentenknoten abgebildeten Quellen in ihrem Verhältnis untereinander hinsichtlich der inhaltlich divergierenden Perspektiven. Das CMAD reserviert ein derartiges sourcingbasiertes Vorgehen ausdrücklich für substanzielle Konflikte, während das CSIM es für jene Fälle vorsieht, in denen Leserinnen und Leser es ausschließen, dass der Dissens auf rein inhaltlichen Gründen basiert.

Einen Sonderfall in den Modellen, bei denen es um die Rolle des Sourcings für möglichst vollständige Repräsentationen geht, bildet das „Cognitive Affective Engagement Model of Multiple Source Use“ (CAEM; List & Alexander, 2017b, ebenfalls wieder aufgenommen im IFMTU von List und Alexander, 2019). In diesem Modell wird Sourcing einerseits als Prozess beschrieben, andererseits wird es als individuelle Disposition im Umgang mit multiplen Dokumenten behandelt. Als stabile Disposition wird Sourcing erst im Verbund mit weiteren motivationalen Personenmerkmalen als Bedingung dafür modelliert, dass tiefergehende und durchaus kritisch-evaluative Sourcingprozesse zu einem gesamthaft tiefen Verstehen multipler Dokumente führen. Damit ist Sourcing nicht nur als behaviorale Größe, sondern auch als relativ stabile personale Eigenschaft Gegenstand im CAEM, wodurch sich die Perspektive auf das Sourcing ausweitert, aber davon unbenommen bleibt, dass es erst über das Sourcing glückt, umfassende Dokumentenmodelle kognitiv zu repräsentieren.

Insgesamt ist Sourcing als zwar eigenständiger, aber im Verbund wirkender Prozess der Kohärenzbildung und -aufrechterhaltung also Gegenstand diverser lesetheoretischer Modellierungen. In dieser Schnittmenge offenbart sich, dass ihm eine wichtige Funktion zukommt, bei komplexen und mehrperspektivischen Sachverhalten ein Gesamtverständnis zu unterstützen, indem die Repräsentation der Quellen als zusätzliche Verstehensebene realisiert wird.

Funktion 2: Inkohärenzvermeidung durch Abwahl und Gewichtung wenig belastbarer Dokumente/Quellen

Das Sourcing fokussiert nicht nur das in der Eingangsdefinition im Teilkapitel [3.1.1](#) angeführte Erkennen von quellenbezogenen Metadaten und deren Verwendung für den Aufbau komplexer Mehrebenen-Repräsentationen als Verstehensprodukte. Vielmehr ist mit dem Beurteilen ein weiterer Prozess angesprochen, der bei der Zuweisung von Glaubwürdigkeit nötig ist und wiederum teilweise die Funktion erfüllt, indirekte Plausibilitätseinschätzungen in komplexen (Wissens-) Gesellschaften vorzunehmen (Kienhues et al., 2020; Kitcher, 2011). Die von den Leserinnen und Lesern zugewiesene Glaubwürdigkeit ist ein mehrdimensional abstützbares Urteil, dass eine Quelle redlich, korrekt und aus epistemischer Sicht zuverlässig informieren kann und dass deshalb ihre Aussage plausibel ist

(Bromme et al., 2018). Bei dem eingangs schon erwähnten Beispiel der widersprüchlichen Aussagen zu den Impfnutzenwirkungen können Personen den Konflikt gemäß der ersten Funktion erkennen und auf die Quellen attribuieren. Ist mit dem Lesen das Ziel einer Entscheidung verbunden, um beispielsweise die Impfung begründbar als Option (nicht) zu wählen, sind evaluative Prozesse nötig, die eine interpretatorische Leistung erfordern, um die Quellen zu bewerten. So können die Leserinnen und Leser beispielsweise darauf achten, ob die Aussage von einer medizinisch ausreichend ausgebildeten Person stammt, ob sich die Person auf die Art von Informationen stützt, die in der Domäne Medizin als angemessen gelten, ob sie in einem seriös wirkenden Kontext erscheint etc.

Die Verwertbarkeit von Quellen aufgrund ihrer zugeschriebenen Glaubwürdigkeit seitens der Leserinnen und Leser wird naturgemäß in jenen Modellen fokussiert, welche eine Entscheidung zur Wahl bzw. Abwahl resp. zur Gewichtung von Quellen und ihren Aussagen akzentuieren. Hierbei sind zwei Modelle anzuführen, nämlich auf der einen Seite das CMAD (Afflerbach et al., 2014) und auf der anderen Seite das CSIM (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014). Beide Modelle gehen von einer quellenbasierten Evaluation aus:

- Das CMAD benennt hierfür einen nicht näher explizierten „Akkuratheitsstandard“, welcher einen Schwellenwert für die eingeschätzte Plausibilität von Aussagen betrifft. Das Sourcing ist dabei ein indirekter Weg, die Akkuratheit von Aussagen mittels Vertrauenswürdigkeitsurteilen zu unterfüttern. Das CMAD geht freilich davon aus, dass diese Herangehensweise an Grenzen stößt, sobald die Unsicherheit in substanziellen Inhalten zu groß wird, um sich alleinig auf das Sourcing zu stützen. Bei (vernachlässigbar wirkenden) Details dient das Sourcing gemäß CMAD hingegen als Option, indirekte Glaubwürdigkeitsurteile zugunsten der vertrauenswürdigeren Quelle zu treffen. Insofern bildet das Sourcing eine eher begrenzt wirksame Heuristik (Afflerbach et al., 2014).
- Das CSIM positioniert das Sourcing als „Zweite-Hand-Einschätzung“ bei der sogenannten *Konfliktregulation* (Stadtler & Bromme, 2014) resp. *Evaluation* konfligierender Aussagen (Braasch & Scharrer, 2020) als finale Phase. Innerhalb dieser Phase können Personen sowohl die epistemische Stärke von Aussagen qualifizierend einschätzen als auch die epistemische Vertrauenswürdigkeit der Quellen, um auf dieser Basis Aussagen oder ganze Dokumente höher oder niedriger zu gewichten.

Die beiden Modelle nehmen damit komplementäre Perspektiven auf die Wahl und Abwahl resp. Taxierung von Quellen ein. Das CMAD setzt zeitlich früher an, ohne dass es einer strengen Phasenlogik folgt, dabei aber den Geltungsraum

des Sourcing limitiert, indem es das Sourcing auf lokale intertextuelle Inkonsistenzen einschränkt. Das CSIM nimmt hingegen gravierenderen Dissens und die Möglichkeit, ihm mit dem Sourcing zu begegnen, in den Blick.

Funktion 3: Überwinden eigener verzerrter Überzeugungen bzw. Wissensbestände durch Berücksichtigung vertrauenswürdiger Inhalte

Die Leseforschung hat sich zunehmend damit auseinanderzusetzen begonnen, in welchem Zusammenhang eher problematisch zu nennende Langzeitgedächtnisinhalte und mentale Modelle zueinanderstehen. Am augenfälligsten ist dies beim unkritischen Umgang mit Desinformation, bei dem unter anderem weltanschauliche Faktoren im Sinne von Überzeugungen (etwa die Einstellungen zu Impfungen) zu verzerrten Repräsentationen führen (Ecker et al., 2022). Dies wird in einem kognitionspsychologischen Teil der Leseforschung unter dem Stichwort „Text-Überzeugungskonsistenz-Effekt“ gesondert behandelt (Wolfe & Griffin, 2017, s. dazu Teilkap. 4.3.2). Diese Forschung befasst sich damit, wie eigene thematische Überzeugungen dazu führen, dass überzeugungsinkonsistente Inhalte aus Texten ohne Korrektur weniger stark in umfassenden mentalen Repräsentationen vorkommen – oder anders formuliert: Personen entwickeln einseitig verzerrte Repräsentationen, die zu ihren Einstellungen passen. Positiv zu Impfungen eingestellte Personen werden demnach eher nach dem Lesen Inhalte in ihrem mentalen Modell integriert haben, die das Impfen unterstützen; bei Personen mit negativer Einstellung gegenüber dem Impfen überwiegen entsprechend die dazu passenden Inhalte in den Repräsentationen. Um diesen Bias zu überwinden, ist gemäß dem „Two-Step-Model of Validation“ (TSMV) ein „strategisches Elaborieren“ nötig (Richter & Maier, 2017). Damit ist ein Bündel diverser Vorgehensweisen gemeint, welche auf die Prüfung der Plausibilität überzeugungsinkonsistenter Aussagen und der Glaubwürdigkeit der dazugehörigen Quellen abzielen. Das Sourcing bildet ergo im TSMV eine unter mehreren Optionen (Richter & Tiffin-Richards, 2024), wobei diese Vorgehensweisen an diverse kognitive und motivationale, aber teils auch kontextuelle Bedingungen geknüpft sind (Richter & Maier, 2017). Das strategische Elaborieren im Allgemeinen und das Sourcing im Besonderen dienen mithin dazu, überzeugungsinkonsistente Inhalte bzw. deren Quelle zu prüfen und auf dieser Basis mehrperspektivische Repräsentationen aufzubauen (Richter et al., 2020).

Einen anderen Zugang als das TSMV wählt das „Knowledge Revision Components Framework – Multiple Documents“ (KReC-MD; Butterfuss & Kendeou, 2021). Es entstammt paradigmatisch der Forschung zu Widerlegungstexten („Refutation Texts“; Kendeou, 2024), in denen normativ unerwünschte Gedächtnisinhalte bzw. Fehlkonzepte (wie unrealistische Überzeugungen zu Impfnebenwirkungen) über eigens verfasste Widerlegungstexte (wie Faktenchecks zu Impfnebenwirkungen) aufgegriffen, aber gezielt mit inhaltlich zutreffenden Aussagen

so kontrastiert werden, dass die individuellen Überzeugungen bzw. Fehlkonzepte dadurch allmählich reduziert und abgeschwächt werden (Kendeou & O'Brien, 2014; Schroeder & Kucera, 2022). Das KreC-MD spezifiziert diese Wirkweise auf der Basis verschiedener aufeinander aufbauender Prinzipien und setzt gezielt das Sourcing ein (Butterfuss & Kendeou, 2021). Es postuliert, dass mehrere Quellen inhaltlich übereinstimmend oder zumindest doch komplementär die eigentliche Widerlegung beinhalten. Die vertrauenswürdigen Quellen und die intertextuelle Widerspruchsarmut – sprich: ein maximal kohärentes Dokumentenmodell – dienen also im Verbund dazu, die Wissensrevision zu ermöglichen.

KReC-MD und TSMV sind damit nicht deckungsgleich, doch sie beziehen sich auf die Korrektur kognitiver Langzeitgedächtnisinhalte, die sich nicht einfach substituieren lassen. Vielmehr fungieren in beiden Modellen die Sourcingprozesse als lesestrategische Varianten, um Inhalte anzuerkennen, welche eigenen Überzeugungen (TSMV) oder fehlerhaften subjektiven Wissensbeständen (KReC-MD) widersprechen. Das KReC-MD geht hierbei so weit, dass es nicht nur zu einer ausgewogenen Repräsentation der Inhalte kommen muss, sondern diese auch in einer Modifikation eigener Gedächtnisinhalte mündet.

3.1.3 Sourcing: mehrdimensionale und vor allem vernetzte Glaubwürdigkeitsevaluationen

Das Sourcing ist nach den bisherigen Ausführungen ein multifunktionaler Prozessverbund. In diesem Verbund ist das Erkennen von Metadaten zur Quelle ein Bestandteil unter mehreren. Insbesondere die zweite Funktion des Sourcings aus dem Teilkapitel [3.1.2.3](#) fokussiert auf die Beurteilung der Metadaten. Damit stellt sich die Frage, was die Bemessungsgrundlage für Quellen ist, wofür es verschiedene kognitions- sowie sozialpsychologische Erklärungsansätze gibt (s. dazu im Detail Bråten et al., 2017). Die Mehrdimensionalität wird in diesem Teilkapitel in drei Schritten behandelt und entfaltet. Im ersten Schritt dient das AIR-Modell dazu, Sourcing als einen Bestandteil umfassender Konzeptionen zum Umgang mit Wissensangeboten zu rahmen ([3.1.3.1](#)). Der zweite Schritt bezieht sich auf zwei Dimensionen von Einschätzungen der Glaubwürdigkeit, nämlich die Expertise und die Vertrauenswürdigkeit von Quellen, samt einem ausführlichen empirischen Exempel, welches zeigt, dass Sourcing dazu dient und dabei hilft, komplexe Fragen erfolgreich zu bearbeiten, wenn Personen digitale Dokumente nutzen ([3.1.3.2](#)). Der dritte Schritt fokussiert auf ein Phänomen des Lesens im Internet, nämlich das laterale Lesen ([3.1.3.3](#)). Eine Zusammenfassung rundet das Teilkapitel und die Darstellung ab ([3.1.3.4](#)).

3.1.3.1 Ein theoretischer Rahmen: das AIR-Modell und das Sourcing als Teil eines Prozessverbunds zum Erreichen epistemischer Ziele

Die zentrale Frage beim Sourcing lautet: Woran macht sich Glaubwürdigkeit fest? Eine mögliche Antwort darauf offerieren Chinn und Rinehart (2016a), indem sie auf eine erkenntnistheoretische und damit eine philosophische Begründung des Reliabilismus zurückgreifen: *Die Glaubwürdigkeit einer epistemischen Quelle bemisst sich daran, dass sie sich auf epistemisch zuverlässige Prozesse stützt, um zu ihrer Aussage zu gelangen.* Einschätzungen, die auf die epistemische Zuverlässigkeit der Quelle abzielen und auf ihnen basieren, setzen also Analysen und damit Wissen voraus, die sich nicht auf singuläre Metadaten stützen, sondern darüber hinausgehend die Wissensproduktion fokussieren und damit voraussetzungsreich sind. Eine theoretische Basis dafür stellt das AIR-Modell von Chinn et al. (2014) dar. Es ist kein auf das Sourcing enggeführtes Modell, sondern basiert auf einem breiten Verständnis von epistemischen Kognitionen und deren Bedeutung für das Erreichen wissensbezogener Ziele wie dem Verstehen, Wissenserwerb, dem Nachvollzug von Erklärungen etc. (s. Philipp, 2024b, für einen kursorischen Überblick). Die im Akronym enthaltenen zentralen Bestandteile des Modells sind:

- „Aims and Values“ (Ziele und Werte): Damit sind individuelle, mit dem Lesen konkret verbundene Ziele (Aims) gemeint, welche sich auf jedwedes motivational grundierte Ziel erstrecken können. Grob unterscheiden lassen sich epistemische Ziele von nicht-epistemischen Zielen, wobei erstgenannte sich auf tieferes Verstehen, Explorieren, den Erwerb gerechtfertigter Überzeugungen oder die Vermeidung falscher Inhalte beziehen können (Chinn et al., 2011). Auf den allgemeinsten Nenner gebracht definieren Chinn und Rinehart (2016b, S. 461) epistemische Ziele als solche „Ziele, die sich auf die Entwicklung einer Art von Repräsentation der Welt beziehen – die Entwicklung einer kognitiven ‚Sichtweise‘ der Welt“. Bezogen auf das Lesen ließen sich damit Annäherungen an objektive, unverzerrte und epistemisch gesicherte Repräsentationen beschreiben.
- „Epistemic Ideals“ (epistemische Ideale): Mit diesem Bestandteil sind individuelle Kriterien gemeint, mittels derer Personen beurteilen, ob sie den angestrebten Zustand erfolgreich erreicht haben. Dafür benennen Chinn et al. (2014) fünf prinzipiell mögliche Kategorien: a) die interne Struktur der Aussagen, b) die logische Widerspruchsarmut zu weiteren externen Aussagen, c) die ausreichende empirische Unterfütterung, d) die Glaubwürdigkeit der Quelle und e) die Verständlichkeit und Präsentation. Sie fokussieren damit auf Begründungen von Aussagen, was sich als ausgesprochen fruchtbares und dynamisches Feld der Leseverstehensforschung im Bereich der Multiple-Documents-Literacy erwiesen hat (s. Teilkap. [4.3.3.1](#) und darin den Abschnitt

„Teildimension 3: Begründung des Wissens (Prozess- und Produktperspektiven“). Besonders die Glaubwürdigkeit der Quelle(n) von Aussagen ist aus dem Kategorienkatalog aus Sicht des Sourcings hervorzuheben.

- „Reliable Processes“ (zuverlässige Prozesse): Die zuverlässigen Prozesse im AIR-Modell basieren auf den Ideen des Reliabilismus, nach denen zuverlässige Prozesse unter bestimmten Bedingungen zu Produkten wie Wissen oder gerechtfertigten Überzeugungen führen (Chinn & Rinehart, 2016b; Chinn et al., 2014). Das Wissen darüber, welche Bedingungen dafür erfüllt sein müssen, dass Prozesse in die anvisierten Produkte münden – z.B. a) wie Experimente funktionieren, die Kausalzusammenhänge prüfen, b) wodurch Argumente innerhalb eines wissenschaftlichen Disputs eine höhere szientifische Überzeugungskraft haben, c) wie Schlussfolgerungen an Daten oder Artefakten belegt werden müssen etc. –, ist eine Basis dafür, Inhalte sinnvoll überprüfen zu können.

Diese drei Komponenten sind für das Sourcing vor allem dahingehend von Belang, dass die Prozesse der Erkenntnisgewinnung zwar deutlich im Fokus stehen, dass sie aber von der lesenden Person lediglich mit einer Restunsicherheit eingeschätzt werden können und darum ein Vertrauensverhältnis zur Quelle nötig ist (Hendriks et al., 2016). Der Grund dafür ist, dass Leserinnen und Lesern das eigentliche fundierte Wissen fehlt. Aus diesem Mangel heraus bildet das Sourcing eine Möglichkeit, zu regelgeleiteten Urteilen zu gelangen, deren Bemessungsgrundlage und Basis nicht mehr die Inhalte sind, sondern der Ursprung der Aussagen. Anders gesagt: Sourcing hilft dabei, Vertrauen in eine Quelle zu etablieren.

Die hierfür nötigen Fähigkeiten sind, selbst wenn sie die Inhalte von Dokumenten nicht primär adressieren, gleichwohl immer noch kognitiv voraussetzungsreich und erfordern mehrdimensionale Prüfungen in variablen Kontexten (Chinn et al., 2020, 2021; Chinn & Rinehart, 2016a; Kienhues et al., 2020). Das Sourcing, welches im AIR-Modell eher implizit aufscheint, stellt insgesamt nur einen von diversen Prüfungsvorgängen (unter dem Mantel der epistemischen Ideale) dar. Metadaten zu Quellen werden damit nicht von der Inhaltlichkeit ihrer Aussagen isoliert, sondern idealerweise in ihrem Zusammenspiel betrachtet und eingeschätzt. Es ist dieser Nexus verschiedener Evaluationen, der aus der Perspektive der epistemischen Kognitionen das Sourcing in einen Funktionszusammenhang rückt, in welchem die Quellenbeurteilung dazu dient, angemessenes Vertrauen in die Korrektheit und Redlichkeit in die Quelle und ihre Aussagen aufzubauen (Sperber et al., 2010).

Um die Einbettung des Sourcings in einen größeren Zusammenhang zu stellen, dient der Kasten „Illustrierendes Beispiel 11“. Er stellt eine Studie vor, die Sourcing als Bestandteil sogenannter „epistemischer Begründungen“ kontextualisiert. Anhand der Studie wird erkennbar, dass Sourcing analytisch trennbar von anderen Prozessen ist, dass es jedoch nicht als Selbstzweck fungiert, sondern – wie

in diesem konkreten Fall – dabei hilft, Quellen auszuwählen, die für das Bearbeiten von konkreten und komplexen Aufgabenstellungen geeignete Informationen zu enthalten versprechen.

Illustrierendes Beispiel 11

Macht es einen Unterschied, ob Personen ihre Auswahl von Internetdokumenten mit oder ohne epistemische Begründungen vorgenommen haben?

A) Was war das Erkenntnisinteresse?

Die Experimentalstudie mit Studierenden aus den USA (List et al., 2016b) verscrieb sich dem Erkenntnisinteresse, welche und wie viele (nicht-)epistemische Begründungen Personen geben, die zwei verschiedene Aufgaben mit demselben digitalen Dokumentenset lösten.

B) Wer wurde untersucht?

An dem Experiment nahmen 31, im Mittel 22-jährige Studierende teil, die mit überwiegender Mehrheit weiblich (87 Prozent) waren und sich primär aus europäischen (58 Prozent) bzw. asiatischen Amerikanerinnen und Amerikanern (32 Prozent) zusammensetzten. Die Testpersonen studierten diverse Fächer; eine Gemeinsamkeit war, dass sie in Veranstaltungen sozialwissenschaftlicher Hauptfächer rekrutiert wurden.

C) Welche Instrumente kamen zum Einsatz?

Die Studie nutzte ein Mixed-Methods- und Within-Design. Die Studierenden absolvierten zunächst zwei Aufgaben, eine offene (Bedeutung der staatlichen Kinderbetreuung für hohe Fertilitätsraten) und eine geschlossene (allgemeine Fertilitätsrate in den USA). Die Aufträge waren als universitäre Aufgaben gerahmt. Für die Bearbeitung beider Aufgaben hatten die Testpersonen Zugang zu einer Bibliothek von acht digitalen Dokumenten, die in einer suchmaschinenartigen Umgebung zur Verfügung gestellt worden waren. Diese Dokumente bestanden aus drei Websites, zwei Buchkapiteln als PDF, einem Zeitungs- sowie einem Zeitschriftenartikel und einem PDF-Zeitschriftenartikel. Diese Dokumente wurden in der Suchmaschinenumgebung mit Snippets dargeboten, welche Namen, Titel, Dokumententypus und eine kurze Vorschau zum Inhalt beinhalteten. Die Aufgabenprodukte wurden hinsichtlich der geschlossenen Aufgabe dichotom kodiert (korrekt bzw. inkorrekt). Die offene Aufgabe wurde mehrdimensional kodiert: Anzahl der Wörter, Anzahl der Argumente (als Indikator für Ausmaß an Evidenz) und Anzahl von Zitationen (als Indikator für abgestützte Evidenz).

Nach der Bearbeitung der Aufgabe fand ein Leitfadeninterview statt, um retrospektiv die Auswahlkriterien systematisch zu erfragen. Die aufgezeichneten Interviews wurden inhaltsanalytisch kodiert, wobei zunächst grob in epistemische und nicht-epistemische Hauptkategorien differenziert und danach taxonomisch feinteiliger unterteilt wurde. Das Ergebnis waren acht Subkategorien bei den epistemischen Begründungen und elf Subkategorien bei

den nicht-epistemischen Begründungen. Diese Subkategorien wurden aufgrund ungleicher Häufigkeiten noch aggregiert (s. Tabelle 7) und die Anzahl der individuell aufgeführten Begründungen mit den Produktmaßen in Varianz- und Korrelationsanalysen ausgewertet.

D) Was waren die Hauptbefunde?

Der Vergleich der beiden Aufgaben ergab, dass sich die Anzahl von Begründungen nicht unterschied, dass aber die nicht-epistemischen Begründungen deutlich häufiger als die epistemischen vorkamen, das Verhältnis lag bei rund 5:1. Die Anzahl der Begründungen lag insgesamt recht hoch und erstreckte sich auf diverse Subkategorien, die Zahlen im oberen Teil der Tabelle 7 geben Auskunft über die Mittelwerte der Hauptkategorien und der beiden Aufgaben. Hinsichtlich der Kategorie A) quellenbasierte Faktoren ist bei den epistemischen Überzeugungen anzumerken, dass mehr als jede zweite epistemische Begründung auf diese gebündelte Kategorie entfiel. Zusätzlich korrelierte die Zahl von epistemischen und nicht-epistemischen Begründungen hoch: Wer Begründungen der einen Kategorie anführte, tat es demnach ebenfalls deutlich mehr bei der anderen Kategorie. Die Korrelationsanalysen offerierten vertiefte Einsichten. Nur bei der offenen Aufgabe hingen die Zahl der genutzten Quellen bei der Beantwortung und die Anzahl der Begründungen mittelstark positiv zusammen. Bei der geschlossenen Frage war das nicht der Fall; für sie war es im Übrigen auch bezüglich der Korrektheit der Antwort unerheblich, wie viele Begründungen die Personen nannten. Im Falle der offenen Aufgabe ließen sich zudem durchgängig positive Zusammenhänge zwischen den epistemischen Begründungen und sämtlichen Produktmaßen nachweisen. Bei den nicht-epistemischen Begründungen gab es diese Zusammenhänge nicht. Darin manifestiert sich, dass Prüfungen und Begründungen zuvorderst bei komplexeren Aufgaben von Vorteil sind.

E) Diskussion

Die Ergebnisse unterstreichen, dass epistemische und nicht-epistemische Begründungen gleichermaßen angeführt werden, selbst wenn der Anteil der epistemischen Begründungen eklatant geringer ausfällt. In dieser Kategorie überwiegen allerdings die auf Sourcing basierenden Urteile. Bei der komplexeren der beiden Aufgaben, deren Lösung nicht binär korrekt bzw. inkorrekt ist und anhand eines einzelnen Fakts gelöst werden kann, erwiesen sich die epistemischen Begründungen in Korrelationsanalysen als vorteilhaft. Sie waren im Gegensatz zu den nicht-epistemischen Begründungen mit zwei Produktmaßen verquickt: Mehr epistemische Begründungen gingen mit mehr Argumenten zu einem komplexen Sachverhalt und mehr Referenzierungen einher. Das spricht dafür, dass ein stärker auf epistemischen Gründen basierendes Analysieren von digitalen Dokumenten günstig ist, wenn die Aufgabe komplex ist.

Tabelle 7: Gegenüberstellung von retrospektiv erfassten epistemischen und nicht-epistemischen Begründungen für die Auswahl von digitalen Dokumenten für die Bearbeitung zweier Aufgaben

Epistemische Begründungen für Informationsauswahlen	Nicht-epistemische Begründungen für Informationsauswahlen
<i>Definition:</i> Basis sind Begründungen, die eine Sensitivität für die Quellen- und Inhaltsmerkmale hinsichtlich der Charakteristika von Wissen und Wissensproduktion nahelegen	<i>Definition:</i> Begründungen, die sich Oberflächenmerkmale wie Inhalt, Relevanz, Verständlichkeit, Verfügbarkeit beziehen
Kategorien bei den epistemischen Begründungen (aus 8 Subkategorien; im Mittel 4,3 Begründungen pro Person, davon 2,6 bei der offenen und 1,8 bei der geschlossenen Aufgabe)* A) Quellenbasierte Faktoren (55 %**) • <i>Dokumententypus</i> • <i>Glaubwürdigkeit</i> der Quelle (Vertrauenswürdigkeit, Zuverlässigkeit, Reputation) • <i>Expertise</i> der Autorinnen und Autoren B) Inhaltliche Faktoren (27 %**) • <i>Akkuratheit</i> der Informationen beim Inhalt • <i>Qualität oder Objektivität</i> der Informationen bzw. deren Quelle (Korrektheit, Verzerrung oder Glaubhaftigkeit der Quelle) • Epistemische Einschätzung zur <i>Einfachheit des Wissens</i> • Evidenzbasierte Einschätzung zur <i>wissenschaftlichen Glaubwürdigkeit</i> anhand von Quellenangaben oder der Wissenschaftlichkeit des Dokuments C) Corroboration (11 %**; gebildet aus nur einer Subkategorie): verifikatorischer Abgleich mit anderen Inhalten aus den anderen Dokumenten bzw. dem eigenen Vorwissen D) Individuelle Einschätzung der Quelle (8 %**; gebildet aus nur einer Subkategorie): Einschätzung der Quelle aufgrund des eigenen Wissens oder eigener Überzeugungen	Kategorien bei den nicht-epistemischen Begründungen (aus 11 Subkategorien; im Mittel 21,7 Begründungen pro Person, davon 9,4 bei der offenen und 12,3 bei der geschlossenen Aufgabe) 1) Informationsgetrieben (37 %**) • <i>Inhalt</i>, den die Person gezielt ansteuert, wie Definitionen oder Beispiele • <i>Antwortgetrieben</i> hinsichtlich des als ausreichend eingeschätzten informatorischen Gehalts für die Antwort auf die Frage • <i>Informationsbedürfnis</i> wie Neugierde hinsichtlich einzelner Informationen 2) Oberflächenmerkmale (27 %**) • <i>Oberflächenmerkmale</i> des Dokuments wie Titel, Beschreibung, Schlüsselwörter • <i>Reihenfolge</i> der Ergebnisse auf der Suchmaschinenseite 3) Einfachheit der Nutzung (10 %**) • <i>Zugänglichkeit</i> des Dokuments, darunter Eigenschaften wie Länge, Lesbarkeit, klare Strukturiertheit, Formatierungen • <i>Habitualisiertes Wissen</i> über Websites und ihre Eignung 4) Zufall (9 %**; gebildet aus nur einer Subkategorie): <i>Zufälligkeit der Auswahl</i> ohne klare Benennung von Kriterien 5) Relevanz (7 %**; gebildet aus nur einer Subkategorie): für die Ausgabenbearbeitung eingeschätzte Nützlichkeit 6) Verknüpfung mit anderen Informationen außerhalb des Dokuments (7 %**) • <i>Vorwissensbasierte Einschätzung</i> des Dokuments • (Oberflächlicher) <i>intertextueller Bezug</i> zwischen den einzelnen Dokumenten

(Definitionen adaptiert von List et al., 2016b, S. 24; Quelle: eigene Darstellung anhand der Daten aus List et al., 2016b, S. 35 f., 38–40 und 54–56; * Abweichungen bei der Gesamtzahl ergeben sich aus Rundungsfehlern; ** Abweichungen bei dem Gesamt von 100 Prozent ergeben sich aus Rundungsfehlern)

3.1.3.2 Expertise- und Vertrauenswürdigkeitsurteile als zentrale Dimensionen der evaluativen Komponente des Sourcing für Glaubwürdigkeitsurteile

Sourcing kann sich wegen des Bezugspunkts der Metadaten auf nahezu jede Informationseinheit über die eigentlichen bedeutungstragenden propositionalen inhaltlichen Elemente aus Dokumenten und über Dokumente beziehen. Das bedeutet: Potenziell kann jedes Metadatum zu einem bedeutungsrelevanten Element bei der Einschätzung der Glaubwürdigkeit von Quellen avancieren – und tut dies aus empirischer Sicht auch (Sun et al., 2019). Entscheidend ist dabei, dass diese Metadaten situations- und domänenspezifisch adäquat ausgewählt und interpretiert sowie mit dem Inhalt des Dokuments bzw. der Aussage kombiniert werden (Goldman et al., 2016).

Glaubwürdigkeitsurteile werden in der Forschung zum Sourcing methodisch unterschiedlich erfasst und verschiedentlich operationalisiert, wobei aus der Produktwarte Fragebögen mit Urteilen zur Rangreihung von Dokumenten entlang ihrer Vertrauenswürdigkeit einen dominanten Zugang darstellen (s. für einen Überblick Anmarkrud et al., 2022). Damit Personen solche Urteile abgeben können, müssen sie in der Regel verschiedene Metadaten erkennen und kombinieren sowie einschätzen. Als hochaggregierte Zuweisungen werden Glaubwürdigkeitsurteile demnach idealtypisch auf mehrere Dimensionen (und diese wiederum auf verschiedene zusammengehörige Metadaten) abgestützt, was sich beispielsweise die Förderung des Sourcing gezielt zunutze macht (Brante & Strømsø, 2018; Philipp, 2020b). Als zentrale Dimensionen der Zuweisung von Glaubwürdigkeit gelten primär aus der Sicht der Persuasionsforschung (Pornpitakpan, 2004), der Experten-Laien-Kommunikationsforschung (Bromme et al., 2018), der Informationsforschung (Choi & Stvilia, 2015), dem Verständnis von epistemischer Vertrauenswürdigkeit (Hendriks et al., 2016) und der Forschung zur „Scientific Literacy“ (Britt et al., 2014) zwei Dimensionen: *Expertise* und *Vertrauenswürdigkeit*. Dies spiegelte bereits die Studie von List et al. (2016b) aus dem Kasten „Illustrierendes Beispiel 11“ aus dem Teilkapitel zuvor wider: Beide Dimensionen ließen sich separat erfassen und wurden unter der Bezeichnung „quellenbasierte Faktoren“ gebündelt. Forschungsüberblicke berichten ebenfalls von der Trennbarkeit und Bedeutsamkeit von diesen Teildimensionen, zum Beispiel bei der Suche nach Gesundheitsinformationen (s. z. B. Sbaffi & Rowley, 2017, sowie Sun et al., 2019). Was ist jedoch mit Expertise und Vertrauenswürdigkeit gemeint?

Dimension 1: Expertise

Expertise bezieht sich auf das „Ausmaß, in dem ein Sprecher/eine Sprecherin als fähig angesehen wird, korrekte Aussagen zu tätigen“ (Pornpitakpan, 2004, S. 244). Demnach geht es vor allem darum, innerhalb der Domäne bzw. des Themas

die entsprechende Fähigkeit bzw. das nötige Fachwissen aufzuweisen (Bromme et al., 2018). Konkret ist damit also gemeint, dass Personen, die sich über eine medizinische Frage informieren wollen, in ihren Expertiseurteilen den Personen mit erkennbarer Ausbildung (Studium der Medizin, ggf. auch Dokortitel) und mit Erfahrung in ihrem Wissensgebiet (Berufspraxis, Forschungserfahrungen, Fachpublikationen, Renommee als Wissenschaftlerin/als Wissenschaftler etwa durch Preise oder Auszeichnungen, aber auch die fachliche Vertiefung innerhalb der Wissensgebiete) eine höhere Expertise zuschreiben. Zudem sind die Distributionsquellen anzuführen, etwa ob die Aussagen in einem Fachverlag und dessen Periodika oder Buchsortiment erscheinen, ob die Beiträge von anderen Wissenschaftlerinnen und Wissenschaftlern vor der Veröffentlichung geprüft worden sind (Peer-Review) etc. Auch hierdurch sind Rückschlüsse auf die inhaltliche Qualität der wissenschaftlich geprüften Aussagen möglich, die dabei helfen, die Expertise einzuschätzen.

Unterfüttert wird die Bedeutsamkeit der Expertise nebst älteren metaanalytischen Befunden (Wilson & Sherrell, 1993) von empirischen Erkenntnissen, aus denen zum Zwecke der Illustration hier die Überblicksarbeit von Sun et al. (2019) stellvertretend steht. In diesem Review wurden 73 Studien mit offenen Aufgabenformaten hinsichtlich der Einschätzung von Gesundheitsinformationen aus dem Internet sekundäranalysiert. Aus diesen Studien wurden 25 Beurteilungskriterien sowie 165 Indikatoren der Beurteilungen extrahiert und geprüft, ob diese Indikatoren zu einer positiven, negativen oder ambivalenten Evaluation führten. Aus den Quelleneinschätzungen schälten sich beim Kriterium der Expertise folgende in der Darstellung bereits wiederum stark aggregierten Indikatoren als maßgeblich heraus: Seitens der Person(en) als Quelle wurden die Profession und Professionalität (sich teils im Sprachgebrauch und dem Befolgen wissenschaftlicher Konventionen manifestierend) sowie die aktive Einbindung von anderen wissenschaftlichen Quellen angeführt. Seitens der Medien, in denen die Dokumente erschienen sind, gaben die Personen Website-Domänen, die auf Bildungs- oder Regierungsorganisationen hinweisen, sowie die Organisationen als Grundlage an. Anders gesagt: Positive Expertiseurteile stützen sich auf mehrere Metadaten ab.

Dimension 2: Vertrauenswürdigkeit

Die *Vertrauenswürdigkeit* ist das komplexere Konstrukt. Pornpitakpan (2004) bezeichnet damit den Grad der zugestandenen Wahrhaftigkeit einer sich äussernden Person gegenüber der rezipierenden Person. In diesem Sinne ist Vertrauenswürdigkeit eine erfahrungs- und wahrnehmungsbasierte Zuschreibung zwischen zwei Akteurinnen bzw. Akteuren, die sich auf eine relationale Größe, das Vertrauen, in einem spezifischen Kontext bezieht (Blöbaum, 2016). Gewendet auf das Lesen geht es also darum, dass die lesende Person der (eingebetteten)

Quelle redliche Absichten unterstellt, was Inhalte und Zwecke der Kommunikation angeht – dieses Vertrauensverhältnis besteht solange, bis Zweifel an der Quelle und ihren Aussagen aufkommen (Sperber et al., 2010). Genau dies ist der Mechanismus, den sich geschickt getarnte Misinformationen bzw. Fake News zunutze machen, indem sie Vertrauenswürdigkeit vorschützen (Egelhofer & Lecheler, 2019; Kendeou et al., 2019).

Blöbaum (2016) hat fünf Merkmale aus einer Forschungssynopse extrahiert, anhand derer sich Vertrauenswürdigkeit gegenüber Quellen bestimmen lässt, und diese Merkmale lasen sich mit den extrahierten Kriterien von Sun et al. (2019) in einigen Fällen vereinen:

- 1) *Kompetenz* – verstanden als die Fähigkeit, die nötigen Aufgaben innerhalb einer Situation ausführen zu können. Dies ist mit der Expertise als eigener Teildimension oben bereits behandelt worden. In der dimensional Beschreibung von Blöbaum (2016) ist Expertise damit eine Facette der Vertrauenswürdigkeit, in anderen Zugängen, die zu Beginn dieses Teilkapitels aufgeführt wurden, hingegen nicht.
- 2) *Integrität* – oder die konsistenten und belastbaren Regeln folgende Herstellung von Wissensaussagen. In dem schon erwähnten empirischen Überblicksbeitrag von Sun et al. (2019) wurden Facetten der Integrität über die Kriterien Akkuratheit (Übereinstimmung mit wissenschaftlich konsensuellen Ergebnissen) und Objektivität (von persönlichen oder kommerziellen Interessen unbeeinflusste Aussagen) abgebildet.
- 3) *Benevolenz* – die unterstellte gute bzw. nicht-schadende Absicht, also die Zuschreibung, moralisch redlich zu agieren und gute Motive zu verfolgen – oder zumindest doch die gleichen Absichten zu hegen wie die Person in der Rezipient:innenrolle. Diese Benevolenz wurde in der Arbeit von Sun et al. (2019) ebenfalls behandelt, nämlich dem Kriterium der als Glaubwürdigkeit bezeichneten Aufrichtigkeit. Diese spiegelte sich zum Beispiel in der Einschätzung von Domänen von Websites und der Organisation hinter Websites wider.
- 4) *Symbole* – das Vorhandensein von Merkmalen wie akademischen Titeln, Empfehlungen, Rankings, weiteren Insignien, die auf eine besondere Macht oder Befugnis verweisen. In der Überblicksarbeit von Sun et al. (2019) sind Aspekte der Symbole gleichfalls auffindbar, etwa indem die Empfehlungen von Dritten (Zertifikate oder Qualitätssiegel) dazu genutzt werden, die Akkuratheit positiv einzuschätzen. Ähnliches gilt für Empfehlungen sowie Verlinkungen von als glaubwürdig eingeschätzten Websites. Ebenso sind eingebettete Quellen sowie Verlinkungen aus den Dokumenten zu anderen vertrauenswürdigen analogen Quellen und Dokumenten Merkmale, die als günstige Indikatoren für Zuschreibungen von Glaubwürdigkeit genutzt werden.

- 5) *Reputation* – die mehrdimensionale Zuschreibung positiver Merkmale, die sich aus der Gemengelage anderer Merkmale wie der Kompetenz speisen.

Vertrauenswürdigkeit ist damit gemäß Blöbaum (2016) ein Konstrukt, dessen jeweils konkrete Verfasstheit sich über mehrere Merkmale bestimmen lässt. Das bedeutet: Vertrauenswürdigkeit bemisst sich daran, dass mittels Metadaten die Ausprägung verschiedener Merkmale bestimmbar wirkt und dadurch ein mehr oder minder vollständiges Gesamturteil für die Vertrauenswürdigkeit zustande kommt.

Ein ausführliches empirisches Beispiel dazu, worin beim erfolgreichen Sourcing die Unterschiede liegen

Wer mittels Metadaten die Expertise und Vertrauenswürdigkeit als wichtige Säulen von Glaubwürdigkeit einschätzen soll, tut dies in aller Regel nicht als Selbstzweck, sondern dazu, indirekte Plausibilitätsurteile zu fällen (Kienhues et al., 2020). In hochkomplexen (medial geprägten) Wissensgesellschaften ist dies eine Konzession daran, dass das Verstehen uneindeutiger, epistemisch unsicherer bzw. mehrperspektivischer Sachverhalte an die Expertise von Spezialistinnen und Spezialisten gebunden ist (Bromme & Goldman, 2014). Insbesondere für Laien mit einem geringen Wissen kommt darum dem Sourcing die Funktion zu, über die metadatenbasierten Glaubwürdigkeitsurteile eine Annäherung an die inhaltlich grundierten Plausibilitätsurteile vorzunehmen (Sharon & Baram-Tsabari, 2020; Sinatra & Lombardi, 2020). Sourcing ist damit *ein* Bestandteil unter mehreren evaluativen Prozessen. Dies wird besonders deutlich an der Studie im Kasten „Illustrierendes Beispiel 12“. In ihr wurden auf verschiedene Weisen die Antworten von zwei Personengruppen kodiert; den Personengruppen glückte bzw. glückte es nicht, die Vertrauenswürdigkeit einer Website zu begründen. Die leistungsstärkeren Personen nutzten sowohl inhaltliche als auch Metadaten, doch die größten Differenzen ergaben sich darin, dass sie beim Sourcing präziser die Metadaten verwendeten und aus Inhalt, Metadaten und einer prüfenden Grundhaltung ihre Antworten speisten. Diese Fähigkeit setzt die Verarbeitung verschiedener interdependenter Informationen voraus, was sich im zweiten Hauptergebnis, einer Taxonomie von Leistungen anhand offener Begründungen, ebenfalls niederschlägt.

Illustrierendes Beispiel 12

Wider die Dichotomie – oder: Worin genau unterscheiden sich Jugendliche, die erfolgreiche Glaubwürdigkeitsurteile fällen, von solchen, die es nicht tun?

A) Was war das Erkenntnisinteresse?

Sind Jugendliche, die Sourcingurteile erfolgreich begründen bzw. nicht begründen, eine homogene Gruppe? Dieser Frage ging eine Mixed-Methods-Studie mit offenen Aufgaben nach, in der gezielt Muster bei beiden Gruppen eruiert wurden (Forzani et al., 2022).

B) Wer wurde untersucht?

Das Sample bestand aus 1.734 Jugendlichen der siebten Klassenstufe, die Stichprobe wurde gezielt hinsichtlich mehrerer Merkmale in zwei US-Bundesstaaten gezogen. Der sozioökonomische Status wurde über die vergünstigte Schulspeisung ausgewiesen, auf die rund 40 Prozent der Schülerschaft aller teilnehmenden Schulen ein Anrecht hatten. Von den mehr als 1.700 Jugendlichen beantworteten 205 Jugendliche eine anspruchsvolle Frage korrekt und wurden deshalb ausgewählt. Von den restlichen Jugendlichen wurde eine gleiche Anzahl randomisiert ausgewählt, welche dieselbe Frage nicht korrekt beantwortet hat. 410 Jugendliche bildeten damit die Stichprobe für eine vertiefte Analyse ihrer korrekten bzw. inkorrekten Antworten.

C) Welche Instrumente kamen zum Einsatz?

Die Jugendlichen absolvierten einen szenariobasierten Lesetest, in dem sie zu konkreten Fragen und Themen innerhalb einer simulierten Internetumgebung Informationen finden, einschätzen, synthetisieren und kommunizieren mussten. Für die Zuweisung der beiden Gruppen war eine Frage entscheidend, bei der es die Glaubwürdigkeit einer spezifischen Website einzuschätzen und vor allem zu begründen galt. Die Website hatte erkennbare und interpretierbare Metadaten. Die Gruppe der besseren Jugendlichen war zur geforderten Evaluation erfolgreich in der Lage, die der schwächeren war es nicht oder nur rudimentär. Die Antworten wurden mehrfach kodiert. Ein erster Zugang bestand darin, dass der Referenzpunkt a) der Inhalt der Website, b) die Metadaten zur Quelle wie Namen und Organisationen oder c) kontextbezogene Informationen wie Zeit des Erscheinens, Dokumentenart oder URLs waren. Mit den beiden letztgenannten Kategorien wurden damit verschiedene Metadatenarten fokussiert. Der zweite Zugang bestand aus einer tiefergehenden Kodierung der Begründungen hinsichtlich der Art der Argumentation zum einen und der kritischen Haltung gegenüber den verschiedenen potenziell nutzbaren Informationen und Metadaten zum anderen. Die finale Kodierung betraf eine Zuteilung zu Gruppen von Personen entlang eines Kontinuums gemäß den Mustern in den Antworten, wie begründet wurde.

D) Was waren die Hauptbefunde?

Die Gruppe der besseren Jugendlichen fiel im Vergleich zu ihren weniger leistungsstarken Peers dadurch auf, dass sie

- deutlich häufiger Metadaten wie die Ausbildung und berufliche Expertise anführten (Referenzpunkte des Urteils);
- in ihrer Begründung präzisere Details als Belege angaben und diese deutlicher mit ihrem Urteil als Abstützung verbanden (Argumentation in der Begründung);
- die Wahrhaftigkeit von Websites prüften und zweiseitige Betrachtungsweisen in ihren Antworten demonstrierten (kritische Haltung).

Die Tiefenkodierung ergab insgesamt sechs Gruppen, welche sich aber ungleich verteilten. Die leistungsstärkeren Schülerinnen und Schüler (mit den korrekten Lösungen) bestand ungefähr im Verhältnis 1:2 aus zwei Teilgruppen (s. Abbildung 6). Die Gruppe der leistungsschwächeren Jugendlichen verteilte sich hingegen auf mehr Ausprägungen (Abbildung 7), nämlich insgesamt vier Gruppen, von denen die größte (Gruppe B) noch unterteilbar war. Leistungsstärkere Jugendliche konnten ihr Urteil themenspezifisch ausdrücken und mit dem Inhalt bzw. der Vertrauenswürdigkeit verbinden (Gruppe E in Abbildung 6). Ein knappes Drittel war zusätzlich dazu in der Lage, über die Belege wie Inhalt bzw. Glaubwürdigkeit das themenspezifische Urteil präzise zu begründen. Diese Gruppe F bildet den normativ positiven Endpunkt des Kontinuums.

Bemerkenswert ist, dass die leistungsschwächeren Jugendlichen nicht nur ebenfalls heterogen waren, sondern dies sogar noch stärker als die leistungsstärkeren Jugendlichen. Ein großer Teil dieser Jugendlichen schätzte die Glaubwürdigkeit gar nicht ein. Sie beurteilten die Relevanz bzw. wiederholten den Inhalt (Gruppe A), benannten nur sehr allgemeine Kontextmerkmale wie das Datum oder die Textsorte (Gruppe C) oder führten ohne thematische Eindeutigkeit inhaltliche oder metadatenbezogene Aussagen an (Gruppe D). Selbst die Personen, die Kürzestbegründungen abgaben (Gruppe B), ließen es an mindestens einem Bestandteil Aussage bzw. Beleg fehlen.

Schülerinnen und Schüler mit korrekter Lösung

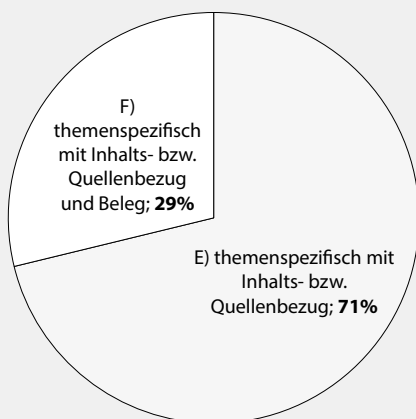


Abbildung 6: Verteilung der erfolgreich begründenden Jugendlichen auf der Basis ihrer Antworttypen (Quelle: eigene Darstellung, basierend auf Forzani et al., 2022, S. 11)

Schülerinnen und Schüler mit inkorrektter Lösung

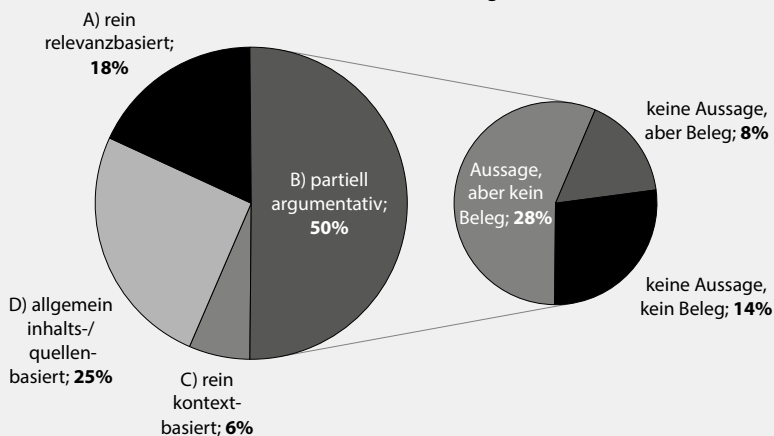


Abbildung 7: Verteilung der nicht-erfolgreich begründenden Jugendlichen auf der Basis ihrer Antworttypen (Quelle: eigene Darstellung, basierend auf Forzani et al., 2022, S. 11)

E) Diskussion

Das wichtigste Ergebnis der Studie ist, dass hinter der binären Zuteilung von erfolgreicher bzw. nicht-erfolgreicher Evaluation der Glaubwürdigkeit einer Webseite keine homogenen Gruppen stecken. Vielmehr erbringen die Jugendlichen unterschiedliche Leistungen. Gleichwohl gelingt es den Jugendlichen mit korrekter Lösung, komplexer zu argumentieren und hierbei für erfolgreich vernetzende Urteile verschiedene Informationen zu verwenden. Ihre Antworten lassen auf eine adäquate Metadatennutzung schließen.

3.1.3.3 Vom Sourcing zum lateralen Online-Lesen: zwischen ökologischer Validität und experimenteller Kontrollierbarkeit

Das Lesen von digitalen Dokumenten aus dem Internet inkludiert stark die evaluative Komponente (Cho & Afflerbach, 2017; Magliano et al., 2017). Diese Facette des digitalen Lesens valide zu erfassen, stellt in der Forschung eine Herausforderung dar, welche Wineburg et al. (2022, S. 895) dezidiert als „Faustianischen Handel“ bezeichnen. Sie kritisieren, dass Studien zum Evaluieren zwecks besserer experimenteller Kontrollierbarkeit statische Dokumentensets aus dem Internet nutzen bzw. zu nutzen vorgeben, welche aus Sicht der ökologischen Validität einen entscheidenden und substanziellen Nachweis aufweisen. Sie spiegeln nicht die nahezu unlimitierten, aber methodisch aufwändiger auszuwertenden Suchmöglichkeiten und Zugänge wider bzw. erzeugen aus der Sicht von Wineburg et al. (2022) bezüglich Sourcing artifizielle Ergebnisse, deren Aussagekraft bzgl. realer Anforderungen fragwürdig wirkt. Das Argumentarium von Wineburg et al. bezieht sich zwar auf die Legitimation ihrer Intervention zum „lateralen Lesen“ (McGrew, 2024). Doch es betrifft im Kern die Merkmale dessen, was ein kompetentes Nutzen des Internets ausmacht, nämlich die Verwendung von Informationen, um informierte Entscheidungen zu treffen, wobei die Grundlage der Entscheidungen nicht aus im Vorfeld festgelegten Sets an (kuratierten) Dokumenten besteht (Breakstone et al., 2024).

Die Perspektive von Wineburg et al. (2022) und damit auch ihre Kritik an der bestehenden Forschungspraxis mit ihrem in der Tragweite zu wenig reflektierten Implikationen ist in Tabelle 8 gegenübergestellt. Die paradigmatischen Sichtweisen, die in der Tabelle kontrastiert werden, haben damit zu tun, dass im Fall des traditionellen Umgangs mit vorgegebenen Sets an Dokumenten die gegebenen Metadaten allein als Grundlage für das Sourcing benutzt werden. Demgegenüber ist die offene Internetrecherche davon gekennzeichnet, dass andere Dokumente mehrdimensional für die Evaluation verwendet werden, wobei die Metadaten einen Ausschnitt daraus bilden und die Leserinnen und Lesern aktiv nach Informationen suchen, die bei der Einschätzung des Sachverhalts helfen.

Tabelle 8: Vergleich von Evaluationsprozessen und -merkmalen im Forschungsparadigma der traditionellen Nutzung multipler Dokumente zum einen und der offenen Internetrecherche zum anderen

Vergleichsdimensionen	Traditionelles Paradigma multipler Dokumente	Paradigma der offenen Internet-suche
Ziel	disziplinäres Wissen mit einem kuratierten Dokumentenset aufbauen	glaubwürdige Informationen in einer offenen Internetrecherche finden, um informierte Entscheidungen zu treffen
Annahmen über die Quellen	Einschätzung alleinig anhand der gegebenen Informationen	keine alleinige Einschätzung anhand der gegebenen Informationen
	Einschätzung der Quelle als Grundlage für Glaubwürdigkeit nutzen	andere Quellen als Grundlagen für Glaubwürdigkeit nutzen
	vorgegebene, fixierte Anzahl an Quellen	unlimitierte Anzahl an Quellen
Ansätze der Quellenevaluation	Engagement beim Sourcing	Engagement beim lateralen Lesen
	alle Quellen/Dokumente lesen	zweifelhafte Quellen/Dokumente gezielt ignorieren

(Quelle: eigene Darstellung, basierend auf Wineburg et al., 2022, S. 896)

Die damit verbundene Implikation ist von großer Tragweite: Es geht in einer Kritik der Operationalisierung des Sourcings in Bezug auf eine artifiziell anmutende Eingrenzung um mehr als ein geringfügiges methodisches Monitum. Die Perspektive auf die offene Internetrecherche gesteht dem Sourcing eine wesentlich umfassendere Rolle zu, indem die Prüfung und Einschätzung der Metadaten (in Verbindung mit inhaltlichen Aussagen in der Kombination) bedeutend komplexer und umfassender modelliert werden (Forzani, 2020). Dabei besteht eine Herausforderung für die lesende Person in der großen Menge an Internetdokumenten, deren Kompilation und Nutzung kontingent ist: Auf welche Internetdokumente im Rahmen ihrer Recherche sie stoßen und wie sie diese nutzen werden, ist für die lesenden Personen nicht vorhersehbar. Darum bemühen Wineburg et al. (2022) explizit Heuristiken als Grundlage für das Erklären eines erfolgreichen Expert:innen-Vorgehens in offenen Problemlösesituationen.

Mit *Heuristiken* bezeichnen Gigerenzer und Gaissmaier (2011, S. 454) „eine Strategie, welche Teile von Informationen ignoriert, um das Ziel einer Entscheidungsfindung schneller, sparsamer und/oder akkurater zu erreichen, als es mit komplexeren Methoden möglich ist.“ Dieser Zugang, Entscheidungen zu treffen, wie sie beispielsweise bei jedem Suchmaschinen-Ergebnis nötig sind, betont die Angemessenheit von Simplifikationen, um in komplexen Kontexten die eigene Handlungsfähigkeit bei ausreichender Genauigkeit in puncto Zielerreichung zu gewährleisten. Entsprechend geht es gemäß Shah und Oppenheimer (2008) darum, den Aufwand bei Entscheidungen zu reduzieren, was auch den biologisch limitierten kognitiven Kapazitäten des Menschen geschuldet ist, vor allem denen des Arbeitsgedächtnisses. Für diese Reduktion an Aufwand stehen mehrere einzelne, allerdings auch kombinierbare Vorgehensweisen zur Verfügung:

- 1) *Prüfung von weniger Hinweisen und Informationen für Entscheidungen*, also eine Priorisierung von wahrgenommenen und relevant wirkenden Informationen, z. B. einer Entscheidung für oder gegen das Auswählen von Informationen, indem beispielsweise Hinweise auf Werbung in Suchmaschinen dazu führen, diese Suchergebnisse zu ignorieren;
- 2) *Verringerung der Schwierigkeit, die mit dem Abrufen und Speichern von Informationen verbunden ist*, indem man Vereinfachungen vornimmt, etwa weil man beim Lesen über zwei Alternativen ganz gezielt nur eine Information pro Vergleichsobjekt berücksichtigt oder sich auf Informationen beschränkt, zu denen man selbst Vorwissen hat;
- 3) *Vereinfachung der Gewichtsregeln für Hinweise*, was meint, dass jemand die Belastbarkeit von Aussagen danach einschätzt, indem er oder sie vertraute Informationen als höher belastbar hält oder die eingeschätzte Reputation eine Quelle dafür nutzt, die Richtigkeit einer inhaltlich neuen Aussage oder bei Widersprüchen zu bescheiden;
- 4) *Integration von weniger Informationen*, also eine gezielte Nutzung einer geringeren Anzahl von Information, als sie potenziell zur Verfügung stehen, also beispielsweise eine Beschränkung auf wenige Metadaten oder die Entscheidung, bei Recherchen über Quellen ab einem Sättigungspunkt die Suche zu beenden;
- 5) *Prüfung von weniger Alternativen*, also eine gezielte Reduktion von potenziellen Vergleichen, etwa indem man bei Suchmaschinen eine Klickbeschränkung dazu nutzt, um Suchergebnisse anzuwählen und zu lesen.

Ihren Wert haben Vorgehensweisen wie die oben beschriebenen, die ihrerseits noch feiner untergliedert und differenziert werden können (Gigerenzer & Brighton, 2009; Gigerenzer & Gaissmaier, 2011; Shah & Oppenheimer, 2008), vor allem in Kontexten wie dem Internet, in welchem Individuen potenziell Myriaden an Informationen vorfinden, die sie aber weder alle verarbeiten können noch müssen noch sollten. Stattdessen erleichtern derartige Heuristiken, und das gilt zumal Expertinnen und Expertinnen in ihren jeweiligen Domänen, Entscheidungen aufgrund der regelgeleiteten Vorgehensweise in dazu passenden Kontexten. Diesen Vorteil verdeutlichen Wineburg et al. (2022) bezogen auf offene Internet-Recherchen am Vorgehen von Faktencheckerinnen und -checkern, das sie vergleichend mit anderen (akademischen) Personengruppen untersucht haben (Wineburg & McGrew, 2019). Faktencheckerinnen und -checker nutzen drei Heuristiken, wenn sie mit Aufgaben konfrontiert sind, die Glaubwürdigkeit von Websites zu evaluieren:

- *Erwartungshaltung klären (Footing)*: Diese Heuristik betrifft eine grundsätzliche Erwartungshaltung gegenüber dem Internet als Ort mit eigenen (ökonomischen) Regeln und Funktionsweisen, etwa dem Suchmaschinenmarketing,

der Funktionsweise von Suchmaschinen, der Verwendung von Suchbegriffen, aber auch dem Nebeneinander von Websites divergierender Qualität und Absichten. Diese Heuristik, welche eine Klärung der Aufgabenumgebung betrifft, dient dazu, die Offenheit von Rechercheaufträgen aufgrund der Gegebenheiten des Internets zu repräsentieren.

- *Positionsbestimmungen (Taking Bearing)*: Aufgrund von Suchmaschinenergebnissen bemühen sich die Faktencheckerinnen und -checker darum, ihren Start- und Zielpunkt zu erkennen bzw. anzupeilen. Dies tun sie, indem sie die Ergebnisse von Suchmaschinen heuristisch einschätzen, sie nutzen die Snippets dazu, vergleichend den potenziellen Ertrag von fokalen Websites zu bestimmen. Sie beschränken sich in der Zahl ihrer Klicks auf die gefundenen Websites und bestimmen, welche der Suchergebnisse eine geringere und eine höhere Qualität für ihr Anliegen haben. Dadurch erhöhen sie ihre Chance, sinnvolle Informationen als solche zu erkennen – in diesem Fall: von potenziell aufzusuchenden Websites.
- *Laterales Lesen (Lateral Reading)*: Dieses seitwärts stattfindende Lesen erfolgt über mehrere parallel geöffnete Browser-Tabs. Faktencheckerinnen und -checker prüfen, wer hinter der jeweiligen Website steht, und recherchieren über die im Impressum verfügbaren Informationen auf anderen Websites (in den parallel geöffneten Tabs) weitere Informationen. Sie lesen also eher *über* die Texte, statt diese (vertikal) vollständig zu lesen. Mit anderen Worten: Sie schätzen mit ebenfalls erst zu recherchierenden Informationen ein, ob sich der Leseaufwand überhaupt lohnt.

Solche Fähigkeiten, wie sie sich prototypisch bei Expertinnen und Experten im Bereich Verifikation und Recherche aus dem Bereich des Journalismus' als Anforderungen des Lesens im Internet darstellen, verdeutlichen, warum Wineburg et al. (2022) eine so deutliche Kritik an der herkömmlichen Operationalisierung des Sourcing im Allgemeinen bzw. der Evaluation von Informationen in Laborsettings im Besonderen formulieren. Denn ohne die faktische Nutzung des Internets für sämtliche der drei Heuristiken wirkt es unwahrscheinlich, adäquate ökologisch valide Beurteilungen vorzunehmen, welche im Kontext des Internets sinnvoll sind. Gleichwohl bringen die für Expertinnen und Experten ihrer Domäne komplexitäts- und aufwandsreduzierenden Heuristiken einen hohen Aufwand des Lesens mit sich, der sich in multiplen Einschätzungen und Recherchevorgängen manifestiert und entsprechend empirisch bei Novizinnen und Novizen nicht bzw. kaum beobachtbar ist (s. Teilkap. [3.4.2](#)).

Um dies noch einmal anhand eines empirischen Ergebnisses abschließend prägnant herauszuarbeiten, enthält der nachstehende Kasten „Illustrierendes Beispiel 13“ die Befunde einer Studie. In ihr ließ sich deutlich demonstrieren, wie das Klären der Haltung (Footing) und der Position (Taking Bearing) als Muster von professionell lesenden Personen aussieht. Insbesondere zeigt sich

darin eine heuristische Vorgehensweise, gezielt selektiv vorzugehen, um Muster zu erkennen und speziell bei offenen und kontroversen Themen per Sourcing Dokumente anzuvisieren, welche der Klärung des uneindeutigen Sachverhalts effektiv und effizient dienen.

Illustrierendes Beispiel 13

Wenn Profis zu einer kontroversen historischen Person recherchieren – Muster im Sourcing und im Kontextualisieren

A) Was war das Erkenntnisinteresse?

Die Studie war Teil eines größeren Forschungszusammenhangs der Stanford History Group, in der verschiedene Personengruppen (Studierende, promovierte Geschichtsfachprofessorinnen und -professoren sowie Faktencheckerinnen und -checker) verschiedene kurze Internetrecherchen und Aufgaben absolvierten (McGrew, 2022). Im Sinne eines Gruppenvergleichs interessierte sich das Forschungsteam, ob sich diese drei Personengruppen in ihrem Vorgehen unterschieden.

B) Wer wurde untersucht?

Das Sample bildeten drei Gruppen. Erstens handelte es sich um 25 Studierende des ersten Studienjahres an der Stanford University (18 bis 19 Jahre, 52 Prozent weiblich), welche eine große Breite an Studienfächern repräsentierten sowie angegeben hatten, mindestens vier Stunden täglich online zu sein. Zweitens partizipierten zehn Professorinnen und Professoren des Fachs Geschichte (40 Prozent weiblich, Durchschnittsalter: 47 Jahre) aus fünf Universitäten. Drittens nahmen zehn Faktencheckerinnen und -checker (40 Prozent weiblich, im Mittel 34 Jahre alt) teil, die aus großen Nachrichten- und politischen Organisationen stammten und deren Jobprofil es mit sich brachte, regelmäßig Informationen zu prüfen, zu verifizieren und zu recherchieren. Die Gruppe der Professorinnen und Professoren sowie der Faktencheckerinnen und -checker wurden wegen der Ähnlichkeit in dem Gesamtverfahren zu einer Gruppe („Profis“) zusammengefasst. Drei Professorinnen bzw. Professoren und eine Person aus dem Pool der Faktencheck-Organisationen wurden wegen unvollständiger Daten aus den Analysen ausgeschlossen.

C) Welche Instrumente kamen zum Einsatz?

Im Zentrum der Studie stand eine offene Aufgabe, bei der die Teilnehmenden eine These zu einer kontrovers eingeschätzten Frauenrechtlerin mittels einer achtminütigen Internetrecherche ohne Vorgaben zum Vorgehen prüfen sollten. Die Prozessdaten, die mittels Lautdenk-Protokollen und Screen-Capturing entstanden, wurden in mehreren Schritten kodiert. Erstens interessierte sich die Forscherin für die Ergebnisse im Sinne des Gesamtverfahrens, wobei danach unterschieden wurde, ob a) zuverlässige und hilfreiche Internetdokumente genutzt wurden und b) ob es zu Verifikationen der Glaubwürdigkeit der Quellen kam. Zweitens wurde analysiert, welche Wege die Personen bei der Lösung nahmen, d. h.

welche Arten von Dokumenten sie wie lange konsultierten. Drittens kodierte die Forscherin, in welchen Fällen die Personen die für das Lesen in der Domäne Geschichte typischen Strategien des Sourcings und der vorwissensbasierten Kontextualisierung anwendeten.

D) Was waren die Hauptbefunde?

Im Vorgehen ähnelten die Professorinnen und Professoren den Faktencheckerinnen und Faktencheckern, zugleich unterschieden sich diese beiden professionell mit Aussagen umgehenden Personengruppen von den Studierenden. Bemerkenswert ist, dass die Verwendung ähnlicher Suchbegriffe in allen Personengruppen zu beobachten war, die Differenz lag in der Nutzung der Ergebnisse der Suchen. Aus den analysierten Daten ließen sich vier Muster extrahieren:

- *Mehrstufiges Prüfen der Zuverlässigkeit bei Profis:* Die Profis – und hier insbesondere die Faktencheckerinnen und -checker – hatten ein Muster im Vorgehen, demzufolge sie gezielt zuverlässige Dokumente aufsuchten und deren Glaubwürdigkeit vertiefend verifizierten. Die Studierenden gaben sich früher auch mit fragwürdigen Dokumenten zufrieden, um die These für sich als geklärt zu wännen.
- *Verstärkte Nutzung der Suchmaschinenergebnisse zur Kontextualisierung:* Profis verwendeten fast dreimal so viel Zeit wie Studierende (letztenannte: zehn Sekunden), ehe sie ein Suchergebnis anklickten. Sie betrachteten mehr Suchergebnisse und nutzten die Snippet-Informationen dazu, sich einen Überblick zu verschaffen. Insbesondere erkannten sie aufgrund ihrer Einschätzungen, dass weitere Suchen nötig sind, und sie führten Suchergebnisse in Gruppen zusammen, um die Divergenz und Konvergenz von Quellen zu erkennen, was ihnen dabei half, die inhaltliche Strittigkeit der These besser einzuschätzen – und welche der Suchergebnisse sie für belastbar hielten. Studierende gaben sich hingegen weniger Zeit bei der Analyse der Suchmaschinenergebnisse und präferierten die ersten Treffer. Statt wie die Profis nutzten sie nicht die Suchmaschine selbst als Quelle von Informationen, sondern eher als Portal für rasche Einstiege in die hoch gerankten Ergebnisse.
- *Intensiveres Sourcing bei den Snippets:* Teil der Nutzung der Suchmaschinenergebnisse war, dass die Profis aufgrund der wenigen Informationen aus den Snippets der Suchmaschine (Titel, URL, kurze Beschreibung, Institution) zu Beurteilungen gelangten, welche Ergebnisse sie für vielversprechend hielten und welche sie ignorierten. Für die Auswahlentscheidungen der Studierenden war neben der Position der Suchmaschinenergebnisse vor allem die vermeintliche inhaltliche Relevanz ausschlaggebend.
- *Divergierende Präferenzen für Primär- und Sekundärquellen:* Bei der Frage nach der Richtigkeit der zu prüfenden These nutzten die Personengruppen unterschiedliche Arten von Dokumenten. Studierende präferierten Primärquellen, d. h. Zeitdokumente aus der Lebenszeit der historischen Persönlichkeit. Profis hingegen konsultierten gezielt Sekundärquellen (also nachträglich entstandene Dokumente über die historische Persönlichkeit) aus ihnen vertrauenswürdig wirkenden disziplinären Kontexten wie digitalisierte Bücher oder Forschungsprojekt-Websites und mieden gezielt die aufwändiger

einzuschätzenden Primärquellen. Diametral anders gingen Studierende vor, welche mit großer Mehrheit Originalaussagen der historischen Persönlichkeit oder Zeitdokumente aufsuchten. Die Profis verließen sich arbeitsökonomisch und -teilig also auf das Urteil von Historikerinnen und Historikern.

E) Diskussion

Diese Studie zeigt – insbesondere durch den Vergleich verschiedener Gruppen – den Effekt der Expertise auf das Vorgehen, wenn sozialwissenschaftlich kontroverse Themen mittels des Internets unter Zeitdruck geklärt werden sollen. Die als Profis bezeichnete Sammelkategorie von zwei verschiedenen Personengruppen mit jeweils eigener ausgewiesener Expertise nutzte die Suchmaschinen anders als Studierende. Profis analysierten länger die Suchergebnisse und verwendeten die Suchmaschinenergebnisse selbst als Orientierungshilfe. Zudem waren sie selektiver, was sie konsultierten, und stellten das Sourcing konsequent in den Dienst, um an vertrauenswürdige Sekundärquellen zu gelangen. Die Studierenden folgten hingegen eher dem Ranking der Suchergebnisse und recherchierten nach Aussagen aus der Originalzeit. Darin manifestiert sich eine markante Differenz: Expertinnen und Experten nutzen bereits die Suchmaschinenergebnisse deutlich analytischer.

3.1.3.4 Zusammenfassung

Sourcing ist ein Verbund verschiedener metadatenbezogener Verarbeitungsprozesse. Diese reichen von dem Erkennen der Metadaten bis hin zu deren Verwendung und insbesondere deren Beurteilung. Metadaten sind Zusatzinformationen, welche auf einer Zusatzebene zu den propositionalen Inhalten von Texten vorliegen und eine bedeutungsergänzende bzw. -modifizierende Funktion erfüllen. Damit kommt den Metadaten eine für die Bedeutungskonstruktion potenziell hohe Bedeutung zu, weil Leserinnen und Leser anhand der Metadaten zu mehrdimensionalen Glaubwürdigkeitsurteilen gelangen können. Diese Heuristik hilft ihnen insbesondere bei epistemisch unsicheren, kontroversen oder auch für die lesenden Personen unvertrauten Informationen dabei, sich über den Umweg einer quellenbezogenen Glaubwürdigkeitseinschätzung der inhaltlichen Plausibilität von Aussagen anzunähern.

Die informatorischen Bezugspunkte des Sourcings bilden Metadaten, die sich entweder auf gesamte Dokumente (Texte und Metadaten als eine Gesamteinheit eines kontextuell gerahmten Diskursteilnehmers) beziehen und darum als „Dokumentenquelle“ bezeichnet werden oder als kleinteiligere Aussagen (wie Zitate oder Paraphrasen) in Sinne einzelner Inhalte in Texten oder Dokumenten vorkommen und darum „eingebettete Quellen“ sind. Diese Unterschiede in der Granularität sind theoretisch herausgearbeitet worden, dennoch ist die grundsätzliche Verarbeitung der Metadaten prinzipiell ähnlich.

Die kognitive Prozessierung der Metadaten erfolgt nicht uniform, da bestehende Theorien und Modelle aus der Leseforschung und weiteren Bezugsdisziplinen die Verwendung und Evaluation der Metadaten innerhalb des Sourcing unterschiedlich darstellen. Bei aller Divergenz der Modellierungen eint sie, dass dem Sourcing eine grundsätzlich unterstützende Leistung attestiert wird, die sich auf das angemessen vollständige Repräsentieren bzw. die fachlich adäquate Nutzung von Dokumenten für den Wissenserwerb erstreckt. Mit dieser Differenzierung wird bereits deutlich, dass sich im konkreten Anwendungs- und Geltungsbereich der Modelle durchaus bedeutende Unterschiede manifestieren. Allerdings lassen sich drei generische Funktionen des Sourcing benennen. Erstens kommt ihm die Rolle zu, bei im- oder expliziten inhaltlichen Widersprüchen und Inkonsistenzen zwischen Dokumenten bzw. innerhalb einzelner Dokumente die Kohärenzbrüche zu überwinden. Das Sourcing der Dokumenten- bzw. eingebetteten Quellen setzt das Erkennen der Metadaten voraus, denn indem der Ursprung von Aussagen auf die Quellen zurückgeführt wird, gelingt es, selbst einander widersprechende Aussagen in einem mentalen Modell zu integrieren. Die zweite Funktion basiert auf evaluativen Anteilen. Leserinnen und Leser nutzen ihre Urteile über Quellen dazu, deren Inhalte zu gewichten. Sie verwenden die Zuweisungen der Glaubwürdigkeit für die Wahl und Abwahl ganzer Dokumente bzw. Teile davon, etwa indem sie aufgrund von Metadaten in Suchmaschinen Ergebnisse ignorieren oder selbst noch im Lektürevorgang in einem teils reziproken Verhältnis von inhaltlicher Einschätzung und Glaubwürdigkeitsbeurteilung Dokumente evaluieren. Das Ziel ist hierbei, glaubwürdigere gegenüber weniger glaubwürdigen Quellen und damit indirekt in ihrer Plausibilität höherwertig erscheinende Inhalte zu privilegieren und sie in der mentalen Modellbildung stärker zu nutzen. Die dritte Funktion basiert ebenfalls auf evaluativen Bestandteilen des Sourcing, welche dabei helfen, die eigenen Überzeugungen und Wissensbestände als automatisch in Anschlag gebrachte Grundlagen für Validierungen von Aussagen zu überwinden. Dadurch gelingt es, ausgewogenere Repräsentationen zu erlangen oder sogar eigene Vorstellungen zu korrigieren und Langzeitgedächtnisinhalte zu modifizieren. Metadaten helfen demnach bei einer Vielzahl von Anlässen dabei, Verstehensprobleme verschiedener Couleur zu lösen oder zumindest doch zu mindern.

Dreh- und Angelpunkt beim Sourcing – zumindest dann, wenn evaluative Prozesse und deren Ergebnisse nötig sind – sind Glaubwürdigkeitsurteile. Die in diesem Kapitel eingenommene Perspektive zielte auf epistemisch fundierte Glaubwürdigkeitsurteile und damit auf die Frage, woran Personen festmachen, dass die Quellen dazu geeignet sind, mutmaßlich als wahr anzunehmende oder zumindest doch als ausreichend belastbar geltende Äußerungen zu tätigen. Derartige Perspektiven sind nicht mehr allein auf das Sourcing beschränkt, sondern nehmen wie das angeführte AIR-Modell umfassende Sichtweisen ein. Das dem Reliabilismus anschlussfähige AIR-Modell bietet hierfür einen Rahmen für das

Sourcing, wobei eine strukturelle Gemeinsamkeit dem Evaluieren gemäß AIR-Modell und dem Sourcing innewohnt: Es handelt sich um mehrdimensionale regelgeleitete Prüfungen, die gegenüber Dokumenten eine epistemisch normative Positionierung legitimieren, indem Ziele, Ideale und Prozesse im Umgang mit Wissen austariert werden. Das Sourcing ist hierbei Bestandteil der epistemischen Ideale, es taucht dort allerdings nur implizit auf. Sourcing als Möglichkeit der Zuweisung der Glaubwürdigkeit ist unter dieser Perspektive ein Akt der multiplen und mehrdimensionalen Prüfung. Dabei gelten zwei Dimensionen als zentral, die übersetzt in Leitfragen als Entscheidungshilfen fungieren: 1) Was qualifiziert eine Quelle, dass sie als ausreichend wissend gelten kann (Expertise)? 2) Welche Merkmale der Quelle verleihen ihr den Status, dass ihren Aussagen jenseits der Inhalte ein Vertrauensvorschuss gebührt (Vertrauenswürdigkeit)? Die zweite Dimension der Vertrauenswürdigkeit erweist sich dabei in sich komplex und beruht auf mehreren potenziell zu treffenden Urteilen, die sich auf diverse Facetten beziehen, darunter die angenommene Integrität und Benevolenz. Die Einschätzung der Vertrauenswürdigkeit nimmt also deutliche Anleihen an sozialpsychologischen Kategorien, die sich nicht mehr allein auf vermeintlich akontextuelles Leseverstehen beziehen, sondern im Gegenteil die Wissenskonstruktion als zutiefst sozialen Verständigungsprozess annehmen.

Diese mehrdimensionalen Einschätzungen des Ursprungs von Aussagen nebst ihrem Kontext gewinnen an Eminenz beim Lesen im Internet. Im lateralen Lesen erweist sich die orientierende Suche nach epistemisch belastbaren Quellen und Dokumenten im Internet als Dreh- und Angelpunkt komplexer Recherchen, um einen zwar heuristischen, aber doch absicherbaren Sättigungsgrad für wissensbezogene Entscheidungen zu erzielen. Expertise- und Vertrauenswürdigkeitsurteile erfolgen innerhalb des gelingenden lateralen Lesens, welches professionelle Faktencheckerinnen und Faktenchecker praktizieren, als stete und adaptive Begleiter von Recherche- und Problemlöseprozessen. In diesem Lesen multipler Dokumente in parallelen Browsertabs manifestiert sich eine konkrete und versierte Variante des Sourcings.

Damit schließt sich der Kreis: Sourcing ist aus der Perspektive des evaluativen Lesens ein heuristischer, die limitierten Fachwissensbestände der Leserinnen und Leser anerkennender, gleichwohl nicht voraussetzungsarmer Prozessverbund. Er dient dazu, Glaubwürdigkeitsurteile zu ermöglichen, welchen eine orientierende und unterstützende Funktion im digitalen Lesen zukommt. Mittels Sourcing nähern sich Leserinnen und Lesern jenen Dokumenten an, die für wissensbezogene und damit epistemische Leseziele geeignet wirken.

3.2 Prozesse und Produkte

In diesem Teilkapitel werden Prozesse und Produkte des Sourcings systematisiert und kompakt dargestellt. Dieser empirische Forschungsüberblick ist nötig angesichts einer Fülle von unterschiedlichen Operationalisierungen in der internationalen Forschung, die zu einer Vielzahl durchaus sehr heterogener Befunde geführt hat. Darum wird als erster Schritt zunächst ein tabellarisch strukturierter Überblick über die verschiedenen Indikatoren aus der Prozess- und Produktwarte gegeben, die dort inhaltlich gebündelt werden, wo die Sachlogik es erlaubt (s. Teilkap. [3.2.1](#)). Die drei anschließenden Teilkapitel widmen sich dann auf dieser Grundlage einzelnen Facetten des Sourcings, nämlich dem Beachten und Benennen von Metadaten ([3.2.2](#)), deren Verwendung ([3.2.3](#)) sowie deren Beurteilung ([3.2.4](#)).

3.2.1 Ein vorgängiger Überblick über Operationalisierungen des Sourcings in Studien

Sourcing ist kein homogenes Konstrukt bzw. kein eindimensionaler Prozess. Dies spiegelt die weite Definition aus Teilkapitel [3.1.1](#) wider. Sie lässt erwarten, dass die in ihr aufgeführten drei Prozesse – Beachten, Beurteilen und Verwenden – in den Prozess- und Produktmaßen sehr unterschiedliche Leistungen im Umgang mit Metadaten abbilden. Was sich in der Theorie andeutet, schlägt sich empirisch nieder: Die Operationalisierung des Sourcings hat zu einer großen Heterogenität in den Studienbefunden geführt (Anmarkrud et al., 2022; Brante & Strømsø, 2018). In ihrem Studienüberblick zu den Zusammenhängen von individuellen Merkmalen und dem Sourcing bündelten beispielsweise Anmarkrud et al. (2022) aus 72 Studien 21 Gruppen von Sourcingmaßen. Dies unterstreicht, dass Sourcing gut untersucht ist – freilich um den Preis der Vergleichbarkeit der Ergebnisse.

Um dieser Heterogenität zu begegnen, bedarf es einer abstrahierenden Sortierung der einzelnen Operationalisierungen in den jeweiligen Studien. Diesen Versuch unternimmt die Darstellung in der Tabelle 9 mit über 100 berücksichtigten Studien. Für diese Tabelle sind zwei darstellerische Hinweise nötig:

- 1) *Vermeidung der Dopplung von Studienergebnissen im Rahmen des Buches.* Die Auswahl der Studien in der Tabelle 9 bzw. deren Behandlung in diesem Kapitel antizipiert die Gefahr der Überschneidung und der Redundanz mit zwei weiteren Bereichen, die in diesem Buch andernorts betrachtet werden. Ein erster Teil der Studien stammt aus einem spezifischen Forschungskontext, in welchem *Sourcingprozesse und -aktivitäten im Kontext der Kohärenzleistungen bei intertextuellen Konflikten* erforscht wurden (s. Abschnitt „Funktion 1: Verstehen selbst widersprüchlicher Aussagen über Sourcingprozesse

auf der Ebene von Intertextmodellen und über den Aufbau von Quelle-Inhalt-Verbindungen“ im Teilkap. [3.1.2.3](#)). Die Studienergebnisse werden darum inhaltlich an einem anderen Ort fokussiert behandelt (s. Teilkap. [4.2.2.3](#)). In wiederum anderen Studien werden *Sourcingleistungen als Bestandteil des Schreibens* behandelt, was neuere Forschungsströmungen reflektiert (Bråten et al., in Druck). Dies betrifft vor allem Referenzierungen von Dokumenten als Produktmaß in offenen Aufgaben, nämlich den Diskurssynthesen (s. Teilkap. [5.2.1.2](#)), und Sourcingprozesse als Bestandteil umfassender hybrider Lese- und Schreibprozesse (s. Teilkap. [5.2.2](#)). Diese Studien und ihre Ergebnisse als zweiter Teil der Studien sind mithin Gegenstand des Kapitels 5.

- 2) *Umgang mit Einzel- vs. Gesamtindizes*. Die Tabelle 9 gibt zwar viele, aber bei weitem nicht alle Operationalisierungen des Sourcings wieder. In weiteren Studien wurde das Produkt des Sourcing mittels größerer Testbatterien und Gesamtindizes abgebildet, welches mehrere verschiedene Einzelleistungen umfasst, z. B. das Erkennen und das Einschätzen von Metadaten (Breakstone, Smith, Wineburg et al., 2021; Britt & Aglinskas, 2002, Experiment 1; Dyoniziak et al., 2023; Forzani, 2018; Hahnel et al., 2016, 2018; Kiili et al., 2021; McGrew et al., 2018; Strømsø et al., 2020). Da solche Gesamtindizes nicht in die darstellerische Form der Tabelle passen, bleiben sie gezielt außen vor. Die Ergebnisse der Studien werden aber dort berichtet, wo sie inhaltlich von Belang sind, nämlich bei den Determinanten und individuellen Unterschieden in Teilkapitel [3.3](#) und bei den Problemlagen in Teilkapitel [3.4](#). Die Frage nach der Aggregation von einzelnen Kategorien mit unterschiedlichen kognitiven Leistungen betrifft nicht nur Tests, sondern sie tritt ebenfalls andernorts, nämlich bei induktiv gewonnenen Verbaldaten, die feinteilig kodiert wurden und dann zu Zwecken der jeweiligen statistischen Analysen aggregiert (in der Regel heißt dies: summiert) worden sind. Davon betroffen sind vor allem Studien mit Lautdenk-Protokollen, die je nach Granularität der Auswertung in ihrer Differenziertheit teils sehr viele, aber eben doch in ihrer Auftretenshäufigkeit eher weniger Fälle der Ausprägung erfassen. In der Tabelle 9 sind solche Studien dort verortet, wo die Auftretenshäufigkeit der Kategorien am größten war.

Tabelle 9: Systematisierung der Sourcing-Operationalisierungen aus Prozess- und Produktperspektive

Umgang mit Metadaten	Prozesse	Produkte
Metadaten beachten/ benennen (s. Teilkap. 3.2.2)	1) Unspezifische Prozessdaten: Eye-Tracking- und Logfile-basiertes Sourcing • <i>Fixation von Metadaten in Eye-Tracking-Studien</i> (Brand-Gruwel et al., 2017; Gerjets et al., 2011; Gottschling & Kammerer, 2021; Gottschling et al., 2019; Hautala et al., 2018; Kammerer et al., 2013; Kammerer & Gerjets, 2014a; Mason, Pluchino & Ariasi, 2014; van Strien et al., 2016) • <i>Logfile-Daten zur Konsultationsfrequenz und -dauer von Metadaten</i> (Hahnel et al., 2019, 2021; List et al., 2017; Londra & Saux, 2023, Experiment 1; Merkt et al., 2017) • <i>Logfile-Daten zur Auswahl von verschieden vertrauenswürdigen Dokumenten</i> (Kammerer & Gerjets, 2014b; Schoor et al., 2024, Experimente 1 und 2; Winter & Krämer, 2012, Studie 2) • <i>Logfile-Daten zur Lesezeit von verschieden vertrauenswürdigen Dokumenten</i> (Bräten, McCrudden et al., 2018; Goldman et al., 2012; Kammerer et al., 2015; Kammerer & Gerjets, 2014b; Kiili et al., 2008; Mason, Junyent & Tornatora, 2014; Rodicio, 2015; Rott & Gavin, 2015; Salmerón et al., 2013; Wertgen & Richter, 2023, Experimente 1 und 2; Wiley et al., 2009, Experiment 1; Winter & Krämer, 2012, Studie 2) 2) Spezifische Prozessdaten: Benennen von Metadaten beim lauten Denken • <i>Identifikation von Metadaten</i> (Cho, Woodward & Li, 2018) • <i>Wiedergabe von Metadaten in Lautdenk-Protokollen</i> (Barzilai et al., 2015; Strømsø & Bräten, 2014; Strømsø et al., 2013)	1) Erinnerungsleistungen bei Metadaten • <i>geschlossene Aufgaben: Erinnerungen an Metadaten</i> (Delgado et al., 2019; Kiili et al., in Druck; Merkt & Huff, 2020, Experimente 1 und 2; Merkt et al., 2017; Salmerón, Sampietro & Delgado, 2020) • <i>geschlossene Aufgaben: korrekte Quelle-Inhalt-Verbindungen</i> (s. für Studien mit konfligierenden Dokumenten die Teilkap. 4.2.1.2 und 4.3.1.2) • <i>offene Aufgaben: korrekte Angabe von explizit erfragten Metadaten in offenen Fragen</i> (Bräten, Salmerón & Strømsø, 2016; Britt & Aglinskas, 2002, Experiment 3; Gottschling & Kammerer, 2021; Gottschling et al., 2019; Kammerer, Meier & Stahl, 2016; Kobayashi, 2014; Salmerón, Gil & Bräten, 2018a, 2018b) 2) Wiedergabe von Metadaten unter Nutzbarkeit der Dokumente in offenen Aufgaben <i>Korrektes Benennen von Metadaten unter Verfügbarkeit der Dokumente</i> (Coiro et al., 2015 = Forzani, 2018; Incognito & Tarchi, 2024; Macedo-Rouet et al., 2013, Experiment 1; Paul et al., 2018, Studie 1; Potocki et al., 2020)

Umgang mit Metadaten	Prozesse	Produkte
Metadaten verwenden (s. Teilkap. 3.2.3)	• <i>Herstellen von Quelle-Inhalt- und Quelle-Quelle-Verbindungen in Lautdenk-Protokollen</i> (Barzilai et al., 2015) • <i>Interpretieren: Inhalte von Dokumenten</i> (Anmarkrud et al., 2014; Strømsø et al., 2013)	• <i>Selektion:</i> Auswahl von Dokumenten/Quellen, teils in Kombination mit Relevanz der Dokumente (Bråten, Brante & Strømsø, 2018; Bråten, McCrudden et al., 2018; Hämäläinen et al., 2021; Hautala et al., 2018; Kammerer & Gerjets, 2012a; McCrudden et al., 2016) • <i>Belegen:</i> Anzahl von Referenzierungen in eigenen Texten (Florit, Cain & Mason, 2020; Rouet et al., 1997; Salmerón, Sampietro & Delgado, 2020; Strømsø & Bråten, 2014; Strømsø et al., 2020), teils mit spezifischen Analysen (Bråten et al., 2015; Kammerer & Gerjets, 2014a; Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2; Kiili, Brante et al., 2020; Londra & Saux, 2023, Experimente 1 und 2; Rouet et al., 1996; Salmerón, Gil & Bråten, 2018a; Strømsø et al., 2013) – s. auch Teilkap. 5.2.1.2 für Referenzierungen als Teil und Produktindikator von Diskurssynthesen • <i>Gewichten:</i> Nutzung von verschiedenen vertrauenswürdigen Dokumenten in eigenen Texten (Bråten, McCrudden et al., 2018; Londra & Saux, 2023, Experiment 1) • <i>Abstützen:</i> Nutzung von Metadaten, um eine eigene Position zu rechtfertigen (Bråten, Strømsø & Andreassen, 2016)
Metadaten beurteilen (s. Teilkap. 3.2.4)	1) Allgemeines Beurteilen: <i>allgemeine Beurteilungsaktivitäten in Lautdenk-Protokollen</i> (Goldman et al., 2012; Jay, 2021; Kammerer et al., 2013; Rott & Gavin, 2015; Ulyshen et al., 2015; Wineburg, 1991) 2) Spezifisches Beurteilen • <i>spezifisches Beurteilen von Metadaten in Lautdenk-Protokollen</i> (Anmarkrud et al., 2013, 2014; Barzilai & Zohar, 2012; Brand-Gruwel et al., 2017; Cho, Han & Kucan, 2018; Cho et al., 2017; Gerjets et al., 2011; Kammerer et al., 2021; Kiili et al., 2008; McGrew, 2021b; Strømsø & Bråten, 2014; Walraven et al., 2009) • <i>Nutzung verschiedener Kriterien der Glaubwürdigkeit einzelner Dokumente in Lautdenk-Protokollen</i> (von der Mühlen et al., 2016)	1) Beurteilungen und Begründungen der Einschätzung ganzer Dokumente • <i>geschlossene Aufgaben: Rangreihung von Dokumenten/Quellen mit Zuweisung einzelner Dokumente/Quellen entlang eines Glaubwürdigkeitskontinuums</i> (Abed & Barzilai, 2023; Anmarkrud et al., 2014; Braasch et al., 2013, 2023; Braasch, Bråten, Strømsø & Anmarkrud, 2014; Bråten et al., 2015; Denton et al., 2022, Studie 2; Goldman et al., 2012; Graesser et al., 2007, Experimente 1 und 2; Kessler et al., 2021; List & Alexander, 2018b; Mason, Junyent & Tornatora, 2014; Mason, Moè et al., 2022; Mason, Scrimin, Tornatora et al., 2018; Rott & Gavin, 2015; Rouet et al., 1996, 1997; Wiley et al., 2009, Experiment 1; Wineburg, 1991) • <i>geschlossene Aufgaben: Beurteilung von Dokumenten/Quellen in Fragebögen auf einer geschlossenen Skala, aber mit Möglichkeit, identische Werte zu verteilen</i> (Abel et al., in Druck; Barzilai & Eshet-Alkalai, 2015; Braasch et al., 2009; Flanagan & Metzger, 2007; Kammerer & Gerjets,

Umgang mit Metadaten	Prozesse	Produkte
		2012b; Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2; Kammerer, Meier & Stahl, 2016; Kanniainen et al., 2019; Kiili et al., 2022; Kobayashi, 2014; List et al., 2017; Londra & Saux, 2023, Experiment 1; Lucassen et al., 2013; Marttunen et al., 2021; Pieschl & Sivyver, 2021; Pilotek et al., in Druck; Tarchi, 2021b; van der Heide & Lim, 2016; van Strien et al., 2016; von der Mühlen et al., 2016; Westerwick, 2013, Studie 1; Winter & Krämer, 2012, Studie 2) • <i>Auswahl von Dokumenten mit Rangreihung der Vertrauenswürdigkeit aus einem Set von Websites</i> (Brand-Gruwel et al., 2017) • <i>Benennung der (un-)zuverlässigsten Dokumente</i> (Mason, Junyent & Tornatora, 2014) • <i>Beurteilung der Verwendbarkeit einzelner, definierter Websites</i> (Braasch et al., 2009; Delgado et al., 2019; Incognito & Tarchi, 2024; Macedo-Rouet, Salmerón et al., 2019; Potocki et al., 2020) • <i>offene Aufgaben: Begründungen für die (Glaubwürdigkeits-)Urteile zu Dokumenten/Quellen</i> (Barzilai & Eshet-Alkalai, 2015; Breakstone et al., 2022; Hämäläinen et al., 2021; Kiili et al., 2019, 2021; Wineburg & McGrew, 2019) bzw. <i>Aussagen in den Dokumenten</i> (Lucassen & Schraagen, 2011; Salmerón et al., 2016, Experimente 1 und 2) • <i>offene Aufgaben: Begründungen für die Reihungen von Dokumenten/Quellen in ihrer Glaubwürdigkeit</i> (Abed & Barzilai, 2023; Braasch et al., 2013, 2023; Braasch, Bråten, Strømsø & Anmarkrud, 2014; Bråten et al., 2015; Denton et al., 2022, Studie 2; Kessler et al., 2021; List & Alexander, 2018b; Mason, Junyent & Tornatora, 2014; Mason, Moè et al., 2022; Mason, Scrimin, Tornatora et al., 2018; Rouet et al., 1996, 1997; Wiley et al., 2009, Experiment 1, aber nur Auswertung einer Kategorie) • <i>offene Aufgaben: Begründungen für die Auswahl von Dokumenten/Quellen</i> (Bråten, Brante & Strømsø, 2018; List et al., 2016b; Stenseth et al., 2019; Strømsø et al., 2020) • <i>offene Aufgaben: Begründung für die Verwendbarkeit definierter Websites</i> (Delgado et al., 2019; Incognito & Tarchi, 2024; Macedo-Rouet, Potocki et al., 2019, Experimente 1 und 2; Potocki et al., 2020; Salmerón et al., 2019)

Umgang mit Metadaten	Prozesse	Produkte
		• offene Aufgaben: Begründungen für die Verwendung von Dokumenten (Lyu et al., 2023) • offene Aufgaben: Einschätzung der Vertrauenswürdigkeit von unbekannten Websites (Wineburg & McGrew, 2019) 2) Beurteilungen und Begründungen der Einschätzung von einzelnen Informationen und Aussagen • geschlossene Aufgaben: binäre Einschätzung der Zuverlässigkeit von Informationen einer Website (Coiro et al., 2015 = Forzani, 2018; Kiili, Leu, Marttunen et al., 2018; Lucassen & Schraagen, 2011) • geschlossene Aufgaben: Beurteilung der Vertrauenswürdigkeit von Informationen von gelesenen Dokumenten (Bråten et al., 2009; Bråten, Strømsø & Salmerón, 2011; Graesser et al., 2007, Experimente 1 und 2; Kammerer & Gerjets, 2014a; Strømsø et al., 2011; Tarchi, 2019, 2021a) • geschlossene Aufgaben: Beurteilung der Schlussfolgerungen in Dokumenten (Bråten, Salmerón & Strømsø, 2016; Strømsø & Bråten, 2017; Strømsø et al., 2017) • geschlossene Aufgaben: Einschätzung der Wichtigkeit einzelner metadatenbezogener Beurteilungskriterien für ein holistisches Glaubwürdigkeitsurteil von Dokumenten (Macedo-Rouet, Salmerón et al., 2019; von der Mühlen et al., 2016) bzw. Informationen (Bråten et al., 2009 = Bråten, Strømsø & Salmerón, 2011; Pieschl & Sivy, 2021; Strømsø et al., 2011; Tarchi, 2019, 2021a) • offene Aufgaben: Begründung für die Einschätzung der Evidenztypen von Aussagen (List et al., 2022) • offene Aufgaben: Begründungen für die Vertrauenswürdigkeitsurteile zu Informationen von gelesenen Dokumenten (Kiili, Leu, Marttunen et al., 2018) • offene Aufgaben: Begründungen für die inhaltlichen Entscheidungen in den eigenen Texten (Paul et al., 2018, Studie 2; Paul et al., 2019)

Umgang mit Metadaten	Prozesse	Produkte
		3) Einschätzungen einzelner Metadaten • <i>Beurteilung einzelner Metadaten auf einer geschlossenen Skala, aber mit Möglichkeit, identische Werte zu verteilen</i> (Anttonen et al., in Druck; Barzilai, Thomm & Shlomi-Elooz, 2020, Experimente 1 und 2; Braasch et al., 2009; Flanagan & Metzger, 2007; Gottschling & Kammerer, 2021; Gottschling et al., 2019, 2020; Herrero-Diz et al., 2019; Incognito & Tarchi, 2024 [keine Auswertung vorgenommen]; Kannianen et al., 2019; Kiili et al., 2022, 2023; Macedo-Rouet, Potocki et al., 2019, Experiment 2; Potocki et al., 2020; van der Heide & Lim, 2016; Wertgen & Richter, 2020, Experimente 1 und 2; Wertgen & Richter, 2023, Experiment 1; Westerwick, 2013, Studie 1) • <i>binäre bis nominale Einschätzung der Expertise von Autorinnen und Autoren von Texten</i> (Coiro et al., 2015 = Forzani, 2018; Kiili, Leu, Marttunen et al., 2018; Macedo-Rouet et al., 2013, Experiment 1) • <i>Auswahl von Gründen für die Evaluation verschiedener Metadaten aus einem Set vorgegebener Antworten</i> (Kiili et al., in Druck) • <i>offene Aufgaben: Begründungen für die Einschätzung einzelner Metadaten</i> (Barzilai, Thomm & Shlomi-Elooz, 2020, Experimente 1 und 2; Coiro et al., 2015 = Forzani, 2018; Kiili et al., 2022; Kiili, Leu, Marttunen et al., 2018; Macedo-Rouet et al., 2013, Experiment 1; Marttunen et al., 2021)

(Quelle: eigene Darstellung, partiell basierend auf der Unterscheidung von Gruppen bei Anmarkrud et al., 2022, S. 772)

Wie aus der gemäß der Sourcing-Definition dreigeteilten Spalte „Umgang mit Metadaten“ ganz links hervorgeht, dient die Untergliederung zugleich als Strukturierung für die folgende Ergebnisdarstellung in den Teilkapiteln. Die Aufteilung des Umgangs mit Metadaten in Beachten, Verwenden und Beurteilen in der Tabelle 9 folgt damit der theoriebasierten Unterscheidung von Prozessen des Sourcings. Zusätzlich – und das spiegeln die Operationalisierungen aus Prozess- und Produktwarte wider – ist eine gewisse zeitliche Logik in der Abfolge der Makroprozesse des Beachtens, Verwendens und Beurteilens impliziert. Solche zeitlichen bzw. präziser formuliert: in ihrem konsekutiven Ablauf sachlogisch aufeinander folgenden Sequenzen sind in Makrophasen-Modellierungen ebenfalls vorgelegt worden (s. z. B. Metzger, 2007), zu denen im weitesten Sinn ebenfalls das DISCM aus dem Teilkapitel [3.1.2.3](#) gezählt werden kann. Exemplarisch kann das illustriert

werden an der Dual-Prozess-Theorie von Metzger (2007). Gemäß der grundsätzlichen Logik von (sozialpsychologisch grundierten) Dual-Prozess-Theorien verarbeiten Personen in Abhängigkeit diverser in- und externer Prozesse eher heuristisch bzw. analytisch. Metzgers (2007) Theorie ist aus einer chrono- bzw. sachlogischen Warte heraus anschlussfähig an die drei oben angeführten Makroprozesse. Sie benennt drei Phasen: erstens der *Kontakt mit einer Website* (mit einer Nähe zum Beachten der Metadaten), zweitens die *Evaluationsphase* (also die Verwendung und mehrdimensionale Prüfung von Metadaten, die gemäß dem Modell heuristisch bzw. analytisch erfolgt) und drittens die eigentliche *Beurteilungsphase* (mit ihrem Urteil und vor allem den Begründungen des Urteils). Diese Phasen wirken in dieser Darstellung diskret und abgeschlossen, jedoch ist realiter eher von Rekursionen und dynamisch interagierenden Prozessgruppen auszugehen. Die Dreiteilung ist entsprechend in Bezug auf die analytische Trennbarkeit fruchtbarer denn als unidirektionale Prozesstrias. Darum wird in der Darstellung der Forschungsbefunde darauf Wert gelegt, die Zusammenhänge zwischen verschiedenen Indikatoren verschiedener Prozessgruppen dort zu berichten, wo sie in den Studien zum Gegenstand der Analyse geworden sind.

3.2.2 Metadaten beachten bzw. benennen

3.2.2.1 Produkte

Produktmaße zum Beachten und Benennen von Metadaten werden entweder als Erinnerungsleistung oder als Aufsuchen der richtigen Metadaten in den Studien operationalisiert. Im Fall der Erinnerungsleistungen lassen sich offene von geschlossenen Tests differenzieren. Das übergreifende Muster der Studien besteht darin, dass eine korrekte Wiedergabe von Metadaten bei nutzbaren Dokumenten größer ist, als wenn Personen aus dem Gedächtnis heraus Metadaten berichten – und hierbei sind geschlossene Aufgaben mit Antwortvorgaben mit besseren Leistungen assoziiert, als wenn die Erinnerungen offen erfragt wurden. Die zeigen im Einzelnen die nun folgenden Ergebnisse im Detail.

1) Erinnerungsleistungen bei Metadaten

Geschlossene Aufgaben: Erinnerungen an Metadaten

In geschlossenen Aufgaben erzielen die Testpersonen in aller Regel relativ hohe Werte. Das ergibt sich sowohl aus Studien, in denen die Anzahl korrekter Lösungen aufsummiert werden (s. Tabelle 10), als auch in jenen Fällen, in denen das

Verhältnis von richtigen zu falschen Antworten bestimmt wird (Delgado et al., 2019; Salmerón, Sampietro & Delgado, 2020). Aus der Tabelle 10 geht ebenfalls hervor, dass die Leistungen in offenen Tests markant niedriger ausfallen.

Tabelle 10: Vergleich von Studien mit offenen und geschlossenen Varianten, Metadaten gedächtnisbasiert zu erfassen

Offene Aufgaben ohne vorgegebene Antwortmöglichkeiten		Geschlossene Aufgaben mit vorgegebenen Antwortmöglichkeiten	
Studie	Prozentualer Anteil korrekter Antworten	Studie	Prozentualer Anteil korrekter Antworten
Bråten, Salmerón und Strømsø (2016)	7,5	Merkt & Huff (2020, Experiment 1)	70,2
Kammerer, Meier und Stahl (2016) – Gruppe ohne Prompts	25,6	Merkt & Huff (2020, Experiment 2)	68,3
Kobayashi (2014)	23,6	Merkt et al. (2017)	56,8
Salmerón, Gil und Bråten (2018a)	34,3		
Salmerón, Gil und Bråten (2018b) – beide Gruppen	50,9		
Mittelwert	28,4	Mittelwert	65,1

(Quelle der Darstellung: eigene Berechnungen)

Von den wenigen Studien mit geschlossenen Antworten ist aus empirischer Sicht jene von Salmerón, Sampietro und Delgado (2020) hervorzuheben. In dieser Studie wurden weitere Produktmaße ermittelt, die in Korrelationsanalysen zur Erinnerungsleistung an Metadaten ins Verhältnis gesetzt wurden. Die Zusammenhänge gestalteten sich bei den Maßen unterschiedlich. Wer sich besser zu-treffend an Metadaten erinnern konnte, konnte auch mehr Quelle-Inhalt-Ver-bindungen korrekt angeben, bildete aber weder mehr oder weniger Inferenzen in Antworttexten noch neigten die Personen dazu, mehr oder weniger Belege in ihren Texten anzuführen. Demnach wirkt es so, als sei das Memorieren von Quelleninformationen stärker mit der kognitiven Repräsentation des Inhalts assoziiert, von anderen Verstehensleistungen aber unabhängig, vor allem denen aus schriftlichen Verstehensprodukten.

Offene Aufgaben: korrekte Angabe von explizit erfragten Metadaten in offenen Fragen

Die offenen Aufgabenformate wirken in den Studien nur auf den ersten Blick über ihre Methodik der Erhebung homogen. Tatsächlich verbergen sich in der Auswertungsvariante teils erhebliche Unterschiede, die zu divergierenden Be-funden führen:

- *Summenwerte in Tests mit diversen potenziell nennbaren Informationen verweisen auf geringe Sourcingleistungen.* Die Ermittlung von Erinnerungsleistungen mit Summenwerten aus korrekt genannten Metadaten (Kammerer, Meier & Stahl, 2016; Kobayashi, 2014) bzw. mit noch feineren Abstufungen dieser Erinnerungsleistungen (Bräten, Salmerón & Strømsø, 2016) ist bislang eher selten durchgeführt worden. Die Ergebnisse, die in Tabelle 10 zusammengetragen sind, verweisen allerdings darauf, dass das Maximum der Leistungen bei einem Viertel des theoretisch erreichbaren Punktwerts lag.
- *Die Art der Erfassung von Metadaten in Studien hängt damit zusammen, wie viele Metadaten korrekt genannt werden.* Dies wird deutlich in jenen Studien, welche die Erinnerungen mittels gänzlich freien und zusätzlich mit Hinweisen zu den Kategorien versehenen Fragen hintereinander erfassten (Gottschling & Kammerer, 2021; Gottschling et al., 2019). Im erstgenannten Fall sind die Personen gebeten worden, alle ihnen erinnerlichen Metadaten niederzuschreiben. Im zweitgenannten Fall wurden Kategorien wie Name, Land, institutionelle Anbindung etc. explizit als Hilfestellung benannt. Die Leistungen bei der zweiten Variante der Erfassung waren laut deskriptiver Statistik jeweils besser, auch wenn die statistische Signifikanz der Unterschiede in den beiden Operationalisierungen nicht eigens geprüft wurde.
- *Personen erinnern sich an verschiedene Arten von Metadaten unterschiedlich genau.* Dies wurde bisher lediglich in den Studien von Kobayashi (2014) und Britt und Aglinskas (2002, Experiment 3) berichtet. In Kobayashis Studie wurden zwei sich inhaltlich widersprechende Texte genutzt, und die offene Aufgabe bezog sich auf beide Texte. Kobayashi verdichtete die Antworten zu fünf Kategorien, die hier für beide Dokumente gebündelt berichtet werden. Mit im Mittel 88 Prozent war die Erinnerung an den Beruf am genauesten, gefolgt von dem Namen des Autors (54 Prozent). Diese beiden Angaben standen als einzige über den jeweiligen Texten. Das Publikationsjahr (19 Prozent), der Publikationsort (12 Prozent) und der Titel des Dokuments (10 Prozent) blieben demgegenüber weniger in Erinnerung. Diese drei Metadaten befanden sich unter dem jeweiligen Text. Die Erinnerungsleistungen wirken demnach abhängig von der Positionierung der Metadaten. Die sozialwissenschaftlich verortbare Interventionsstudie von Britt und Aglinskas (2002) testete bei der geförderten Gruppe nach dem Posttest in einer Intervention, an wie viele unterschiedliche Metadaten (bzw. genauer: Bestandteile des sog. Dokumentenknotens, s. Teilkap. [4.1.2.3](#)) sich die Jugendlichen eine Woche nach einer Fördermaßnahme mit einer digitalen Sourcing-App erinnerten. In dieser App standen sieben Dokumente mit deutlich erkennbaren Metadaten in einer digitalen Bibliothek zur Verfügung. Ausgewiesen wurden die Anteile der Personen, die sich an einzelne Metadaten korrekt erinnert hatten. Der Name des Autors/der Autorin (55 Prozent), die (berufliche) Position dieser Person (82 Prozent) und das Datum des Dokuments (64 Prozent)

wurden von mehr als der Hälfte der Testpersonen wiedergegeben. Andere Metadaten wie die Motive der Autorin bzw. des Autors (45 Prozent), die Perspektive zum zeithistorischen Geschehen (36 Prozent) oder der Dokumententypus (27 Prozent) waren nur noch kleineren Teilen der Stichprobe präsent.

Mögen diese Studien bisher die Ausnahme sein, so legen sie doch nahe, dass es sich lohnt, genauer zu differenzieren, welche Metadaten salienter sind und in Memorierungsleitungen häufiger auftauchen.

- *Studien mit eher liberaler Auswertung zeichnen ein anderes Bild mit hohen Sourcingleistungen bei Erinnerungsleistungen.* Dieses Befundmuster stammt aus Studien von Gottschling et al. (2019) sowie Gottschling und Kammerer (2021). In diesen Studien wurde nicht bestimmt, wie viele Metadaten korrekt wiedergegeben worden waren. Stattdessen wurden anhand der (in)korrekten Antworten ermittelt, wie hoch der Anteil der Personen war, die für die beiden jeweils verwendeten kurzen Dokumente Metadaten korrekt memoriert hatten. Gezählt wurde also, dass ausreichend genau, aber nicht wie häufig Metadaten richtig waren: Es genügte damit, dass von mehreren erfragten Metadaten nur eine Antwort pro Dokument zutreffend war. In beiden Studien wurden die Metadaten in sehr freier Form und zusätzlich noch in einer weiteren Form mit Hinweisen zur Kategorie von Metadatum (wie Name oder institutionelle Zugehörigkeit) erfragt. Da sich diese Studien dem Verstehen konfligierender Dokumente widmeten und gezielt die Metadaten wie der Expertise variiert wurden, wurden mehrere Experimentalgruppen unterschieden. Die Anteile der Personen, die ausreichend korrekt beide Dokumente mit Metadaten erinnert hatten, lag bei der offenen Aufgabe in beiden Studien zwischen 48 und 67 Prozent; bei der Variante mit Hinweisen zwischen 56 und 81 Prozent. Diese letztgenannten Quoten entsprachen relativ genau dem Bereich der Antworten, der in einer weiteren Studie (Gottschling & Kammerer, 2021) mit einem geschlossenen Test mit Quelle-Inhalt-Verbindungen erfasst worden war, nämlich zwischen 69 und 83 Prozent.

Diese Ergebnisse auf der Basis offen erfasster Erinnerungsleistungen machen darauf aufmerksam, dass die Auswertung entscheidend dafür ist, ob die Leistungen als niedrig oder hoch zu gelten haben. Liberalere Tests kommen zu anderen Ergebnissen als spezifischere und genauere Auswertungen. Mit dieser Genauigkeit der Analysen deutet sich zudem an, dass verschiedene Metadaten unterschiedlich gut in Erinnerung bleiben – dies ist ein erhebliches Forschungsdesiderat.

2) Wiedergabe von Metadaten unter Nutzbarkeit der Dokumente in offenen Aufgaben

Einige Studien maßen keine Erinnerungsleistungen, sondern gewährten den Testpersonen Zugriff auf die Dokumente. Dabei wurde analog zu den Studien mit den offenen Fragen entweder nur ein Metadatum erfragt, nämlich die Namen der Verfasserin bzw. des Verfassers (Coiro et al., 2015; Macedo-Rouet et al., 2013, Experiment 1; Potocki et al., 2020), oder aber mehrere weitere Metadaten wie das Erscheinungsjahr mussten korrekt berichtet werden (Incognito & Tarchi, 2024; Paul et al., 2018, Studie 1). Die studienübergreifende Auswertung der Ergebnisse ist nur bedingt möglich. Teils liegt das daran, dass keinerlei deskriptive Statistik berichtet wird (Incognito & Tarchi, 2024; Potocki et al., 2020). Dort, wo sie ausgewiesen wurde, sind die Anteile richtiger Antworten bzw. damit auch der Personengruppe mit der richtigen Antwort hoch: Sie liegen je nach Population in der Studie zwischen 81 bis 90 Prozent (Coiro et al., 2015; Macedo-Rouet et al., 2013, Experiment 1; Paul et al., 2018, Studie 1).

3.2.2.2 Prozesse

Die folgende Darstellung der Prozessdaten folgt einer Zweiteilung, die sich vordergründig aus den Datenquellen bzw. den Operationalisierungen des Erkennens von Metadaten speist, indem Eye-Tracking- und logfilebasierte Daten in einem ersten Schritt und Lautdenk-Protokollanalysen in einem zweiten Schritt fokussiert werden. Diese Trennung ergibt sich daraus, dass im ersten Fall erkennbar ist, dass Sourcing stattgefunden hat, der zweite methodische Fokus – zumal in der Kombination mit weiteren Prozessdaten – erlaubt hingegen potenziell vertiefte Einsichten. Diese analytische Trennung wird auch später in diesem Buch noch eine Rolle spielen (s. Teilkap. [5.2.1.1](#)).

1) Unspezifische Prozessdaten: Eye-Tracking- und Logfile-basiertes Sourcing

Die Darstellung der Prozessdaten aus der Tabelle 9 im Teilkapitel [3.2.1](#) insinuiert eine gewisse Homogenität bei der Operationalisierung über nicht-reaktive Verfahren. Die Studien fallen jedoch in Erkenntnisinteressen und den sich daraus ergebenden Operationalisierungen so heterogen bis disparat aus, dass sich für eine lesbare Darstellung der Befunde Aggregationen als nötig erweisen.

Perspektive 1: Funktionen des Sourcings

Die Logfile-Daten – und teils die Eye-Tracking-Indikatoren – erteilen als nicht-reaktive und vor allem zeitlich, d. h. in ihrer Sequenz gut auswertbare Datenquellen Auskunft über verschiedene Zeitpunkte, wann Sourcingaktivitäten auftauchen,

was wiederum Informationen darüber gibt, zu welchem Zweck das Sourcing erfolgt. Besonders deutlich wird dies an der Unterscheidung, die Hahnel et al. (2019) in ihrer logfilebasierten Operationalisierung vorgenommen haben. Sie differenzieren neben einem allgemeinen Sourcing-Indikator drei Subformen: 1) proaktives Sourcing, 2) aufgabenbezogenes Sourcing und 3) wiederholtes Sourcing.

- 1) *Proaktives Sourcing*: Die erste Form wird deshalb proaktiv genannt, weil die Konsultation von Metadaten erfolgt, bevor ein Dokument gelesen wird und damit als heuristischer Indikator für den mutmaßlich folgenden Inhalt dient (Hahnel et al., 2019). Dies ist anschlussfähig an jene Studien, in welchen die Konsultationsmuster der Metadaten vor dem Lesen der Dokumente erfasst wurden (Hahnel et al., 2021; List et al., 2017), bzw. an die Eye-Tracking-Studien, die Muster bei der Fixation der Metadaten vor der Auswahl von potenziell zu lesenden Dokumenten prüften, z. B. ob Personen überhaupt die Metadaten vor der Auswahl der Dokumente betrachtet hatten (Kammerer & Gerjets, 2014b, Studien 1 und 2).
- 2) *Aufgabenbezogenes Sourcing*: Diese Variante reservieren Hahnel et al. (2019) für jene Anlässe, in denen es Teil der Aufgabe war, nach spezifischen Metadaten als Teil des Dokuments zu suchen. Prozessbezogene Studien dazu sind noch rar, doch die produktbezogenen Studien mit ihren Fragen zu einzelnen Metadaten (und damit ein relevanzbasiertes Verarbeiten der Metadaten, s. dazu das Teilkapitel „Offene Aufgaben: korrekte Angabe von explizit erfragten Metadaten in offenen Fragen“, [3.2.2.1](#)) sind Ausdruck eines derartigen Sourcings. Aus den Studien mit Prozessen lässt sich als direktes Beispiel jene von Hahnel et al. (2019) benennen. Die Studie von Brand-Gruwel et al. (2017) lässt sich indirekt dieser Variante des Sourcings zuordnen, denn in der Studie sollten Personen aus einem Set von Websites die fünf vertrauenswürdigsten selektieren, und es wurde per Eye-Tracking geprüft, ob Metadaten in den ausgewählten Dokumenten fixiert wurden.
- 3) *Wiederholtes Sourcing*: Diese dritte Form zeichnet sich dadurch aus, dass Metadaten mehrfach konsultiert (und prozessiert) werden. Diesem rekursiven Sourcing kommt die Funktion zu, die Verbindung zwischen Quelle und Inhalt zu koaktivieren und zu intensivieren, was über die erneute Zuwendung zu den Metadaten erfolgt (Hahnel et al., 2019). Speziell bei der kognitiven Verarbeitung von konfligierenden Dokumenten wird das Bearbeiten von Kohärenzbrüchen mithilfe des Sourcings im Allgemeinen und hier der Zuwendung zu den Metadaten im Besonderen theoretisch postuliert und empirisch untersucht (s. Teilkap. [4.2.2](#)). Wenn Personen mehrfach Metadaten entweder betrachten (Gottschling et al., 2019; Hautala et al., 2018) oder konsultieren (Hahnel et al., 2019) bzw. wenn die Studien Gesamtdauern beim Eye-Tracking (Gerjets et al., 2011; Kammerer et al., 2013; Mason, Pluchino & Ariasi, 2014; Salmerón, Gil & Bråten, 2018b; van Strien et al., 2016) bzw.

in den Logfile-Daten (Londra & Saux, 2023, Experiment 1) ausweisen, kann aufgrund längerer und häufigerer Konsultationen der Metadaten angenommen werden, dass Personen sie im Arbeitsgedächtnis reaktivieren.

Die drei genannten Funktionen des Sourcings – sofern sie sich mit unspezifischen Indikatoren gesichert als Funktion bezeichnen lassen – sind hilfreich, die operativen Zugänge zum Konstrukt als Rahmen zu nutzen. In den Studien zeichnet sich ab, dass das wiederholte Sourcing einen Schwerpunkt bildet, während das aufgabenbasierte Sourcing auffällig selten untersucht wurde.

Perspektive 2: Aggregate im Navigationsverhalten

Die Taxonomie zu den Funktionen des Sourcings bildet freilich nur einen Ausschnitt dessen ab, wie das Sourcing mit nicht-reaktiven Prozessmaßen erfasst wird. In anderen Studien, zumal solchen mit Logfile-Daten, werden Prozesse des Sourcings anders abgebildet, indem das Navigationsverhalten in den Blick gerät. Denn nicht-reaktiv erhobene Prozessmaße erlauben es, *die zeitliche Zuwendung bzw. die Auswahl von einzelnen Dokumenten zu untersuchen*, die a priori eine Zuweisung zu einem Kontinuum oder zu Gruppen von unterschiedlich vertrauenswürdigen Dokumenten erhalten haben. Indem es also definierte Differenzen gibt, wird es möglich, anhand der Prozessdaten Rückschlüsse darauf zu ziehen, dass Metadaten als heuristische Hinweise genutzt wurden, um eine hinreichend genaue Einschätzung zu treffen (Fogg, 2003). In dieser Hinsicht ähnelt es dem proaktiven Sourcing aus der Aufzählung oben. Allerdings wird hier über die erfolgte Auswahl und die relative Zuwendung kein direktes Maß wie beim proaktiven Sourcing operationalisiert, sondern ein eher indirektes. Denn die treffsicherere Auswahl von vertrauenswürdigen Dokumenten bzw. die längere Verarbeitungszeit sind lediglich als Indizien für erfolgtes Sourcing zu werten. Inwiefern es zu evaluativen bzw. lesestrategischen Entscheidungen gekommen ist, ist ohne weitere Daten kaum zu entwirren. Hat man jedoch weitere Daten, ergeben sich Muster.

- Ein erstes Muster stammt aus Studien, die zeigen, dass Auswahlhäufigkeiten (Schoor et al., 2024, Experimente 1 und 2; Winter & Krämer, 2012, Studie 2) und Lesezeiten (Bräten, McCrudden et al., 2018; Rodicio, 2015; Winter & Krämer, 2012, Studie 2) bei Dokumenten von Autorinnen und Autoren mit unterschiedlicher Expertise systematisch zugunsten fachlich geeigneterer Quellen variieren. Ähnliche Resultate erbrachten Studien, in denen die allgemeine Glaubwürdigkeit der angebotenen Dokumente variiert worden war (Goldman et al., 2012; Kammerer et al., 2015; Kammerer & Gerjets, 2014b, Studien 1 und 2; Rott & Gavin, 2015; Salmerón et al., 2013). Eine Tendenz in den Ergebnissen dieser Studien ist somit, dass *(leistungsstärkere) Personen*

bei hoher Salienz der Differenzen der Glaubwürdigkeit der Dokumente oder durch optimierte Darstellungen in Suchmaschinen dazu neigen, glaubwürdige Dokumente häufiger bzw. länger zu rezipieren. Komplementär dazu ist der Befund aus dem Paradigma der Validierungsforschung, nach welchem weniger glaubwürdige Dokumente länger gelesen werden, was als Unterbrechung automatisch ablaufender Leseprozesse gewertet wird (Wertgen & Richter, 2023, Experiment 1, partiell Experiment 2).

- Andere Studien wiederum nutzten Daten aus dem Navigationsverhalten dazu, den Wechsel zwischen unterschiedlich glaubwürdigen Websites auszuzählen und daraus Verhältnisse zu bilden (Goldman et al., 2012; Wiley et al., 2009, Experiment 1), den Anteil der Lesezeit von verschiedenen vertrauenswürdigen Websites auszuweisen (Kiili et al., 2008) und aggregierte Rekursionsmuster zu abstrahieren (Wiley et al., 2009, Experiment 1). Hierdurch ist es möglich, die proportionale Nutzung verschieden glaubwürdiger Dokumente mittels einer Kennzahl zu quantifizieren. Diese Kennzahlen hängen systematisch positiv mit Leistungsdaten zusammen (Wiley et al., 2009, Experiment 1) bzw. differenzieren Leistungsgruppen (Goldman et al., 2012). Das bedeutet: *Je stärker sich in den Logfiles ein Verhalten manifestiert, welches eine Präferenz der glaubwürdigeren Dokumente nahelegt, als desto günstiger schlägt sich dies in Leistungs(gruppen)daten nieder.*

Wenn das Navigationsverhalten bei verschiedenen Dokumenten vergleichend analysiert wird oder aber aus dem Navigationsverhalten Muster extrahiert werden, ergeben sich weitere Perspektiven auf das Sourcing, die zusammengefasst eine Sensitivität für Dokumenten- bzw. Personenunterschiede nahelegen.

Zusammenhänge unspezifischer Prozessdaten mit anderen Prozessdaten

In der bisherigen Darstellung waren die einzelnen Prozesse nach Funktionen bzw. nach Aggregationsgrad resp. Zusammenhang mit anderen Variablen separat ausgewiesen. In vereinzelt Studien wurden sie in den Zusammenhang mit anderen Prozessdaten gestellt. Durch die insgesamt große Forschungslücke ist es kaum möglich, generelle Aussagen zu treffen, doch gibt es erste Hinweise:

- Je länger Personen die URLs von Webseiten in Suchmaschinen betrachteten, desto mehr evaluative Kommentare äußerten sie in Bezug auf die konkret fixierten Ergebnisse (Kammerer et al., 2013). Das Inspizieren von Metadaten geht also damit einher, dass es zu mehr Beurteilungen kommt.
- Die Zeiten, die Personen auf verschiedenen vertrauenswürdigen Websites verbringen, korrelieren dergestalt, dass die Lesezeit bei glaubwürdigeren Dokumenten zulasten jener bei unglaubwürdigeren Quellen geht (Kammerer et al., 2015). Anders gesagt: Die Personen widmen sich eher einer Art von

Gruppen von Websites: entweder eher den glaubwürdigeren oder den weniger glaubwürdigen.

- Proaktives Sourcing und Switches zwischen den Dokumenten korrelieren positiv (Hahnel et al., 2021): Wer häufig Metadaten vor dem Lesen des Inhaltes eines Dokuments anwählt, neigt demnach dazu, häufiger zwischen den Dokumenten zu wechseln, was als Hinweis auf ein textvergleichendes Lesen (Corroboration) wertbar ist.

Selbst bei der zu konzедierenden geringen Menge an erforschten Zusammenhängen eint die Befunde, dass Sourcingleistungen positiv mit anderen Prozessen korreliert sind. Je selektiver und fokussierter in Bezug auf die Wahrnehmung von Metadaten die Leserinnen und Leser vorgehen, desto günstiger wirkt ihr Verhalten bei anderen Prozessen.

2) Spezifische Prozessdaten: Benennen von Metadaten beim lauten Denken

Lautdenk-Studien sind bislang selten durchgeführt worden, dennoch offerieren die wenigen Untersuchungen bereits reichhaltige und differenzierte Befunde, die sich aus den zum Teil hochspezifischen Untersuchungsfokussen und Analysemethoden ergeben. Verdichten lassen sich aus diesen Studien folgende Punkte:

- *Das Erkennen von Metadaten kommt verglichen mit anderen Prozessen des Sourcings häufiger vor.* Dies ist das (deskriptive) Ergebnis aus Studien, die verschiedene Formen des Sourcings bei konfligierenden Dokumenten (Barzilai & Eshet-Alkalai, 2015; Strømsø & Bråten, 2014; Strømsø et al., 2013) oder das Erkennen von Metadaten in offenen Settings (Cho, Woodward & Li, 2018) als Teil der interessierenden Verbaldaten auswerten. Besonders markant wird das in der Studie von Barzilai et al. (2015), es ist die Studie mit der bislang feinteiligsten Analyse der verschiedenen Sourcingprozesse. Das Forschungsteam ermittelte, dass von allen in ihrer Funktion differenziert kodierten Sourcingaktivitäten das Repräsentieren (Wiedergeben) von Metadaten fünf Sechstel allen Sourcings ausmachte, was statistisch signifikant häufiger vorkam als das Evaluieren von Metadaten.
- *Sourcing kann implizit oder explizit erfolgen, die Anteile variierten jedoch deutlich.* Drei Studien (Barzilai et al., 2015; Strømsø & Bråten, 2014; Strømsø et al., 2013) prüften, ob die Probandinnen und Probanden die Metadaten genau in ihren Lautdenk-Protokollen benannt hatten (als „explizites Sourcing“ bezeichnet) oder ob sie zwar erkennbar Metadaten repräsentierten, diese aber nicht genau wiedergaben (darum als „implizites Sourcing“ deklariert). Durch diese binäre Unterscheidung lassen sich die Anteile von ex- und impliziten Benennungen von Metadaten ausweisen. Die Anteile der expliziten Sourcingäußerungen betrugen je nach Studie hohe 58 Prozent, was sowohl statistisch

signifikant (Strømsø & Bråten, 2014) als auch nicht-signifikant war (Strømsø et al., 2013), oder niedrige 21 Prozent (Barzilai et al., 2015, statistisch signifikanter Unterschied). Abgesehen von diesen Differenzen, die auch mit den hohen Differenzen der absolut getätigten Sourcingäußerungen zu tun haben mögen, lässt sich zumindest attestieren, dass ein genaues Benennen der Metadaten beim lauten Denken kein durchgängiger Modus ist. Diese Differenz von im- und expliziten Sourcing betrifft ebenfalls das Sourcing in Produkten (s. dazu Teilkap. [3.2.3.1](#) und dort den Abschnitt „Belegen“).

- *Die jeweiligen Metadaten werden unterschiedlich häufig benannt.* Welche Metadaten in den Lautdenk-Protokollen explizit oder implizit aufscheinen, ist selten fokussiert und in den Studien zudem unterschiedlich operationalisiert worden. Dennoch deuten sich Differenzen dahingehend an, dass Personen Perspektive und Expertise häufiger thematisieren als Publikationszeitpunkt und mögliche Motivation der Quelle (Barzilai et al., 2015). Wenn andere Metadaten kodiert wurden, so zeigt sich, dass Personen Publikationsorgane und Institutionen sowie Dokumentenart teils gleichrangig, teils häufiger als die Person des Autors bzw. der Autorin ansprechen (Strømsø & Bråten, 2014). Dies ist mitunter noch zusätzlich damit verschränkt, auf welches Dokument sich die Personen in ihren Äußerungen konkret bezogen haben.
- *Das Repräsentieren von Metadaten variiert nach jeweiligem Dokument.* Obwohl viele Studien, zumal aus der Multiple-Documents-Literacy-Forschung, gezielt konfligierende Dokumentensets mit experimentell manipulierten Metadaten einsetzen, ist relativ unterbelichtet, welche Metadaten bei welchem Dokument repräsentiert werden. Vereinzelt jedoch gibt es Untersuchungen, die auf Dokumentenebene solche Auswertungen vornehmen (Strømsø & Bråten, 2014; Strømsø et al., 2013). Diese Studien kommen zum Ergebnis, dass nicht alle Dokumente gleich viele Sourcingaktivitäten beim Benennen hervorrufen. So elizitierten beispielsweise die Dokumente mit der stärksten inhaltlichen Exponierung zu einem kontroversen Sachverhalt deutlich mehr sourcingbezogene Äußerungen als neutrale. Hinzu kommt, dass innerhalb des Sourcings pro Dokument unterschiedliche Gewichtungen vorkamen. So wurden je nach Dokument Schwerpunkte beim Sourcing erkennbar, die sich teils auf die Person des Autors, das Publikationsorgan, die publizierende Organisation bzw. den Dokumententypus bezogen. Diese Ergebnisse unterstreichen, dass Leserinnen und Leser eine gewisse Sensitivität aufweisen, worauf sie ihre Aufmerksamkeit beim Sourcing richten.
- *Es ist noch unklar, ob das alleinige Erkennen von Metadaten systematisch mit anderen Verarbeitungen gemäß Lautdenk-Protokollen zusammenhängt.* Die Studien sind hierfür in ihrer Zahl zu wenig vorgelegt worden und fallen in ihrer Konsistenz der Befunde disparat aus. Von besonderem Belang ist in diesem Zusammenhang, dass das erste Erkennen von Metadaten, welches bereits erste oberflächliche Anteile des Evaluierens in sich trug, in Chos,

Woodwards und Lis Studie (2018) mit Jugendlichen auf spezifische Weise mit anderen Prozessmaßen korrelierte: Es hing nämlich vor allem mit weiteren Sourcingprozessen des ersten Einschätzens von Metadaten und dem Abgleich mit dem eigenen Vorwissen zusammen. Es handelte sich um in der Verarbeitung sachlogisch eher früh auftretende Prozesse. Andere Zusammenhänge des Erkennens von Metadaten mit tiefenorientierten und analytischen Prozessen ließen sich in besagter Studie nicht nachweisen. Die Studie von Barzilai et al. (2015) mit Studierenden kam hingegen zum Schluss, dass das Repräsentieren (Wiedergeben) von Metadaten mit anderen Facetten des Sourcings (darunter dem Verbinden von Inhalt und Quelle sowie dem Vergleich von Quellen) bzgl. der Häufigkeit korrelierte. Die Korrelationen waren hier allerdings durchgängig und substantziell, d. h. Repräsentieren und andere Sourcingprozesse waren eng miteinander verknüpft.

Die Lautdenk-Protokoll-Studien verdeutlichen, dass das Erkennen von Metadaten im Verlauf der Bearbeitung von lesebezogenen Aufgaben ein mitunter häufiger, in sich jedoch heterogener Sourcingprozess ist. Personen divergieren in dem Grad, wie sehr sie explizit oder implizit die Metadaten benennen, welche Metadaten sie beachten und bei welchen Dokumenten sie es tun. Die Studien selbst geben im Gesamt noch kein klares Bild ab, ob das Erkennen von Metadaten systematisch und musterhaft mit anderen Prozessen korrespondiert. Es existiert mithin eine Forschungslücke.

3) Zusammenhänge von Prozess- und Produktmaßen

Die Verbindungen von Prozessindikatoren und Produktmaßen sind in der bisherigen Forschung nur sporadisch ausgewiesen worden und bestehen aus vielen fragmentarischen Ergebnissen, die sich kaum zu einem kohärenten Gesamtbild zusammenfügen lassen. Darum erfolgt an dieser Stelle notgedrungen eine reine Aggregation zu Nullergebnissen und zu positiven Zusammenhängen.

- *Nullergebnisse:* Die reine Konsultation von Metadaten gemäß Logfiledaten führt nicht zu besseren Erinnerungsleistungen bei Metadaten in einem geschlossenen Test (Merk et al., 2017). Ähnliches gilt für die Effekte von Fixationszeiten und nicht das Sourcing direkt betreffenden Produktindikatoren wie ausgewogenen inhaltlichen Darstellungen (Kammerer et al., 2013). Die Lesezeit von verschiedenen glaubwürdigen Websites weist zudem keinen korrelativen Zusammenhang mit der inhaltlichen Entscheidung in eigenen schriftlichen Antworten (Kammerer et al., 2015) oder Wissenstests (Rodicio, 2015) auf.
- *Positive Zusammenhänge:* Das wiederholte Sourcing (wiederholtes Fixieren der Metadaten gemäß Eye-Tracking) prädiziert zum einen Erinnerungsleistungen in offenen Aufgaben (Gottschling & Kammerer, 2021).

Zum anderen hängt es nur partiell mit Erinnerungsleistungen zusammen (Salmerón, Gil & Bråten, 2018b). Die Studie von Salmerón, Gil und Bråten (2018b) ist darüber hinaus wichtig, weil sie demonstrierte, dass offenbar nicht jedes der vielen eigens untersuchten Sourcingmaße mit den Erinnerungsleistungen in einem positiven korrelativen Zusammenhang steht. Darum stellt sich die Frage, welche Art der Eye-Tracking-Maße theoretisch begründbar mit welchem Produktindikator positiv oder negativ korreliert. Navigationsmaße wie die Tendenz, verstärkt die glaubwürdigen Websites aufgesucht zu haben, korrelieren mit besseren Rangreihungsleistungen (Wiley et al., 2009, Experiment 1). Das bedeutet: Sich mehr im Gesamtprozess der Aufgabenbearbeitung den glaubwürdigeren Websites gewidmet zu haben, findet eine Entsprechung darin, später die glaub- von den weniger glaubwürdigen Dokumenten treffsicher zu unterscheiden. Daneben gibt es positive Effekte von Fixations-, Auswahl- und Konsultationsmaßen des Sourcings hinsichtlich der Referenzierungen in eigenen schriftlichen Aufgabenprodukten. Diese sind näher beschrieben in Teilkapitel [5.2.2.2](#) (und dort im Abschnitt „Spezifische Leseprozesse 2: Sourcingprozesse“).

Alles in allem bestehen erste Hinweise darauf, dass das Erkennen von Metadaten mit positiven Effekten auf später erhobene Produktmaße einhergeht. Generalisierende Aussagen sind gleichwohl derzeit angesichts des Mangels an empirischen Befunden sowie den vielen Nullergebnissen nur eingeschränkt möglich.

3.2.3 Metadaten nutzen

Produkt- und Prozessmaße des Sourcings aus der Funktion des Nutzens von Metadaten wirken gemäß der Darstellung in Tabelle 9 randständig. Dieser erste Eindruck täuscht jedoch, denn das Nutzen von Metadaten ist ein Bereich der Forschung mit relativ viel Empirie, deren Befunde zum Zweck der Redundanzvermeidung in diesem Buch jedoch an zwei anderen Stellen vorgestellt werden. Zum einen wird das Nutzen von Metadaten als *Mittel der Kohärenzwiederherstellung* erachtet und erforscht. Dies ist Gegenstand von Teilkapitel [4.2.2.2](#). Dort wird deutlicher werden, dass das Vorhandensein von intertextuellen Konflikten mit einer erhöhten Aktivität im Sourcing und mehr Referenzierungen in Produkten einhergeht. Zum anderen ist das Sourcing sowohl aus Prozess- als auch als Produktwarte Bestandteil des *Herstellens von Diskurssynthesen* (s. Teilkap. [5.2.1.2](#)). Die prozessfokussierte Empirie, die sich aus diesem Forschungsbereich speist, wird entsprechend darum im Teilkapitel [5.2.2.2](#) behandelt. Zum gegenwärtigen Stand der Forschung ist das sonstige Nutzen der Metadaten klar ein Desiderat, sodass die vorliegenden Studien nur fragmentarisch Auskunft geben, wie Personen – jenseits des Bemerkens und des Evaluierens – mit Metadaten umgehen.

3.2.3.1 Produkte

Die Nutzung der Metadaten ist ein heterogenes Feld, aus dem die Studienergebnisse in vier Bereichen bzw. Funktionen präsentiert werden: erstens bei der Auswahl von Dokumenten, zweitens beim Belegen von Aussagen mit Metadaten, drittens beim Gewichten von Aussagen bzw. Dokumenten und viertens als Abstützung für Entscheidungen.

Selektion

Die Frage, ob Variationen in der Auswahl von Dokumenten mit der Expertise der Autorinnen und Autoren erklärbar sind, ist in einer Handvoll Studien genauer untersucht worden. Die Resultate sind eindeutig: *Dokumente mit einer hohen Expertise der Quelle wählen Personen generell häufiger aus*, und zwar sowohl bei vertrauten als auch unvertrauten Themen (Bråten, McCrudden et al., 2018; McCrudden et al., 2016). Dabei gibt es teils Interaktionseffekte von Relevanz des Inhalts und Expertise der Quelle, wie es Bråten, McCrudden et al. (2018) und McCrudden et al. (2016) verdeutlichten: Bei unvertrauten Themen selektieren Personen eher solche Dokumente, bei denen die Quelle eine hohe Expertise aufweist und das Dokument inhaltlich relevant ist. Bei thematisch wenig relevanten Dokumenten ist der Effekt der Expertise hingegen weniger bedeutend. Ein wichtiger Befund beider Studien bestand darin, dass der Effekt der Expertise bei der Auswahl bei den unvertrauten Themen deutlich stärker war als bei den vertrauten Themen (Bråten, McCrudden et al., 2018; McCrudden et al., 2016). Anders gesagt: Bei den unvertrauten Themen wählten Personen mehr Dokumente aus, die von Personen mit höherer Expertise stammten. Diese verstärkte Zuwendung zu den Dokumenten mit hoher Expertise bei Themen mit geringer Vertrautheit des Themas bezeichnen McCrudden et al. (2016) als *Hypothese der Gebundenheit an die Autorinnen und Autoren*. Sie steht in Einklang mit der Gebundenheit von Laienpersonen an epistemische Autoritäten bei Themengebieten, in denen es eine gesellschaftlich hohe kognitive Arbeitsteilung gibt (Bromme & Goldman, 2014).

Teils wurden Expertise und Relevanz nicht als einzelne Faktoren betrachtet, sondern miteinander kombiniert. Studien mit diesem Zugang fanden die Tendenz, sich auf hohe Expertise und thematische Eignung gleichermaßen zu stützen, ebenfalls vor (Hämäläinen et al., 2021). Damit deutet sich an, dass inhaltliche und metadatenbezogene Einschätzungen zusammenwirken.

Belegen

Leserinnen und Leser zitieren mehr glaubwürdige als weniger glaubwürdige Dokumente, sie stützen sich also mehr auf glaubwürdigere Quellen und weisen dies mit Referenzierungen nach. Dies wurde bei naturwissenschaftlichen Themen daran deutlich, dass neutrale Dokumente aus staatlichen Institutionen überproportional häufig zitiert wurden (Bråten et al., 2015; Kiili, Brante et al., 2020, Thema 2). Diese Tendenz wurde sowohl für Dokumenten- als auch für eingebettete Quellen beobachtet (Londra & Saux, 2023, Experimente 1 und 2). Zur Wahrheit gehört indes auch, dass Personen teils ungeeignete ideologisch-religiöse Organisationen zitieren, also ein eher problematisches Sourcing betreiben (Kiili, Brante et al., 2020, Thema 1).

Die Zuwendung zu den vertrauenswürdigen Dokumenten spiegelt sich in Studien aus der Domäne Geschichte, also in den *Gesellschaftswissenschaften*, wider. Denn es ergaben sich unterschiedlich starke Zuwendungen zu einzelnen Dokumenten, etwa eine starke Nutzung und Referenzierungen von offiziellen Zeitdokumenten bei historischen Themen (Rouet et al., 1996, 1997). Anders gesagt: Die Dokumente mit der größten zeitlichen Nähe zum historischen Ereignis wurden überproportional häufig zitiert.

Gewichten

Nur sehr wenige Studien prüften, ob Dokumente aus Quellen mit unterschiedlicher Expertise in schriftlichen Antworten in unterschiedlichen Verhältnissen vorkamen. Es gibt es jedoch Hinweise, dass Personen in eigenen Antworten sensibel sind für die Expertise gemäß der Metadaten von Dokumenten: Personen nutzten mehr Dokumente von Autorinnen und Autoren mit höherer Expertise für ihre Antworten, wobei sie analog zu den Befunden bei der Selektion bei unvertrauten Themen eine deutliche Präferenz für Dokumente haben, die inhaltlich relevant sind und von einer Person mit hoher Expertise stammen (Bråten, McCrudden et al., 2018).

Abstützen

Ob die eigene Position in schriftlichen Dokumenten mit Metadaten unterfüttert wird, ist eine erstaunlich selten erforschte Frage. In puncto Gesundheitsentscheidungen ermittelten diesbezüglich Bråten, Strømsø und Andreassen (2016), dass die von ihnen untersuchten Personen nach dem Lesen eine gesundheitsrelevante Frage in keinsten Weise mit den Metadaten aus den gerade zuvor gelesenen Dokumenten begründeten.

3.2.3.2 Prozesse

Die Verwendung von Metadaten aus prozessualer Sicht ist wenig erforscht. Dennoch gibt es erste Hinweise aus Studien mit Lautdenk-Protokollen, dass Leserinnen und Leser Metadaten und Quellen als Ganzes nutzen, um

- den Inhalt von Dokumenten anhand der Quelleninformationen zu prognostizieren (Vorhersagen; Strømsø et al., 2013);
- den Inhalt von Dokumenten mit deren Ursprung zu verbinden (Quelle-Inhalt-Verbindungen herstellen; Barzilai et al., 2015), z. B. um den Inhalt zu interpretieren, was bei Anmarkrud et al. (2014) ein gewichtiger Bestandteil einer übergeordneten Sammelkategorie (Evaluieren) war und bei Strømsø et al. (2013) eine eigene, wenn auch empirisch wenig vorkommende Kategorie bildete;
- die Perspektiven von Quellen bzw. deren Aussagen miteinander zu vergleichen (Quelle-Quelle-Verbindungen aufbauen; Barzilai et al., 2015).

Das bedeutet: In den drei oben berichteten Studien lassen sich anhand der Kodierungen verschiedene Funktionen nachweisen (s. dazu auch Cho & Afflerbach, 2017). Durch die divergierenden Operationalisierungen und die Aggregation der Daten sind jedoch kaum studienübergreifende Befundmuster ableitbar.

Vereinzelt wurde die prozessuale Verwendung von Metadaten mit anderen Prozessmerkmalen ins Verhältnis gesetzt: Barzilai et al. (2015) stellten substantielle Zusammenhänge der Quelle-Inhalt- bzw. Quelle-Quelle-Verbindungen untereinander fest: Wer das eine mehr tut, tut es auch bei der anderen Kategorie. Außerdem korrelierten sowohl die Quelle-Inhalt- als auch die Quelle-Quelle-Verbindungen stark mit anderen Sourcingaktivitäten, die sich auf Metadaten bezogen. Anders gesagt: Sie bilden aus variablenzentrierter Warte einen Bestandteil eines umfassenden Sourcings. Die Zusammenhänge der Verwendung von Metadaten in Prozessen und Produktindikatoren sind ebenfalls noch in anderen Studien erforscht worden. Da Produktmaße aber häufig über längere schriftliche Antworten in Diskurssynthesen gewonnen wurden, werden sie an anderer Stelle berichtet (s. Teilkap. [5.2.2.2](#) und dort Abschnitt 2) Sourcingprozesse).

3.2.4 Metadaten evaluieren

Das Beurteilen der Metadaten ist der am umfassendsten erforschte Bereich des Sourcings und wird vor allem aus der Produktperspektive heraus intensiv untersucht. Darum wird die folgende Forschungsübersicht hierauf einen sehr deutlichen Akzent legen ([3.2.4.1](#)). Die Prozessperspektive ist Gegenstand einer sehr viel kürzeren Darstellung ([3.2.4.2](#)).

3.2.4.1 Produkte

Die Beurteilung von Produkten lässt sich gemäß dem Fokus der qua Sourcing zu beurteilenden Einheiten differenzieren: 1) ganze Dokumente (samt Begründungen), 2) einzelne Aussagen aus Dokumenten (samt Begründungen) sowie 3) die Metadaten selbst. Diese bereits in der Tabelle 9 im Teilkapitel [3.2.1](#) dargestellte Logik betrifft bereits die Erfassung auf der Ebene der Erhebungsinstrumente bzw. der Operationalisierung bei der Datengewinnung. Die Art der Rohdatenerfassung bildet naturgemäß lediglich einen ersten Schritt. Wie die Daten analysiert und zuvor transformiert wurden, ist hingegen ein Folgeschritt, bei dem verschiedene Studien mit identischer Datengewinnung unterschiedliche Wege beschreiten, Variablen zu bilden. Aus diesem Grund und dem Schwergewicht in der produktbezogenen Empirie bei der Evaluation von Metadaten werden in diesem Teilkapitel die Methodik bei der Variablenbildung und die Befunde jeweils blockweise dargestellt, um so die Komplexität im Gegenstand und dessen Entfaltung auf ein Minimum zu beschränken.

1) Beurteilungen und Begründungen der Einschätzung ganzer Dokumente

Die Beurteilung von ganzen Dokumenten als Einheit von Text und Metadaten hat in der Forschung besondere Zuwendung erfahren. Darum ist die Zahl an Studien vergleichsweise hoch. Die in diesem Abschnitt vorgenommene Zusammenschau der empirischen Befunde erfolgt zweigeteilt. In einem ersten Schritt werden die Ergebnisse zur Glaubwürdigkeitsbeurteilung behandelt (1a), um im Anschluss daran auf die Begründungen zu fokussieren (1b).

1a) Zuweisen von Glaubwürdigkeit für gesamte Dokumente

Zur Methodik: Welche Produktindikatoren bei den Glaubwürdigkeitsurteilen werden erfasst bzw. aus den Daten gebildet?

Die Einschätzung von Glaubwürdigkeit ganzer Dokumente wird in Studien üblicherweise darüber vorgenommen, dass Personen nach dem Lesen multipler Dokumente auf einer Skala taxieren, wie sie die Glaubwürdigkeit entweder eindimensional oder operationalisiert über mehrere Dimensionen beurteilen. Die Einschätzung auf einer eindimensionalen Skala kann dabei auf zweierlei Art erfolgen. Zum einen gibt es Studien, welche die Personen mittels einer Rangreihung dazu auffordern, die Dokumente vom glaubwürdigsten bis zum unglaublichsten Dokument zu ordnen. Der Punktwert des Dokuments entspricht damit seiner Position der individuellen Einschätzung auf einem Kontinuum, und jedes Dokument kann damit nur einen Wert zugewiesen bekommen, das heißt, es gibt keine Möglichkeit, Dokumente als gleichermaßen vertrauenswürdig

zu erachten. Das ist – zum anderen – bei der zweiten Variante möglich. Denn ebenso gibt es Studien, die es den Probandinnen und Probanden gestatten, für jedes Dokument einen Wert zu vergeben, der unabhängig von einer Rangreihung ist. Damit zwingt die erste Variante dazu, die gesamte Skala auszunutzen und forciertere Entscheidungen zu treffen.

Aus diesen Werten der Glaubwürdigkeitsurteile – sei es mit forcierter Rangreihung und damit Werten bei einzelnen Dokumenten, die durch die Reihung von anderen Dokumenten abhängig sind, sei es mit unabhängigen und numerisch potenziell durchaus deckungsgleichen Urteilen – werden je nach Studie unterschiedliche Indikatoren gebildet resp. Auswertungen vorgenommen:

- 1) *Vergleich aller Dokumente bzw. bestimmter Gruppen untereinander.* Hierbei werden die Mittelwerte der zu beurteilenden Dokumente untereinander verglichen und signifikante Unterschiede zwischen den einzelnen Dokumenten ermittelt (Abed & Barzilai, 2023; Anmarkrud et al., 2014; Bråten et al., 2009, 2015; Flanagan & Metzger, 2007; Lucassen & Schraagen, 2011; Strømsø et al., 2011). Teils wird nur die deskriptive Statistik zu den Mittelwerten der holistischen Urteile zu dem jeweiligen Dokument in den Experimenten berichtet (Braasch et al., 2023; Denton et al., 2022, Studie 2; List & Alexander, 2018b; Tarchi, 2019, 2021a, 2021b). Entscheidend ist bei aller Unterschiedlichkeit: In diesen Studien sind mit verschiedenen Granularitätsgrößen Vergleiche zwischen Dokumenten(gruppen) möglich.
- 2) *Vergleich der individuellen Leistungen mit dem Urteil von Personen mit hoher Expertise.* Eine andere Form der Auswertung besteht darin, dass die Rangreihungen der Testpersonen mit einer Musterlösung von Expertinnen und Experten abgeglichen und die Übereinstimmung der Lösungen bepunktet werden (Brand-Gruwel et al., 2017; Mason, Scrimin, Tornatora et al., 2018; von der Mühlen et al., 2016). Eine hohe Punktzahl ergibt sich in diesem Falle aus einer Annäherung an das Urteil von Expertinnen und Experten.
- 3) *Vergleich von Differenzwerten a priori festgelegter Dokumentengruppen im Sinne einer Diskriminanz (Gruppentrennung).* Andere Studien nutzen aggregierte Maße, in denen die Urteile zur Glaubwürdigkeit zusammengefasst wurden. Dabei werden in der Regel zwei Mittelwerte gebildet, die sich auf vorab definierte Dokumente beziehen. Der erste Mittelwert wird aus den Glaubwürdigkeitsurteilen aller glaubwürdigeren bzw. als besser geeignet geltenden Dokumente berechnet, der zweite aus den Urteilen zu den weniger glaubwürdigen bzw. geeignet wirkenden Dokumenten. Vom ersten Wert wird der zweite subtrahiert, was dann einen Differenzwert ergibt. Je größer dieser Differenzwert ausfällt, desto besser sind die Personen dazu in der Lage, Gruppen von Dokumenten zu unterscheiden, also zu diskriminieren. Diverse Studien bedienen sich dieser Auswertungsmethodik (Abel et al., in Druck;

Braasch, Bråten, Strømsø & Anmarkrud, 2014; Braasch et al., 2009, 2013; Macedo-Rouet, Salmerón et al., 2019; Pieschl & Sivyer, 2021; Wiley et al., 2009, Experiment 1).

- 4) *Summenwerte der Benennung bzw. Beurteilung der vertrauenswürdigen bzw. am wenigsten vertrauenswürdigen Websites.* Diese Variante bildet eine Art Vorstufe jener Analysen aus dem Spiegelstrich zuvor. In vorab definierten Website-Zusammenstellungen sollen die Vertreter auf den Endpunkten des Vertrauenswürdigkeitsspektrums benannt werden. Für jede korrekte Antwort erhalten die Studienteilnehmerinnen und -teilnehmer eine Punktzahl, die aufsummiert wird. Dadurch wird erfasst, wie gut Personen differenzieren können, und zwar anhand zweier Punktwerte. Dieses Verfahren ist vergleichsweise selten angewendet worden (Fendt et al., 2023; Graesser et al., 2007, Experimente 1 und 2; Mason, Junyent & Tornatora, 2014).

Die obige Aufzählung verdeutlicht, dass die Glaubwürdigkeitseinschätzungen einzelner Dokumente auf unterschiedliche Weise ausgewertet werden können. Dabei ist entscheidend, ob die einzelnen Dokumente die Einheit des fokalen Interesses bilden (Nr. 1 in der Liste oben) oder ob die Trennung von weniger und stärker vertrauenswürdigen Dokumenten fokussiert werden soll, was mit einem Vergleich der Lösungen mit einem Expert:innen-Benchmark erfolgt (Nr. 2), als Differenzwert im Sinne der Trennschärfe ausgedrückt wird (Nr. 3) oder – deutlich seltener – als Mittelwert für jede Gruppe (Nr. 4). Damit sind letztlich zwei Aussagen anhand solcher Analysen möglich: Zum einen lässt sich die Glaubwürdigkeit einzelner Dokumente aus Sicht der Testpersonen eruieren, zum anderen ermöglichen die reduktiven Verdichtungen auf einzelne Indikatoren weitere Analysen mit einer erheblich geringeren Anzahl von (abhängigen) Variablen.

Ergebnisse: Welche Muster bei der Beurteilung von Dokumenten gibt es?

Die Studien, in denen die Teilnehmerinnen und Teilnehmer zur Beurteilung ganzer Dokumente aufgefordert wurden, haben unterschiedliche Auswertungen vorgenommen. Diese beziehen sich erstens auf die Rangreihungsergebnisse der unterschiedlich glaubwürdigen Dokumente und die Zusammenhänge zwischen den Urteilen. Zweitens wurden Personenunterschiede bereits mehrfach geprüft. Beide Perspektiven werden in der nachfolgenden Darstellung berücksichtigt.

- *Personen weisen verschiedenen Dokumenten in Globalurteilen der Glaubwürdigkeit unterschiedliche Rangplätze zu und differenzieren vor allen an den Polen.* Diese Differenzierung gelingt insbesondere an den Endpunkten der experimentell zusammengestellten Dokumentensets, also bei den Dokumenten, deren Glaubwürdigkeit ganz gezielt so manipuliert wurde, dass sie sehr (un-) glaubwürdig wirken sollen (Anmarkrud et al., 2014; Braasch, Bråten, Strømsø

& Anmarkrud, 2014; Bråten et al., 2009; Denton et al., 2022, Studie 2; Graesser et al., 2007, Experimente 1 und 2; Kammerer, Meier & Stahl, 2016, Kontrollgruppe; Rouet et al., 1997). Teils sind die Personen allerdings nicht dazu in der Lage, die Dokumente in ihrer Glaubwürdigkeit ausreichend voneinander trennen zu können (Macedo-Rouet, Potocki et al., 2019, Experiment 1, teils auch Experiment 2; Mason, Junyent & Tornatora, 2014, Kontrollgruppe). Bei den Dokumenten im Mittelfeld mit weniger eindeutig zu fällenden Glaubwürdigkeitsurteilen liegen die Urteile dichter beieinander (Anmarkrud et al., 2014; Kammerer, Meier & Stahl, 2016, Kontrollgruppe; Mason, Scrimin, Tornatora et al., 2018). Diese geringe Trennschärfe im mittleren Segment wirft die Frage auf, warum sich Glaubwürdigkeitsurteile in der Mitte so stark ähneln, was wiederum die Begründungen für die Urteile in ihrem Erklärungsbeitrag als weiteres Evaluationsmaß betont.

- *Neutralen Informationsquellen mit hoher inhaltlicher Expertise wird eine höhere Glaubwürdigkeit – sowohl in Urteilen über das Gesamtdokument als auch bei der Glaubwürdigkeit der Informationen – attestiert.* Dies zeigt sich darin, dass informierende Überblickstexte (Anmarkrud et al., 2014; Bråten et al., 2009, 2015; Tarchi, 2019, 2021a, 2021b), Dokumente von staatlichen Organisationen bzw. Institutionen (Abed & Barzilai, 2023; Anmarkrud et al., 2014; Bråten et al., 2009, 2015; Kammerer & Gerjets, 2014a; Mason, Scrimin, Tornatora et al., 2018; Tarchi, 2019, 2021a, 2021b), Forschenden (Strømsø et al., 2011; Tarchi, 2019, 2021a, 2021b) bzw. auf ein Thema spezialisierten Personen (Denton et al., 2022, Studie 2), empirische Beiträge bei Wissenschaftsthemen (von der Mühlen et al., 2016) sowie Qualitätsmedien (Abed & Barzilai, 2023; Flanagan & Metzger, 2007; List & Alexander, 2018b) gegenüber anderen Dokumenten höhere Werte bei der Einschätzung erhielten. Darin deutet sich eine systematisch aufscheinende Sensibilität an, Dokumenten mit spezifischen Metadaten, die auf Neutralität bzw. Expertise verweisen, mehr Glaubwürdigkeit zu unterstellen. Dies steht in Einklang mit den beiden zentralen Dimensionen Expertise und Vertrauenswürdigkeit, die aus theoretischer Sicht als besonders bedeutend für Glaubwürdigkeitseinschätzungen gelten (s. Teilkap. [3.1.3.2](#)).
- *In Dokumentensets mit kontroversen Dokumenten werden jene Dokumente als geringer in der Glaubwürdigkeit eingeschätzt, die stark eine eigennützige bzw. extreme Position vertretend bzw. dabei noch stark polemisch wirken.* Diesen Effekt beobachteten Forschungsteams bei solchen Themen wie Gefahren von Handystrahlung (Anmarkrud et al., 2014; Bråten et al., 2015; Mason, Scrimin, Tornatora et al., 2018), Impfungen (Tarchi, 2019, 2021a, 2021b), Klimawandel (Bråten et al., 2009; Strømsø et al., 2011), Wirksamkeit von Nahrungsergänzungsmitteln (Kammerer & Gerjets, 2014a) und Veränderlichkeit von Arbeitsgedächtnisleistungen (Braasch et al., 2023). Es sind im Lichte dieser Befunde also nicht nur die Metadaten selbst, die anscheinend

mit den Glaubwürdigkeitsurteilen zu tun haben, sondern überdies die Inhalte bzw. deren Ausgewogenheit.

- *Die Glaubwürdigkeitsurteile bestehen teils unabhängig voneinander, teils sind sie untereinander korreliert – entscheidend wirken eine ähnliche Glaubwürdigkeit bzw. die inhaltlichen Positionen der Dokumente.* Auffällig ist, dass Personen dazu neigen, glaubwürdige Dokumente bzw. Informationen aus den glaubwürdigsten Dokumenten sehr ähnlich zu beurteilen, das heißt, es gibt hohe Interkorrelationen zwischen den Glaubwürdigkeitseinschätzungen bei diesen Dokumenten (Bråten et al., 2009; Tarchi, 2019, 2021a, 2021b). Allerdings ist dies nicht durchgängig ermittelbar gewesen (Anmarkrud et al., 2014).

In einigen Studien ließ sich anhand der *inhaltlichen Positionen der Dokumente* zusätzlich prüfen, ob die Perspektiven auf den Inhalt (in Verbindung mit weiteren Metadaten) eine Rolle für die Zusammenhänge spielen. So hat Tarchi (2019, 2021a) – nicht aber Tarchi (2021b) – beim Thema Impfen festgestellt, dass Personen gegenläufige Glaubwürdigkeitsurteile abgeben: Wenn sich impfkritische Personen (in diesem Fall ein Blogger bzw. ein Alternativmediziner) äußern, werden diese Dokumente negativer beurteilt, während die Testpersonen offiziellen positiv oder neutral berichtenden Dokumenten eine höhere Glaubwürdigkeit attestierten. Ähnliches berichten Anmarkrud et al. (2014) für ein Dokumentenset zur Handystrahlung, in dem das auffallend kritische Dokument in seinem Urteil negativ mit den Rangreihungen anderer Dokumente korrelierte. Noch prononcierter erwies sich dieses Muster in der Studie von Braasch et al. (2023): Bei Dokumenten, die die Verbesserung des Arbeitsgedächtnisses durch ein Training aus gegenläufigen Perspektiven darstellten, hingen die Glaubwürdigkeitsurteile negativ zusammen. Die Einschätzungen von neutralen Texten sind ebenfalls negativ korreliert mit den Glaubwürdigkeitseinschätzungen von Dokumenten, die entweder positive oder negative Perspektiven auf den Gegenstand entworfen haben. Sind die Darstellungen komplexer und uneindeutiger – so wie beim Thema Klimawandel in der Studie von Bråten et al. (2009) – sind die Glaubwürdigkeitsurteile insgesamt positiv korreliert. Je weniger gegenläufig die inhaltlichen Positionen sind, desto homogener scheinen die Urteile in ihrer Gänze auszufallen.

- *Personengruppen, die anhand verschiedener Merkmale gebildet werden können, unterscheiden sich in ihren Urteilen.* Ein erster Marker für Differenz ist die Unterscheidung von *Expertinnen und Experten* sowie *Laienpersonen*, deren Ergebnisse miteinander verglichen werden. Erstgenannte Personengruppe trifft präzisere Vertrauenswürdigkeitsurteile bei der Bestimmung von Websites in Suchmaschinen (Brand-Gruwel et al., 2017), unbekannten Websites mithilfe von Recherchen (Wineburg & McGrew, 2019), Sets von historischen Dokumenten (Wineburg, 1991) oder kurz präsentierten Dokumenten rund um ein psychologisches Thema (von der Mühlen et al., 2016). Außerdem sind sich Expertinnen und Experten in ihren Urteilen als Gruppe

wesentlich einiger, als Laienpersonen es sind, die stark divergierende Glaubwürdigkeitsurteile fällen (Wineburg, 1991). Zu guter Letzt zeigen sich Expertinnen und Experten sensibler beim Inhalt als Laien. Wenn Dokumente Fehler enthalten, attestieren Expertinnen und Experten den Dokumenten eine geringere Glaubwürdigkeit, während Laienpersonen undifferenziertere Glaubwürdigkeitsurteile abgeben (Lucassen & Schraagen, 2011).

Ebenfalls untersucht wurden – zweitens – Personen mit *unterschiedlichen Wissensbeständen*. Rouet et al. (1997) berichteten davon, dass Studierende mit mehr Domänenwissen Primär- und Sekundärtexte zu historischen Themen anders beurteilen als Personen mit geringerem Vorwissen: Personen mit höherem Wissen sprachen Primärtexten eine höhere Glaubwürdigkeit zu und Lehrbuchtexten eine geringere. Fendt et al. (2023) ermittelten leicht positive Effekte von formaler Ausbildung und der Fähigkeit, vertrauenswürdige und weniger vertrauenswürdige Dokumente einzuschätzen. In ihrer Studie mit Studierenden differenzierten Bråten, Strømsø und Salmerón (2011) zwei Gruppen, nämlich eine mit hohem und eine mit geringem Wissen. Personen mit höherem Vorwissen unterschieden sich im Vergleich zu Personen mit geringerem Vorwissen zwar nicht durchgängig in ihren Glaubwürdigkeitsurteilen über die Dokumente, aber sie waren kritischer gegenüber den Dokumenten mit erkennbar einseitigen Interessen, z. B. Aussagen von Ölfirmen zur Thematik Klimawandel, und gestanden Quellen wie Zeitungen und Forschungszentren eine höhere Glaubwürdigkeit zu. Es gibt also Differenzen, die sich an einzelnen Dokumenten und deren Evaluationen manifestieren. Zudem bestehen – drittens – Differenzen zwischen *Altersgruppen*. So sind Sekundarschuljugendliche des siebten Schuljahrs im Gegensatz zu solchen aus neunten und elften Klassen nicht dazu in der Lage, vertrauenswürdige und nicht-vertrauenswürdige Internetquellen voneinander zu unterscheiden (Pieschl & Sivyer, 2021). Zu ähnlichen Befunden kamen Abel et al. (in Druck) mit Schülerinnen und Schülern der siebten gymnasialen Jahrgangsstufe und solchen aus der Sekundarstufe II.

Erwähnenswert ist in diesem Zusammenhang viertens, dass Einschätzungen zu den vertrauenswürdigeren Dokumenten *Personen mit kognitiven Beeinträchtigungen* eher schwer fallen (Delgado et al., 2019; Salmerón et al., 2019). Damit lässt sich festhalten: Jüngere Personen, solche mit kognitiven Beeinträchtigungen und Novizinnen und Novizen bzw. Personen mit geringerem Vorwissen sind in Glaubwürdigkeitseinschätzungen weniger leistungstark.

Insgesamt lässt sich konstatieren: Die Beurteilung der Glaubwürdigkeit von Dokumenten ist vor allem als relatives Maß erforscht worden, indem multiple Dokumente einzuschätzen waren. Personen unterscheiden Dokumente trennschärfer, wenn die Dokumente aufgrund der Metadaten und/oder ihrer inhaltlichen Positionierung unterscheidbar wirken. Neutralen Dokumenten von epistemischen

Autoritäten sprechen Leserinnen und Leser eine höhere Glaubwürdigkeit zu. Dieses Muster wird teils deutlich in Studien, die nicht nur Rangreihungen, sondern auch korrelative Zusammenhänge geprüft bzw. berichtet haben. Hinzu kommen noch Personenunterschiede: Ältere Personen, solche ohne diagnostizierte kognitive Beeinträchtigungen und solche mit höherem bis sehr hohem Vorwissen können Dokumente besser in ihrer Glaubwürdigkeit einschätzen.

1b) Begründungen für Einschätzungen, Entscheidungen und Reihungen

Diverse Studien bedienen sich offener oder auch geschlossener Fragen, um die verschiedenen Urteile zu den Dokumenten begründen zu lassen. Dabei variieren die Bezugspunkte, nämlich warum Personen Dokumente ausgewählt oder sie für verwertbar gehalten haben, sie in ihrer Glaubwürdigkeit mit einem bestimmten Wert versehen haben oder das Dokument als Ganzes oder einen Teil davon – eine Information, eine Schlussfolgerung, aber auch ein Metadatum – beurteilt haben. Demnach unterscheiden sich die Studien in ihrem evaluativen Bezugspunkt. Sie divergieren zusätzlich darin, wie die mittels geschlossener oder offener Fragen erhobenen Begründungen ausgewertet wurden.

Zur Methodik: Welche Produktindikatoren bei den Begründungen werden erfasst bzw. aus den Daten gebildet?

Die Urteile zu den Begründungen werden mit offenen oder geschlossenen Instrumenten gewonnen, wobei die offenen Aufgabenformate nicht nur häufiger zum Einsatz gekommen sind, sondern bei ihnen auch mehr und unterschiedliche Varianten der Datenauswertung genutzt wurden. Zunächst zu den geschlossenen Formaten, deren Ansatz sich auf den allgemeinsten Nenner gebracht wie folgt beschreiben lässt: *Geschlossene Formate fokussieren auf das Bestimmen eines Wertes innerhalb einer definierten Range, sodass die variablenzentrierten Zugänge die Lage zwischen Minima und Maxima fokussieren.* Dieser Zugang bei der Erhebung und Auswertung der Begründungen ist damit nicht weit entfernt von Globaleinschätzungen von Dokumenten sowie den sich daraus ergebenden Analysemöglichkeiten. Studien, die diesen Zugang gewählt haben, sind in der Tabelle [11](#) zuoberst im Teil A dargestellt (ab S. [173](#)).

Die *offenen Formate*, um Begründungen zu elizitieren, sind nicht homogen in der Analyse verwendet worden. Je nach Erkenntnisinteresse und Anlage der Studien lassen sich die Studien verorten, und zur Strukturierung dienen die folgenden Leitfragen:

- *Welche Arten von Begründungen und damit welche Gruppen von Begründungen lassen sich in offenen Aufgaben wiederfinden?* Ein häufig gewählter Zugang in Studien war, aufgrund inhaltsanalytischer Kodierungen der

Begründungen mindestens zwei Kategorien, nämlich inhalts- und quellenbasierte Begründungen, auszuzählen und damit ein metrisches Maß für Begründungen zu erhalten (z. B. Bråten, Brante & Strømsø, 2018; Strømsø et al., 2020). Zwei Kategorien bilden damit die Minimalvariante, um die Begründungen zum einen zu vergleichen und zum anderen, um die Auftretensmuster zu bestimmen. Weitere Studien sind noch differenzierter, weil sie mehr als zwei Subkategorien benennen und teils noch auf Subkategorieebene ausweisen, wie häufig die einzelnen Nennungen waren. Ein Beispiel hierfür ist die Interventionsstudie von Braasch et al. (2013). In ihr formierten sieben Subkategorien die Hauptkategorie inhaltsbezogener Begründungen (in der Reihenfolge ihrer Häufigkeit (in der Kontrollgruppe) waren dies: 1) Güte der Erklärung, 2) Verständlichkeit, 3) Stil, 4) Interessantheit, 5) Glaubhaftigkeit, 6) Relevanz, 7) argumentative Qualität). Bei der Hauptkategorie der quellenbasierten Begründungen wurden folgende fünf Subkategorien addiert: a) Person des Autors/der Autorin, b) qualifizierende Merkmale der Autorin/des Autors, c) Dokumententypus, d) Publikationsorgan sowie e) Datum der Veröffentlichung. Diese Zugänge in der Auswertung sind also heterogen und haben damit selbstverständlich Auswirkungen darauf, welche inhaltliche Aussagekraft die Studien haben. Insgesamt ist es anhand der Daten möglich, Rangreihungen bei den Häufigkeiten der einzelnen Begründungen vorzunehmen, und diese Ergebnisse enthält Tabelle 11 im unteren Teil B (ab S. 174).

- *Nominal oder metrische Variablen bei den rein sourcingbezogenen Begründungen: Begründen die Personen überhaupt bzw. in welchem Ausmaß mit Metadaten?* Die Auswertungen bei den offenen Antworten variieren stark. Dies hat damit zu tun, dass die Studien bei den Auswertungen unterschiedliche Granularitätsgrößen gewählt haben, wie es die Anzahl der Kategorien in der Tabelle 11 verdeutlicht. Dass die Begründungen erhoben wurden, heißt aber nicht, dass sie auf gleiche Weise ausgezählt und analysiert wurden. So gibt es auf der einen Seite solche Studien, die zwar mehrdimensional die Begründungen in offenen Fragen erfassen, aber entweder nur die Antworten in Bezug auf quellenbezogene Begründungen auswerten, indem sie diese als metrische (Barzilai & Eshet-Alkalai, 2015; Hämäläinen et al., 2021; Kiili et al., 2019, 2021; Mason, Junyent & Tornatora, 2014; Wiley et al., 2009, Experiment 1) oder binäre Variable behandeln (Abed & Barzilai, 2023; Incognito & Tarchi, 2024; Macedo-Rouet, Potocki et al., 2019; Macedo-Rouet, Salmerón et al., 2019; Marttunen et al., 2021; Potocki et al., 2020). Entsprechend geht es diesen Studien darum, wie viele oder ob Merkmale der Quellen für Begründungen bemüht wurden (diese letztgenannte Gruppe von Studien sind nicht Gegenstand der Tabelle 11). Damit sind hier Begründungen gezielt auf eine einzige Variable hin verdichtet worden und häufig Teil von übergeordneten Gesamtindizes geworden.

- *Versehen Personen ihre Begründungen mit kritischen Einschätzungen bzw. mit positiven oder negativen Urteilen?* Die Begründungen lassen sich nicht nur nach Kategorien ihrer Bezugspunkte, sondern auch nach ihrer Valenz inhaltlich analysieren, was zusätzliche Einblicke erlaubt. Dass die Entscheidungen bei Beurteilungen mittels offener Fragen erfasst werden, ist weit verbreitet. Ebenfalls ist das Zuweisen zu verschiedenen Arten von Begründungen Usus. Einen empirisch jedoch noch selten beschrittenen Weg hat die Studie von Braasch et al. (2023) eingeschlagen. In ihr wurden die offenen Begründungen für Rangreihungen doppelt spezifisch ausgewertet. Es wurden die Begründungen zunächst danach unterschieden, ob sie sich auf die Dokumente erstreckten, welche den Sachverhalt – die Trainingseffekte auf Arbeitsgedächtnisleistungen – inhaltlich neutral, positiv oder negativ darstellten. Die Begründungen, die auf inhaltsbezogene oder vertrauenswürdigkeitsbezogene Aspekte hin analysiert und zu diesen beiden Gruppen zusammengefasst wurden, erhielten noch eine Valenz-Kodierung. Das heißt, es wurde kodiert, ob die jeweilige Begründung eine positive, negative oder neutrale Ausprägung aufwies. Daraus ließen sich Summenwerte bilden. Eine weitere Studie von Braasch, Bråten, Strømsø und Anmarkrud (2014) ist in diesem Zusammenhang noch erwähnenswert, weil sie ähnlich wie diejenige von Braasch et al. (2023) die Begründungen in den offenen Antworten ausgewertet hat und sich vor allem auf die quellenbasierten Glaubwürdigkeitseinschätzungen konzentriert. Diese wurden pro beurteiltem Dokument danach ausgewertet, ob es kritische Einschätzungen der Glaubwürdigkeit gab, bei denen es für die weiteren Analysen unerheblich war, dass diese eine positive oder negative Valenz aufwiesen. Aus dem Summenwert aller Dokumente wurde ein Index der Glaubwürdigkeitseinschätzungen der Begründungen berechnet. Damit nutzen beide Studien eine Zusatzkodierung, die über die Bezugspunkte der Begründungen hinausgeht und die Valenz der Einschätzungen direkt aufnimmt.
- *Welche Fehlleistungen erbringen Personen bei den Begründungen?* Den bisher berichteten Zugängen wohnt die Logik inne, dass sie sich dem Ausmaß glückender Begründungen widmeten. Einige Studien setzten bei der Auswertung offener Fragen jedoch einen anderen und letztlich komplementären Akzent. Sie nahmen spezifische inhaltsanalytische Auswertungen für Problemlagen vor. Dies erfolgte in Studien, in denen Personen kein vorgegebenes Set an Dokumenten hatten, sondern für ihre Urteile auf das Internet für dezidiert erlaubte Recherchen zurückgreifen konnten (Breakstone et al., 2022; Coiro et al., 2015). Es interessierten dann die Fehlleistungen, die zu Kategorien zusammengefasst wurden. Diese Studien beleuchten also, welche Schwierigkeiten bei der Begründung bestehen. Diese Studien sind besonders bedeutend, um Problemlagen besser benennen zu können, was im Teilkapitel [3.4](#) Gegenstand ist.

Die Auswertung der Begründungen erfolgt also insgesamt unterschiedlich. Neben der Frage, ob es überhaupt zu Begründungen kommt und in welche Kategorie sie gehören, lassen sich diese Begründungen unterschiedlich auswerten, nämlich hinsichtlich der Frage, wie (un-)kritisch bzw. wie adäquat hinsichtlich der Aufgabenstellung sie ausfallen. Somit eröffnen sich diverse inhaltliche Auswertungsstrategien und unterschiedliche Optionen, aus den Begründungen Variablen zu bilden.

Ergebnisse: Welche Muster bei den Begründungen gibt es?

Trotz der verschiedenen Operationalisierungen bei den Begründungen in Zahl und Granularitätsgröße in den jeweiligen Studien lassen sich mehrere studienübergreifende Muster klar benennen:

- *Begründungen für die verschiedenen Urteile zu den Dokumenten erfolgen mehrdimensional und oftmals unabhängig voneinander.* Dieser Befund geht aus deskriptiver Sicht bereits deutlich aus den vielen Studien aus Tabelle 11 hervor. Wenn Personen ihre Entscheidungen begründen, dann tun sie dies in der Regel nicht allein auf der Grundlage von nur inhaltlichen oder ausschließlich metadatenbasierten Kriterien im Sinne einander ausschließender Zugänge. Bereits die Studie von List et al. (2016b), die im Kasten „Illustrierendes Beispiel 11“ vorgestellt wurde, zeigte dieses Muster für epistemische und nicht-epistemische Begründungen in der deskriptiven und bivariaten Statistik. Je nach Granularitätsgröße gibt es Unterschiede, allerdings ist auffällig, dass die Begründungen unterschiedlicher Kategorien eher selten und eher moderat statistisch zusammenhängen (Braasch et al., 2013, 2023; Bråten, Brante & Strømsø, 2018) bzw. sogar negative Zusammenhänge von inhaltlichen und metadatenbasierten Begründungen bestehen (Kessler et al., 2021). Freilich existieren Gegenbeispiele, die mit wenigen und hochaggregierten Summenwerten pro Kategorie sehr hohe und substanzielle Korrelationen haben belegen können, etwa List et al. (2016b) für epistemische und nicht-epistemische Begründungen. Eine andere Perspektive auf die Einzelkriterien erlauben jene Studien, in denen Personen einschätzen sollten, wie wichtig verschiedene Metadaten für ihre Vertrauenswürdigkeitsurteile über Informationen aus den gelesenen Dokumenten waren, und in denen die Ergebnisse über alle Dokumente hinweg in Korrelationsanalysen ausgewertet wurden (Bråten et al., 2009; Pieschl & Sivyer, 2021; Tarchi, 2019, 2021a). Hier hingen diese Einschätzungen teils substanziell zusammen, nämlich vor allem die drei Kriterien Inhalt, Publikationstypus und -organ. Auffällig war zudem, dass die Wichtigkeit der eigenen Meinung in diesen Studien auffällig geringer mit anderen Kriterien korrespondierte und im Licht der Korrelationskoeffizienten eher als eigenständige Dimension wirkt. Wenn sich die Betrachtungsebene allerdings auf einzelne Dokumente verengt,

dann sind die soeben genannten Muster nicht mehr ohne weiteres auffindbar, sondern bestehen teils nur noch bei einzelnen Dokumenten (Bråten, Strømsø & Salmerón, 2011).

- *Es gibt musterhafte Präferenzen für Kriterien bei den Einschätzungen – meta-datenbezogene Einschätzungen sind nur eine und nicht notwendigerweise die wichtigste Grundlage.* Die Studien mit geschlossenen oder offenen Frageformaten divergieren, was nicht nur in der Erfassung (s. Tabelle 9), sondern auch in den Auswertungskategorien deutlich wird (s. Tabelle 11). Darum sind musterhafte Präferenzen allein auf dieser Grundlage nur mit hoher Abstraktion und dem Eingeständnis von hoher Divergenz der Rangreihungsurteile bestimmbar.

In den Studien, in denen der *Inhalt* eine Kategorie war, befand sich diese Kategorie weit vorn in den Rangreihungen, häufig auf dem ersten Rang. Inhaltliche Aspekte spielen – auch teils in Verbindung mit der Verwertbarkeit der Informationen für ein Ziel (Relevanz) – eine vordergründige Rolle. Wo differenzierter die Kategorie des Inhalts kodiert wurde, waren Kriterien mit hohem kognitivem Aufwand wie der Dokumentenvergleich (Corroboration) oder der Abgleich mit anderen Informationen in ihrer Position erheblich geringer eingestuft. Demgegenüber wirken heuristische Einschätzungen zum Inhalt als gangbarer Weg, die Begründungen zu untermauern, z. B. über die Einschätzung der Faktizität von Dokumenten bzw. über die angenommene Korrektheit der Inhalte. Eine Ausnahme in diesem Befundmuster ist die Studie von der Mühlens et al. (2016), in der inhaltliche Aspekte auffällig selten genannt wurden, was sich methodisch erklären lässt, da die Teilnehmerinnen und Teilnehmer nur maximal eine Minute das jeweilige Dokument bearbeiten konnten. Dies war gezielt so geplant worden, um insbesondere heuristische Reaktionen einzufangen, sodass eine inhaltliche Auseinandersetzung wie in den anderen Studien nicht möglich war. Wo die *Relevanz* ein Kriterium war, war sie ein in den Rangreihungen oftmals hoch gewichtetes Kriterium. Die instrumentelle Verwertbarkeit der Dokumente war insbesondere in Studien, in denen die Auswahl bzw. die Verwendbarkeit Bezugspunkte der Begründungen darstellten, das zentrale Kriterium der Begründungen. Das *Sourcing* als Sammelkategorie ist eine Begründung, welche durch ihre breite Streuung in der studienübergreifenden Analyse auffällt. Bei singulären Indizien des Sourcings, teils auch bei der hochaggregierten Gesamtkategorie war das Sourcing sowohl das bedeutendste Kriterium, teils aber auch in Relation stark untergeordnet. Die Person der Autorin bzw. des Autors und die Art des Dokuments bzw. der Kontext, in welchem es erschienen ist, sind die am häufigsten extrahierten Kategorien. Aufgrund der unklaren Datenlage und der heterogenen Operationalisierung (s. Teil A in Tabelle 11) kann zu den Begründungen mit klarem Sourcing-Bezug nur konstatiert werden, dass es unterschiedlich stark in seiner Eminenz für Begründungsgrundlage genutzt wird. Sourcing ist damit eine unter mehreren Optionen, selten jedoch die zentrale.

Tabelle 1.1: Überblick über Begründungen für Urteile in Studien mit Fokus auf evaluative Einschätzungen

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächen-merkmale	Weitere Merkmale
A) Geschlossene Erhebungsmethoden							
Bråten et al. (2009) = Bråten, Strømso und Salmerón (2011)	I/A	6	• Dokumententypus: ② • Publikationsorgan: ③ • Autorin: ⑤	①		Datum: ⑥	eigene Meinung: ④
Strømso et al. (2011) – Mittelwert über beide Dokumente	I/A	4	• Dokumententypus: ② • Autorin: ④	①			eigene Meinung: ③
Tarchi (2019)	I/A	6	• Publikationsorgan: ② • Dokumententypus: ③ • Autorin: ④	①		Datum: ⑥	eigene Meinung: ⑤
Tarchi (2021a)	I/A	6	• Publikationsorgan: ② • Dokumententypus: ③ • Autorin: ④	①		Datum: ⑥	eigene Meinung: ⑤
Von der Mühlen et al. (2016) – Expert:innen	G	6°	• Textgenre: ① • Informationen zu Autor:innen: ② • Titelinformationen: ③ • Qualität der Referenzierungen: ④	Argumentationslinie: ⑤		Anzahl der Referenzierungen: ⑥	

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Von der Mühlen et al. (2016) – Noviz:innen	G	6°	• Textgenre: 1 • Informationen zu Autor:innen: 2 • Titelinformationen: 3 • Qualität der Referenzierungen: 5	Argumentationslinie: 4		Anzahl der Referenzierungen: 6	
B) Offene Erhebungsmethoden							
Abed & Barzilai (2023) – Videos; bei zwei Kategorien identische Werte (9)	R	10	Expertise der Quelle: 8	• Qualität der Erklärung: 2 • Vollständigkeit: 6 • inhaltliche Validität: 9	7	• Nutzungserfahrung: 1 • visuelle Qualität: 3 • Verständlichkeit: 4 • sprachliche Qualität: 5 • Tonqualität: 9	
Abed & Barzilai (2023) – Websites; bei zwei Kategorien identische Werte (4)	R	9	Expertise der Quelle: 7	• Qualität der Erklärung: 2 • Vollständigkeit: 5 • inhaltliche Validität: 8	4	• sprachliche Qualität: 1 • Nutzungserfahrung: 3 • visuelle Qualität: 4 • Verständlichkeit: 6	

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Barzilai, Thomm & Shlomi-Ellooz (2020) – Experiment 1: vertrautes Thema	EM: Autorin	4	Sourcing: ❶	- Diskurs (inhaltliche Analyse): ❸ - Corroborati-on: ❹			Vorwissen: ❷
Barzilai, Thomm & Shlomi-Ellooz (2020) – Experiment 1: unvertrautes Thema	EM: Autorin	4	Sourcing: ❶	- Diskurs (inhaltliche Analyse): ❷ - Corroborati-on: ❸			Vorwissen: ❹
Barzilai, Thomm & Shlomi-Ellooz (2020) – Experiment 2: vertrautes Thema	EM: Autorin	4	Sourcing: ❷	- Diskurs (inhaltliche Analyse): ❸ - Corroborati-on: ❹			Vorwissen: ❶
Barzilai, Thomm & Shlomi-Ellooz (2020) – Experiment 2: unvertrautes Thema	EM: Autorin	4	Sourcing: ❷	- Diskurs (inhaltliche Analyse): ❶ - Corroborati-on: ❸			Vorwissen: ❹
Braasch et al. (2013) – Kontrollgruppe	R	3	❷	❶			keine Begründung: ❸

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Bräten, Brante & Strømsø (2018) – Thema Kernenergie	A	2	1	2			
Bräten, Brante & Strømsø (2018) – Thema Klimawandel	A	2	2	1			
Bräten et al. (2015)	R	5	• Autor:in: 2 • Dokumententypus: 3 • Publikationsorgan: 4	1			eigene Meinung: 5
Breakstone et al. (2022) – Fake-News-Aufgabe; bei zwei Kategorien identische Werte (7)	G	8	• fehlender Autor:in: 1 • keine wissenschaftliche Quelle: 7		2	• Domain-Name: 3 • Typus Website: 4 • Erscheinungsbild: 7	• Generelle Ablehnung von Medien: 5 • Vertrauen in Zeitungen: 6
Breakstone et al. (2022) – Minimum-wage-Aufgabe	G	10	• keine wissenschaftliche Quelle: 4 • Einschätzung der Website als Nonprofit-Organisation: 7 • fehlender Autor:in: 8	Abgleich mit anderen Quellen bei Einzelinformationen: 6	1	• Domain-Name: 2 • Impressum: 3 • Links als Hinweis: 5 • Aktualität: 9 • Erscheinungsbild: 10	

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Delgado et al. (2019) – Pretest	V	3	③	①			Vorwissen: ②
Kessler et al. (2021)	R	3	allgemeine Quellenmerkmale: ③	- Anzahl korrekter Konzepte: ① - Anzahl inkorrektter Konzepte: ②			
Kiili et al. (2019), Studie 1	G	2	②	①			
Kiili et al. (2019), Studie 2	G	2	②	①			
Kiili, Leu, Marttunen et al. (2018) – akademische Website	I/A	3	Expertise Autor:in: ①	②			irrelevante Begründungen: ③
Kiili, Leu, Marttunen et al. (2018) – kommerzielle Website	I/A	3	Expertise Autor:in: ③	①			irrelevante Begründungen: ②

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
List & Alexander (2018b)	R	7	• Ziele Autorin: 2 • Autorin: 4 • Merkmale Autorin: 5	• Einschätzung Faktizität: 1 • Vollständigkeit der Informationen: 3 • Corroborations: 6		Aktualität: 7	
List et al. (2016b) – bei zwei Kategorien identische Werte (6)	A	10	• Einfachheit der Nutzung: 3 • tiefenorientiert quellenbasierte Faktoren: 4 • individuelle Einschätzung der Quelle: 9	• informationsgetrieben: 1 • inhaltliche Faktoren: 7 • Corroborations: 8	6	Oberflächenmerkmale der Quellen: 2	• Zufall: 5 • Verknüpfung mit anderen Informationen außerhalb des Dokuments: 6
Lucassen & Schraagen (2011) – Expert:innen	I/A	4	3			2	• eigenes Wissen: 1 • Anderes: 4
Lucassen & Schraagen (2011) – Noviz:innen	I/A	4	2			1	• anderes: 3 • eigenes Wissen: 4
Lyu et al. (2023)	A	5	epistemische Kriterien: 4	1	Aufgabenstellung: 5	Formatierung, Verständlichkeit: 2	anderes: 3

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/ Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro- kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz- kriterien	Oberflächen- merkmale	Weitere Merkmale
Macedo-Rouet et al. (2013, Experiment 1)	EM: Expertise	3	1	2			kein Bezug: 3
Macedo-Rouet, Potocki et al. (2019, Experiment 1)	V	8	Kompetenz Autor:in: 7	Informationsmenge: 3	1	- Lesbarkeit: 2 - Aktualität: 3	- Zufriedenheit mit Text: 4 - anderes: 5 - eigene Meinung: 6
Macedo-Rouet, Potocki et al. (2019, Experiment 2) – bei einer Kategorie identische Werte (2)	V	8	Kompetenz Autor:in: 4	Informationsmenge: 3	1	- Lesbarkeit: 2 - Aktualität: 6	- Zufriedenheit mit Text: 2 - anderes: 5 - eigene Meinung: 7
Mason, Scrimin, Tornatora et al. (2018)	R	4	2	- Inhalt: 1 - Corroboration: 4			eigene Meinung: 5
Rouet et al. (1996) – Gruppe Primärdokumente	R	4	- Dokumententypus: 2 - Autor:in: 3	1			eigene Meinung: 4
Rouet et al. (1996) – Gruppe Sekundärdokumente	R	4	- Autor:in: 2 - Dokumententypus: 3	1			eigene Meinung: 4
Rouet et al. (1997) – Expert:innen	R	4	1	2	3		anderes: 4

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Rouet et al. (1997) – Noviz:innen	R	4	②	①	④		anderes: ⑤
Salmerón et al. (2016, Experiment 1)	I/A	3	②	①			Kombination Inhalt und Quelle: ③
Salmerón et al. (2019)	V	3	③	②			kein Bezug: ①
Stenseth et al. (2019) – Thema Kernenergie; bei drei Kategorien identische Werte (② [2x], ③ [3x] und ⑤ [2x])	A	9	- Person Autor:in: ① - Glaubwürdigkeit Autor:in: ③ - Motive Autor:in: ⑤ - Zusammenhang von Expertise und Verständigkeit: ⑤		- Ziel basale Informationen: ② - Ziel multiple Perspektiven: ② - Ziel Diskussion anregen: ③		- Vorwissen (fehlendes Wissen): ③ - Vorwissen (vorhandenes Wissen): ④
Stenseth et al. (2019) – Thema Klimawandel; bei zwei Kategorien identische Werte (⑥ und ⑦)	A	9	- Glaubwürdigkeit Autor:in: ② - Person Autor:in: ③ - Zusammenhang von Expertise und Verständigkeit: ⑥ - Motive Autor:in: ⑦		- Ziel basale Informationen: ① - Ziel multiple Perspektiven: ⑤ - Ziel Diskussion anregen: ⑥		- Vorwissen (vorhandenes Wissen): ④ - Vorwissen (fehlendes Wissen): ⑦

Studie	Auswahl (A) Reihung (R), Globalurteil (G) ohne Reihung, Informationen bzw. Aussagen (I/A), Verwendung/Verwendbarkeit (V), einzelne Metadaten (EM)	Anzahl Kriterien (Makro-kategorien)	Quellenmerkmale	Inhaltliche Merkmale	Relevanz-kriterien	Oberflächenmerkmale	Weitere Merkmale
Strømso et al. (2017) – Thema Aspartam	I/A	4	Eingebettete und Dokumentenquellen: ④	③			• persönliche Meinung: ① • andere Quellen: ②
Strømso et al. (2017) – Thema Mobilfunk	I/A	4	Eingebettete und Dokumentenquellen: ④	②			• persönliche Meinung: ① • andere Quellen: ③
Strømso et al. (2020)	A	2	①	②			

(Legende: Dunkel hervorgehobene Zahlen in den Zellen geben an, auf welchem Rang sich die jeweilige Kategorie gemäß Häufigkeitsverteilung/den Mittelwerten bei offenen Antworten bzw. beim Mittelwert in geschlossenen Antworten in der jeweiligen Studie befand. Legende: ° Kriterien a) Stil, b) Struktur, Präsenz von c) Tabellen und d) Abbildungen, e) Layout, f) Objektivität und g) eigene Meinung wurden in der Studie zwar erfasst, aber wegen des Fokus' auf signifikante Unterschiede dort nicht berichtet)

- *Die Begründungen fallen für einzelne Dokumente spezifisch aus.* Die Begründungen lassen sich nicht nur danach differenzieren, welche Kriterien Personen mehrdimensional anwenden und welche Schwerpunkte sie – jeweils auf der Aggregatebene der Gruppen von Studiensamples – allgemein setzen. Vielmehr ist es auch möglich, Eigenheiten bei den Begründungen auf der Ebene einzelner Dokumente zu eruieren.

Ähnlich wie bei den holistischen Einschätzungen ganzer Dokumente gemäß Punkt 1a weiter oben gibt es studienübergreifende Muster. Bei offiziellen und neutralen Dokumenten erfolgten mehr inhaltsbezogene Begründungen, bei eher subjektiv gefärbten Dokumenten eher metadatenbezogene (Bråten et al., 2015; Rouet et al., 1996, 1997). Wo die metadatenbezogenen Analysen in den Begründungen genauer erfolgten, zeigten sich Tendenzen von Profilen bei den Dokumenten(gruppen). Neutrale bzw. offizielle Quellen im Sinne von epistemischen Autoritäten (wie Lehrbücher oder Aussagen von Forschungsinstitutionen) werden über das Publikationsorgan bzw. die Autorinnen und Autoren als integer wirkend begründet, während inhaltlich einseitig darstellende Quellen (wie Zeitschriften oder Blogs) wegen des Publikationsorgans und/oder bei Zweifeln an der Autorin/dem Autor eher kritisch betrachtet werden (Bråten et al., 2015; Bråten, Strømsø & Salmerón, 2011; Strømsø et al., 2011) bzw. mit der eigenen Meinung abgeglichen werden (Mason, Moë et al., 2022, Kontrollgruppe; Mason, Scrimin, Tornatora et al., 2018).

Hinzu kommt, dass Personen bei inhaltlichen Einschätzungen besser dazu in der Lage sind, Dokumente als (in)adäquat einzuschätzen, als sie es bei der Expertise der Autorinnen und Autoren bzw. der Aktualität von Aussagen zu leisten vermögen. Dies ist das Ergebnis zweier Experimente von Macedo-Rouet, Potocki et al. (2019), in denen vier Dokumente so gestaltet wurden, dass sie einem Kriterium nicht genügten, z. B. der Verfasser keine erkennbare Qualifikation zum Thema hatte bzw. ein Dokument deutlich älter war als alle anderen im Set. Geprüft wurde, ob die Teilnehmenden aus den Klassenstufen 5 bis 10 in den Begründungen dasjenige Kriterium passend zu dem Text anführten, nach welchem flankierende Fragen explizit fragten. Dies war bei der Expertise und Aktualität in nur einem Bruchteil des ersten Experiments (unter 4 Prozent), und maximal für 36 Prozent der Begründungen im zweiten Experiment der Fall. Zusammengefasst: Personen weisen zum Teil die Fähigkeit auf, ihre Begründungen an die Dokumente anzupassen. Die gleichen Kriterien (hier verstanden als Kategoriebezeichnungen) nutzen Personen hierbei unterschiedlich, sei es in Bezug auf die Quellenmerkmale, sei es in der Valenz (s. dazu weiter unten). Jedoch fällt es Personen nicht immer leicht, ungeeignete Dokumente anhand explizit erfragter Kriterien ausreichend treffsicher einzuschätzen.

- *Die Begründungen sind nicht kontextlos, sondern erfolgen in Verbindung mit den Merkmalen der Aufgaben.* Einzelne Studien prüften (List et al., 2016b) bzw. berichteten (Breakstone et al., 2022; Kiili, Leu, Marttunen et al., 2018),

wie sich Begründungen je nach *Aufgabenstellung* und zur *Verfügung* gestellten Dokumenten unterschieden. Trotz der dünnen Datenbasis zeichnen sich erste abstrahierbare Muster ab, dass die Personen anhand der Aufgabenerfordernisse anders vorgehen: Sie prüften weniger, wenn der Aufgabentypus einfachere Antworten verlangte (List et al., 2016b) bzw. stützten sich in den Begründungen mehr (und teils oberflächlich) auf die salienten Metadaten bzw. deren Fehlen (Breakstone et al., 2022). Außerdem gaben sie bei Quellen mit erkennbarer Expertise mehr metadatenbezogene und bei Quellen mit kommerziellen Interessen mehr inhaltliche Begründungen (Kiili, Leu, Marttunen et al., 2018). Eine andere Variante der Variation von Aufgaben besteht darin, dass die *Themen* in Hinblick auf die unterstellte Vertrautheit bzw. Unvertrautheit der Testpersonen hin verändert und dann die Effekte dieser Aufgabenstellungen geprüft wurden (Barzilai, Thomm & Shlomi-Elooz, 2020, Studien 1 und 2; Bråten, Brante & Strømsø, 2018; Stenseth et al., 2019). In diesen Studien offenbarte sich das Muster, dass die Personen bei den vertrauten Themen eher vorwissensbasiert begründeten und bei unvertrauten Themen stärker den Inhalt und die Metadaten anführten (Barzilai, Thomm & Shlomi-Elooz, 2020, Studien 1 und 2). Wenn nur die Kategorien Inhalt und Metadaten fokussiert wurden (Bråten, Brante & Strømsø, 2018), begründeten die Personen in vertrauten Themen mehr über den Inhalt, bei den unvertrauten Themen hingegen über die Metadaten. In einer ähnlichen Tendenz findet sich dieses Ergebnis selbst dann noch wieder, wenn erheblich mehr Kriterien ausgewertet wurden (Stenseth et al., 2019). Vertrautere Themen elizitieren anscheinend stärker inhaltsbezogene Prüfungen – ein Ergebnis, das sich in den unterschiedlichen Vorgehensweisen von Personen mit unterschiedlichem Vorwissen ebenfalls widerspiegelt (s. dazu weiter unten die personenbezogenen Auswertungen).

- *Die Globalurteile und die metadatenbezogenen Begründungen hängen in Studien mit geschlossenen Antwortformaten nur zum Teil zusammen.* Bessere Sourcingleistungen bei dem Erkennen der vertrauenswürdigeren Dokumente korrelierten einerseits mit der größeren Zuwendung zu den Kriterien Inhalt (Braasch et al., 2009, nur in der Gruppe der Leistungsstärkeren) bzw. generell dem Anführen von normativ gewünschten Quellenmerkmalen in den Begründungen (Wiley et al., 2009, Experiment 1). Bemerkenswert ist aber klarerweise überdies, dass andererseits in Studien häufig keinerlei Zusammenhänge mit den anderen erfassten Begründungsmöglichkeiten bestanden (Braasch et al., 2009) bzw. es überhaupt keine statistisch auffälligen Korrelationen (Kessler et al., 2021) resp. regressionsanalytischen Zusammenhänge (Abed & Barzilai, 2023) zwischen den Begründungen und Rangreihungsurteilen gab.
- *Rangreihungsbegründungen, die spezifisch nach Valenz kodiert wurden, verdeutlichen, dass evaluative Aussagen musterhaft mit anderen Verstehensleistungen sowie untereinander korrelieren.* Im Textteil oben zu den Spezifika

bei den Kodierungen wurden die variablenzentrierten Valenzkodierungen thematisiert, also die zusätzliche Bestimmung, ob Reihungsbegründungen mit neutralen, positiven, negativen bzw. kritischen Ausprägungen erfolgt waren. Dieses bisher nur zuweilen vorgenommene Vorgehen (Braasch et al., 2023; Braasch, Bräten, Strømsø & Anmarkrud, 2014) liefert Erkenntnisse zum Verhältnis von Außenmaßen des Verstehens und den Begründungen. Die erste Studie (Braasch, Bräten, Strømsø & Anmarkrud, 2014) zählte aus, wie viele der Rangreihungsbegründungen Indizien für eine kritische Einschätzung beinhalteten. Der Gesamtwert pro Person korrelierte zum einen substanziell mit der Fähigkeit, vertrauenswürdige von weniger vertrauenswürdigen Dokumenten zu unterscheiden, und zum anderen mit der Anzahl zentraler Konzepte in eigenen Diskurssynthesen. Das bedeutet: Kritische Einschätzungen korrespondieren mit anspruchsvollen Verstehensleistungen des Evaluierens, aber auch des Verwendens intertextueller Makropropositionen. Die bisher ausführlichste und kleinteiligste Analyse mittels Korrelationen stammt aus der zweiten Studie (Braasch et al., 2023) und eröffnet vertiefte Einsichten, wenn man zum einen nur auf die Gruppen von Dokumenten fokussiert, die den Sachverhalt (Beeinflussbarkeit des Arbeitsgedächtnisses durch Trainings) positiv, neutral oder negativ darstellen. Zum anderen lässt sich prüfen, ob die Zusammenhänge konsistent über die Gruppen der Dokumente bestehen. Zunächst zu den Befunden, welche die Korrelationen *innerhalb einzelner Dokumentengruppen* (positive, neutrale oder negative inhaltliche Perspektive) betreffen: Bei den Dokumenten mit positiver und neutraler Perspektive korrelieren die positiven resp. negativen inhaltlichen Rangreihungsurteile negativ, d. h. wer mehr positiv über den Inhalt sein oder ihr Urteil begründet, tut es weniger mit negativ getönten Begründungen. Bei dem Dokument mit negativer Perspektive gab es dieses Muster in schwächerer Form ebenfalls, wenn auch nicht statistisch überzufällig. Außerdem bestanden Auffälligkeiten bei den Begründungen mit der Vertrauenswürdigkeit insofern, als bei den beiden Dokumentengruppen mit positiver bzw. negativer inhaltlicher Perspektive die Zusammenhänge jeweils anders ausfielen: Im Falle eines Dokuments (mit positiver inhaltlicher Perspektive) waren die positiven und negativen Vertrauenswürdigkeitsbeurteilungen unkorreliert, beim Dokument mit negativer inhaltlicher Perspektive hingegen positiv korreliert. Das heißt: Je nach Dokumentengruppe (positiv, neutral, negativ) und je nach Polung der Valenz (positiv bzw. negativ) und je nach Bezugspunkt der Begründung (Inhalt bzw. Vertrauenswürdigkeit) sind allein auf der Ebene der Dokumentengruppe differenzielle Zusammenhänge beobachtbar. Daneben zeigten sich *über die Dokumentengruppen hinweg Auffälligkeiten*. Rangreihungsbegründungen über den Inhalt erfolgen in der Regel homogen. Inhaltsbezogene Urteilsbegründungen in einer Richtung ziehen sich über die Dokumentengruppen hinweg. Wer also das Rangreihungsurteil innerhalb

der positiven getönten Dokumente mit positiven Urteilen versieht, nutzt eher negativ gefärbte Inhalte, die negativ darstellenden Dokumente in ihrer Glaubwürdigkeit zu beurteilen. Dieses Muster findet sich zudem bei den Begründungen zur Vertrauenswürdigkeit wieder, aber nicht so deutlich ausgeprägt über alle Dokumentengruppen hinweg. All diese Befunde verdeutlichen, dass der genaue Blick auf das Zusammenspiel verschiedener Variablen sich lohnt. Allerdings wurden in der Studie keine weiteren Analysen mit zusätzlichen Leistungsdaten vorgelegt, sodass unklar bleibt, ob verschiedene Valenzen bei den Begründungen differenzielle Leistungseffekte bedingen.

- Nicht nur die bisher dominant präsentierten variablenzentrierten Auswertungen sind in den Studien erfolgt, sondern es wurden auch *personenbezogene Analysen* im Sinne von Gruppenvergleichen von Testpersonen vorgenommen. Die Analysen in den Studien setzten entweder darauf, a priori definierte Gruppen wie Altersgruppen (Pieschl & Sivyer, 2021; Potocki et al., 2020; Salmerón et al., 2016), Gruppen von Personen mit hoher fachlicher Expertise sowie solche mit Laien-Status (Lucassen & Schraagen, 2011; Rouet et al., 1997; von der Mühlen et al., 2016) bzw. Vorwissensdifferenzen (Rouet et al., 1997) zu vergleichen, oder aber es gibt Gruppen, die erst durch die Experimente selbst entstehen – sei es mit divergierenden Wissensbeständen (Bråten, Strømsø & Salmerón, 2011), sei es als Leistungsgruppen in der korrekten Sourcing- bzw. inhaltlichen Leistung (Braasch et al., 2009; Wiley et al., 2009, Experiment 1), sei es über Clusteranalysen (Denton et al., 2022, Studie 2, s. Kasten „Illustrierendes Beispiel 14“).

Je älter Personen sind, desto besser werden sie (tendenziell) darin, Metadaten in ihre Begründungen einzubinden. Die Studienlage ist zwar noch ausbaufähig, doch das genannte Muster offenbart sich an verschiedenen Stellen: Jüngere Kinder zeigen sich weniger sensitiv für die Inhalte von Dokumenten und begründen vor allem über ihre eigene Meinung (Salmerón et al., 2016, Experimente 1 und 2) und erheblich weniger über die Verbindung von Inhalt und dessen Quelle (Salmerón et al., 2016, Experiment 2). Je älter die Testpersonen (hier: höhere Klassenstufen in der Sekundarstufe) waren, desto deutlicher nutzten sie das Kriterium des Autors bzw. der Autorin dafür, zwischen vertrauenswürdigen und weniger vertrauenswürdigen Dokumenten zu unterscheiden (Pieschl & Sivyer, 2021), wobei die Fähigkeit zur Differenzierung von erkannter Expertise mit steigendem Alter größer ist (Macedo-Rouet, Potocki et al., 2019, Experiment 2; Potocki et al., 2020). Ein weiterer Alterseffekt besteht darin, dass Personen sich mit zunehmendem Alter darin unterscheiden, mehr quellenbasierte Begründungen und weniger inhaltliche Kriterien anzuführen, um ihre Glaubwürdigkeitsurteile zu untermauern (Potocki et al., 2020).

Die Vergleiche zwischen *Laien und Expertinnen/Experten* sind mehrfach untersucht worden. Lucassen und Schraagen (2011) ermittelten beispielsweise, dass sich Laien und Expertinnen bzw. Experten darin unterschieden,

dass die letztgenannte Gruppe mehr Inhalte prüfte, während die Laienpersonen stärker auf Oberflächenmerkmale und das Sourcing auswichen. Teils sind diese Differenzen zwischen Expertinnen und Experten sowie Laienpersonen aber nur geringfügig (Rouet et al., 1997), selbst wenn sich in Tabelle 11 die Rangreihungen unterschieden. Die Rangreihungsergebnisse dürfen zudem nicht darüber hinwegtäuschen, dass die Kriterien teils erheblich wichtiger für die Expertinnen und Experten waren als für die Novizinnen und Novizen, was ein markantes Resultat aus der Studie von der Mühlens et al. (2016) war. Es bestehen also durchaus Unterschiede zwischen den Personengruppen mit hoher und geringer Expertise, doch sind diese mitunter weniger evident, als man vermuten könnte. Personen mit unterschiedlichem *Vorwissen* sind eine weitere Differenzkategorie in der Empirie. Diese Personengruppen divergieren partiell in der Wichtigkeit, die sie einzelnen Kriterien zugewiesen haben. Dies geht aus der bereits bei den Glaubwürdigkeitsurteilen erwähnten Studie von Bråten, Strømsø und Salmerón (2011) hervor. Laut dieser Studie differenzierten Personen mit mehr Vorwissen (statistisch signifikant) stärker bei der Wichtigkeit der Kriterien, indem sie Publikationsorgan und Dokumententypus sowie ihre eigene Meinung als teils wichtiger einschätzten als andere Kriterien. In der Gruppe jener Personen mit geringerem Wissen waren die Unterschiede zwischen den Kriterien nivelliert. Rouet et al. (1997) stellten ebenfalls in deskriptiven Analysen fest, dass Personen mit einem ersten Studienabschluss im Fach Geschichte ihre Begründungen stärker über den Dokumententypus absicherten. Es wirkt demnach nicht so, als würden die Differenzen Vorwissensgruppen in jeglicher Hinsicht scharf trennen; vielmehr bestehen Unterschiede an einzelnen Punkten. Die Eigenheiten von *Leistungsgruppen* sind bisher nur selten untersucht worden. In einer Studie von Braasch et al. (2009) wurden Jugendliche gebeten, Websites einzuschätzen. Je nachdem, welche Leistungen sie erbrachten, unterschieden sich die Personen in ihrer Nutzung von Kriterien. Gut zwischen den Dokumenten differenzierende Personen nutzten vor allem inhaltliche Kriterien wie eine kurze Zusammenfassung bzw. den Titel. Schwächer differenzierende Personen stützten ihre Urteile hingegen mehr auf Metadaten. Anders gesagt: Sie nutzten zwar das Sourcing, aber dies ging nicht mit einer adäquaten Evaluation der Dokumente einher. Das Sourcing war zwar bedeutend für ihre Leistungen, doch diese Leistung war inadäquat. Wiley et al. (2009, Experiment 1) ermittelten unter Studierenden ebenfalls Unterschiede zwischen Leistungsgruppen sowohl in den korrekt als zuverlässig als auch als unzuverlässig erkannten Websites sowie im Index, der die Differenz aus beiden Ergebnissen bündelte. Statistisch signifikante Differenzen fanden sich lediglich in zwei der acht ausgewerteten Kategorien für die Begründungen der Zuweisung der Websites zu einer Rangreihungsposition. Die erfolgreicheren Leserinnen und Leser gingen häufiger auf die Erklärungsgüte

des Inhalts zum einen und die Metadaten zum anderen ein. Die weniger erfolgreichen Leserinnen und Leser fokussierten deutlich mehr auf die Relevanz der Website, selbst wenn dieser Befund nicht statistisch signifikant war.

Ein letzter Zugang im Personenvergleich basiert auf ähnlichkeitsbasierten Analysen von Personen anhand ihrer Begründungen. Diese Variante ist ausführlicher Gegenstand im nachstehenden Kasten „Illustrierendes Beispiel 14“. Dieser methodische Zugang erlaubt eine weitere Differenzierung, in der die Begründungen die Basis der Gruppenbildung sind. Dieses Beispiel macht deutlich, dass sich in den Studien latente Gruppen mit eigenen Profilen bei den Begründungen verbergen.

Illustrierendes Beispiel 14

Lassen sich Personen danach unterscheiden, auf welche Art sie Begründungen geben?

A) Was war das Erkenntnisinteresse?

Dem Zusammenhang von korrekten Glaubwürdigkeitsurteilen und den Kriterien bei der Begründung der Globalurteile ging eine Experimentalstudie aus Kanada nach (Denton et al., 2022, Studie 2). Insbesondere interessierte sich die Studie dafür, ob sich Personen darin systematisch unterschieden, welche Kriterien sie nutzen.

B) Wer wurde untersucht?

An der Studie nahmen 43 Personen teil, die überwiegend in ihrem zweiten Jahr an einer kanadischen Universität sozialwissenschaftliche Fächer studierten. Von ihnen waren 65 Prozent Frauen, das mittlere Alter lag bei knapp 19 Jahren. Mit 57 Prozent dominierten Personen europäischer Abstammung das Sample, gefolgt von asiatischen (16 Prozent), afrikanischen (13 Prozent) und weiteren Gruppen.

C) Welche Instrumente kamen zum Einsatz?

In der Studie wurden zwei Zeitungsartikel zum Thema medizinisch assistierter Suizid als Stimuli verwendet. Diese beiden Texte wurden von den Forschenden kriterienbasiert beurteilt und dadurch der vertrauenswürdigeren von den beiden bestimmt. Die erste Aufgabe bestand mithin für die Testpersonen darin, eine der Musterlösung entsprechende korrekte Zuordnung vorzunehmen.

Ihre Urteile begründeten die Studierenden in einer freien Antwort. Diese Begründungen wurden inhaltsanalytisch kodiert, und die Antworten zu Makrokategorien verdichtet. Von diesen gab es sechs: 1) die Aktualität, 2) die Relevanz, 3) die Expertise, 4) die Akkuratheit, 5) die Ausgewogenheit und 6) weitere Kriterien. Für jede Person wurde ausgezählt, wie viele Begründungen zu jeder Makrokategorie gehörten, und auf dieser Basis wurde eine Clusteranalyse durchgeführt. Deren Ziel bestand darin, Gruppen einander ähnelnder Personen zu finden. Die Cluster wurden daraufhin statistisch überprüft, ob sie sich in den Globalurteilen unterschieden.

D) Was waren die Hauptbefunde?

Bezogen auf die Korrektheit der Lösungen waren sieben von zehn Personen dazu in der Lage, die korrekte Lösung zu geben. Die Personen mit korrekter Lösung gaben 5,6, jene mit inkorrektur Lösung 4,5 Begründungen an. Die Ergebnisse der Clusteranalysen sind in Abbildung 8 dargestellt und zeigen die Profile der vier Gruppen. Demnach gab es vier unterschiedlich große Cluster, von denen sich drei (Nr. 1, 3 und 4) darin unterschieden, dass sie jeweils bei einer Hauptkategorie besonders häufige Nennungen aufwiesen, die sich auch in der Clusterbezeichnung niederschlugen. Lediglich das Cluster 2 gab mit den Akkuratheits- und Ausgewogenheitsurteilen zwei Kriterien in ähnlichem Ausmaß an. Die vier Cluster unterschieden sich nicht statistisch in der Korrektheit der Lösung, wenngleich Cluster 3 die besten Werte aufwies, gefolgt von Cluster 1, 2 und 4.

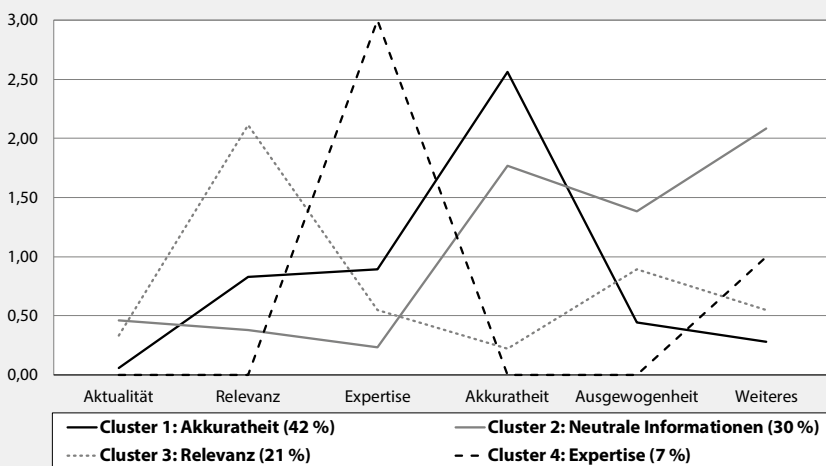


Abbildung 8: Häufigkeitsbasierte Ausprägung (Mittelwerte) verschiedener offen erfasster Beurteilungskriterien bei Glaubwürdigkeitsurteilen in vier Clustern (Quelle: eigene Darstellung, basierend auf Denton et al., 2022, S. 342)

E) Diskussion

Ein wichtiges Ergebnis, welches zugleich ein Desiderat der Sourcingforschung markiert, ist die Profilierbarkeit von Personen, die sich in ihren Begründungen, nicht aber in der auf einer sehr geringen Anzahl von Dokumenten basierenden Globaleinschätzungen zur Glaubwürdigkeit unterscheiden. In ihren offenen Antworten manifestieren sich Schwerpunkte, die sich auf inhaltliche Korrektheit (die größte Gruppe, Cluster 1), die Relevanz von Informationen (Cluster 3) oder die Expertise der Quellen (Cluster 4) beziehen. Zudem sind Kombinationen von anderen Begründungen musterhaft erkennbar (Cluster 2). Dieser personenzentrierte Ansatz, der bei einer kleinen Stichprobe keine statistisch bedeutsamen, aber inhaltlich anregende Befunde ergab, erscheint für weitere Forschungsbemühungen fruchtbar – zumal unter Berücksichtigung weiterer Variablen und mit ökologisch validen Settings wie echten Internetdokumenten.

2) *Beurteilungen und Begründungen der Einschätzung von einzelnen Informationen und Aussagen*

Beurteilungen

Im Vergleich zu Dokumenten als Bezugspunkt der Einschätzungen sind einzelne Informationen bzw. Aussagen deutlich in der Forschung unterrepräsentiert. Dabei ist die *Beurteilung der Informationen bzw. Aussagen* grundsätzlich sehr ähnlich operationalisiert worden wie bei der Einschätzung ganzer Dokumente. In den Studien (Bråten, Salmerón & Strømsø, 2016; Bråten et al., 2009; Kammerer & Gerjets, 2014a; Strømsø & Bråten, 2017; Strømsø et al., 2011, 2017; Tarchi, 2019, 2021a) sollten die teilnehmenden Studierenden auf einer Skala die Frage beantworten, in welchem Ausmaß sie den Informationen bzw. Schlussfolgerungen in den einzelnen Dokumenten vertrauen. Die Ergebnisse aus diesen Studien wurden bereits im Abschnitt 1a) weiter oben berichtet, um Redundanzen in der Ergebnisdarstellung zu vermeiden, zumal die Studienbefunde zur Evaluation einzelner Informationen sich relativ nahtlos einfügen lassen in die Resultate der Beurteilung ganzer Dokumente.

Andere Studien wählten einen vereinfachten Zugang zu der Auswertung, indem sie mit Ja-/Nein-Antworten die Glaubwürdigkeit der Informationen auf Websites einschätzen ließen. In diesen Studien lagen die Quoten der richtigen Antworten sehr weit auseinander. Bei einer Studie mit Siebtklässlern betrug sie ein Viertel (Coiro et al., 2015). Kiili, Leu, Marttunen et al. (2018) differenzierten diesen Befund mit 12- und 13-Jährigen aus, indem sie als Ergebnis festhielten, dass die Einschätzung der kommerziellen Website in ihrer Studie zu knapp über 50 Prozent korrekt erfolgte, während sie bei der wissenschaftlichen Website bei über 90 Prozent lag. Es ist mithin herausfordernd für die Jugendlichen, kritisch mit kommerziellen Darstellungen umzugehen.

Erwähnenswert ist noch abschließend ein methodischer Zugang, den Graesser et al. (2007) in ihrer Interventionsstudie gewählt haben. Sie baten die Teilnehmenden, für die gelesenen Websites die Zuverlässigkeit der Informationen auf einer Skala zu beurteilen. Verglichen wurden dann zwei Gruppen von je drei (un-)zuverlässigen Websites. In beiden Experimenten schätzten die Testpersonen unabhängig von der Zugehörigkeit zur Experimental- und Kontrollgruppe die vertrauenswürdigeren Dokumente korrekter ein.

Begründungen

Wendet man sich den Begründungen für die Einschätzungen zu, lassen sich mehr und unterschiedliche Zugänge in der Methodik attestieren. Basierend auf den Einschätzungen der *Wichtigkeit einzelner Kriterien für das Urteil über*

die *Vertrauenswürdigkeit von Informationen* wurden die Urteile mit mehreren vorgegebenen Antworten, also in geschlossenen Tests, auf unterschiedliche Art behandelt:

- Der erste Zugang bestand darin, bei einzelnen Dokumenten die Begründungen auszuwerten, was jedoch bisher selten vorgekommen ist (List et al., 2022; Strømsø et al., 2011).
- Die zweite Variante ist ähnlich, allerdings werden hier über alle Dokumente hinweg die Mittelwerte berechnet, also die mittlere Wichtigkeit eines Kriteriums für alle gelesenen und beurteilten Dokumente ermittelt (Bråten et al., 2009; Tarchi, 2019, 2021a).
- Die dritte Variante differenziert den zweiten Zugang noch einmal aus, indem die Wichtigkeitsurteile für Gruppen von Dokumenten gemittelt und voneinander subtrahiert werden. Dass bedeutet, dass ein Koeffizient entsteht, der die Differenz der Mittelwerte über die Kriterien bei den vertrauenswürdigen bzw. nicht vertrauenswürdigen Websites bezeichnet und darum von der Fähigkeit zeugt, Kriterien differenziert anzuwenden (Pieschl & Sivyer, 2021). Dieser Wert geht also inhaltlich in die gleiche Richtung wie Diskriminanzwerte bei Globalurteilen ganzer Dokumente (s. o., Nr. 1a in der Aufzählung in diesem Teilkapitel).
- Eine vierte Variante wurde von Macedo-Rouet, Salmerón et al. (2019) vorgeschlagen. Sie ließen für die Auswahl von Links mit geschlossenen Items auf einer Skala einschätzen, ob die jeweiligen Metadaten einen positiven oder einen negativen Einfluss auf ihr Urteil hatten. Faktorenanalytisch ließen sich zwei Konstrukte ermitteln: Einfluss der Expertise und Einfluss von verzerrten Interessen der Autorinnen und Autoren. Diese Variante bildet die theoretisch angenommenen Glaubwürdigkeitsdimensionen Expertise und Vertrauenswürdigkeit ab (s. Teilkap. [3.1.3.2](#)).

Die Ergebnisse der Studien sind in ihrem Gesamt aufgrund der Heterogenität nicht immer eindeutig, dennoch kann man folgende zwei Muster benennen:

- *Die Wichtigkeit einzelner Kriterien in Verbindung für die Einschätzung der Glaubwürdigkeitsurteile ganzer Dokumente wirkt gemäß Studien mit geschlossenen Antwortformaten unklar, jedoch sind autor:innenbezogene Metadaten tendenziell bedeutsam.* Dahinter steckt das empirische Muster, dass die Befunde je nach Studie mehr oder minder konsistent ausfallen. Pieschl und Sivyer (2021) konnten beispielsweise nur eruieren, dass die Differenzierung von in ihren Aussagen glaubwürdigeren und unglaubwürdigeren Dokumenten nur mit zwei Kriterien korrelierte: dem Inhalt und der Person (Autorin/Autor). Bei den restlichen Kriterien bestanden keine Zusammenhänge mit den Leistungen. Strømsø et al. (2011) berichteten in ihrer Studie mit einer Analyse

von zwei Dokumenten, dass nur bei einem der Dokumente die Bedeutsamkeit der Autorin bzw. des Autors positiv mit dem Glaubwürdigkeitsurteil über einen populärwissenschaftlichen Text korrelierte, während es mit demselben Kriterium bei dem zweiten Dokument, einem Zeitungsartikel, einen leicht negativen, aber statistisch nicht-signifikanten Zusammenhang gab.

Andere Studien mittelten die Bedeutsamkeitseinschätzungen der Kriterien über alle zu beurteilenden Dokumente, bildeten also eine Skala, die sie dann mit dem Glaubwürdigkeitsurteil über die Dokument-Aussagen statistisch ins Verhältnis zueinander setzten (Bräten et al., 2009; Tarchi, 2019, 2021a). In diesen Studien zeigten sich tendenziell Muster: Die Person der Autorin bzw. des Autors, der Publikationstypus und das Publikationsorgan erweisen sich erstens studienübergreifend als positiv korreliert mit der Einschätzung der Glaubwürdigkeit vieler Dokumente in den Dokumentensets. Zweitens sind neutral informierende Dokumente und die Wichtigkeit diverser Kriterien häufiger und durchgängiger positiv korreliert, als dies bei stark einseitigen und in ihrer Gesamtbeurteilung eher unzuverlässig wirkenden Dokumenten der Fall ist. Tatsächlich sind die Gesamturteile bei diesen Dokumenten auffällig seltener bis gar nicht mit den vorgegebenen Kriterien und deren Einschätzung assoziiert.

Die Befundlage ist damit gemischt. Allenfalls die Person bzw. Institution hinter den Aussagen wirken aktuell als ein Kristallisationspunkt, dessen Wichtigkeit für höhere Glaubwürdigkeitseinschätzungen im Nachgang nach der Einschätzung teils konsistent in den Studien vorfindbar war.

- *Personen unterscheiden verschiedene Arten von Evidenzen von Informationen, haben aber partiell Probleme dabei, ausreichend akkurate bzw. differenzierte Einschätzungen und Begründungen zu geben.* Ein noch wenig erforschtes Terrain hat die experimentelle Studie von List et al. (2022) fokussiert. In ihr wurden Studierende in einem Within-Design gebeten, verschiedene kurze Aussagen (fingierte Zeitungsmeldungen) gemäß vier Evidenztypen zu klassifizieren und in ihrer Überzeugungskraft einzuschätzen, wobei sich die Aussagen in ihrer wissenschaftlichen Aussagekraft voneinander unterscheiden und reihen lassen: anekdotische Darstellungen, deskriptive Darstellungen, korrelative Darstellungen sowie kausale Darstellungen. Hierbei waren die Studierenden am ehesten in der Lage, anekdotische Darstellungen korrekt und in ihrer Überzeugungskraft als geringer zu erkennen, während sie bei den kausalen Zusammenhängen die höchsten Fehlerquoten hatten. Hinzu kommt, dass sich bei der ebenfalls erfragten und über offene Antworten erhobenen Begründung für die Qualität der Belege insbesondere bei der Einschätzung der anekdotischen Darstellungen Eigenheiten ergaben: Für anekdotische Schilderungen führten die Testpersonen die Relevanz mehr an und dafür weniger die Methoden oder die Interpretierbarkeit von quantitativen Daten. Die Unterschiede innerhalb der drei anderen Evidenztypen und teils über diverse Kategorien waren vernachlässigbar gering.

Zusätzlich interessierte man sich noch dafür, welche der vier Evidenztypen die Personen in einen Forschungsbericht übernehmen würden und wie die Personen ihre Auswahl begründen. Auch hier waren die anekdotischen Darstellungen auffällig: Sie kamen am seltensten vor. Bezogen auf die vier ausgewerteten induktiv gebildeten Kategorien wurde nur eine als statistisch auffällig häufiger bei einem Evidenztypus angeführt: Beim Typus kausaler Evidenz benannten die Testpersonen häufiger quellenbezogene Gründe.

Im Lichte dieser Studie sind die Personen also primär dazu in der Lage, den epistemisch schwächsten Evidenztypus, die rein anekdotische Darstellung, ausreichend zu erkennen; zum Teil treten im Vergleich der anderen Evidenztypen Trennschärfe-Probleme auf. Dieses Problem, Urteile ausreichend differenziert zu begründen, setzt sich zusätzlich bei Fragen der Verwendungsabsicht für einzelne Evidenztypen fort. Insgesamt verweist die singuläre Studie auf Problemlagen, die sich im Licht der Studien mit geschlossenen Formaten so nicht bzw. anders darstellen.

3) *Einschätzungen einzelner Metadaten*

Die letzte Gruppe von Ergebnissen und Studien fokussiert auf die gleichsam geringste Korngröße: die der Einschätzung von einzelnen Metadaten. Ihnen widmet sich dieser Abschnitt; in ihm werden nur die Ergebnisse aus den geschlossenen Fragen berichtet, da die qualitativ ausgewerteten Ergebnisse andernorts, nämlich vor allem bei den Problemlagen (s. Teilkap. 3.4) Gegenstand sein werden bzw. bei der häufigkeitsbasierten Auszählung bereits in Tabelle 11 berichtet worden sind.

Methodisch ist bei den geschlossenen Aufgabenformaten anzumerken, dass die Einschätzungen auf der Basis von Nominierungen in Vergleichen von Dokumenten erfolgten, wobei sich das Skalenniveau unterschied. Es reichte davon, dass von zwei Quellen diejenige mit der höheren Expertise anzuführen war (Macedo-Rouet et al., 2013, Experiment 1), bis hin zu ganzen Skalen, auf denen die Zuverlässigkeit von Websites mit einzelnen Items einzuschätzen war (Westerwick, 2013, Studie 1). Die Studienbefunde lassen sich zu drei Ergebnislinien verdichten:

- *Bei experimentell auf eindeutige Kontraste angelegten Stimuli in Form von Dokumenten mit überschaubar vielen Metadaten, deren Differenz nachvollziehbar ist, gelingt es Personen, die Unterschiede zu erkennen und zu benennen.* Dies betrifft mehrere Metadaten: Expertise der Quellen (Anttonen et al., in Druck; Gottschling et al., 2020; Kiili et al., 2022, 2023; Macedo-Rouet et al., 2013, Experiment 1; Potocki et al., 2020; Salmerón et al., 2016, Experiment 2; Strømsø et al., 2011; Wertgen & Richter, 2020, Experiment 1), Benevolenz der Quellen (Anttonen et al., in Druck; Flanagan & Metzger, 2007; Gottschling et al., 2020; Kiili et al., 2022, 2023), eine Kombination von Vertrauenswürdigkeit und Benevolenz (Gottschling

& Kammerer, 2021; Gottschling et al., 2019) sowie die Zuverlässigkeit des Publikationsorgans (Herrero-Diz et al., 2019; Kiili et al., 2022; Wertgen & Richter, 2023, Experimente 1 und 2; Westerwick, 2013, Studie 1).

Sind die Kontraste hingegen nicht so deutlich angelegt, weil die einzuschätzenden Metadaten aus ökologisch valideren Kontexten wie selbstaufgesuchten Websites stammen, haben jugendliche Personen teils erhebliche Probleme, die Expertise einzuschätzen. Dazu war beispielsweise in der Studie von Coiro et al. (2015) nur ein knappes Drittel in der Lage. Kiili, Leu, Marttunen et al. (2018) kommen zu einem ähnlichen Ergebnis: In ihrer Studie bezweifelten fast zwei Fünftel der Personen nicht die Expertise der Organisation, die eine kommerzielle Website hergestellt hatte.

- *Die Einschätzungen erfolgen nicht nur nach den Metadaten allein, sondern unter Rückgriff auf externe Aussagen und interne Ressourcen.* Personen nutzen Kontextinformationen. Sie verwenden beispielsweise die Übereinstimmung mit anderen Aussagen dazu, die Expertise und Vertrauenswürdigkeit einzuschätzen: Je konsistenter die Aussagen von Quellen mit anderen Aussagen sind, desto höher werden beide Dimensionen der Glaubwürdigkeit bei einzelnen Dokumenten beurteilt (van der Heide & Lim, 2016). Gleiches gilt für plausiblere, mit dem Weltwissen übereinstimmende und mit den eigenen thematischen Überzeugungen konsistente Aussagen: Wenn Personen textuelle Aussagen für inhaltlich stimmiger und mit den thematischen Überzeugungen kongruenter halten, beurteilen sie die Quelle als vertrauenswürdiger (Wertgen & Richter, 2020, Experiment 1; Wertgen & Richter, 2023, Experimente 1 und 2).
- *Es bestehen Personenunterschiede bei der Einschätzung von Metadaten.* Beispielsweise haben jüngere Personen mit *schlechterem Leseverstehen* Probleme damit, die Metadaten zur Expertise korrekt einzuschätzen (Macedo-Rouet et al., 2013, Experiment 2, Kontrollgruppe). Auch in Leistungsgruppen beim Evaluieren fallen Unterschiede dahingehend auf, dass leistungsschwächere Personen die Wichtigkeit von Metadaten zum Publikationsorgan und zur Autorin/zum Autor höher einschätzen, dass dies ihnen aber nicht dabei hilft, Dokumente zu differenzieren (Braasch et al., 2009). Anders gesagt: Sie machen ihr Urteil von Metadaten abhängig, was ihnen aber nicht für das Resultat dient. Daneben fällt es Personen schwerer, bei *unvertrauten Themen* die Vertrauenswürdigkeit von Quellen trennscharf zu unterscheiden, was bei vertrauteren Themen weniger der Fall ist (Barzilai, Thomm & Shlomi-Elooz, 2020, Experimente 1 und 2). Ähnliches gilt für persönliche Einstellungen zu einem Thema (Kiili, Leu, Marttunen et al., 2018). Dabei besteht die Auffälligkeit, jenen Quellen mehr zu vertrauen, die der eigenen Perspektive mehr entsprechen (s. dazu Teilkap. [3.3.3](#)). Ein weiterer Faktor ist das *Alter* der Personen, wobei sich ein höheres Alter nicht zwangsläufig bei allen metadatenbezogenen Evaluationen niederschlägt.

So beobachteten Potocki et al. (2020), dass die Fähigkeit zur Einschätzung von Expertise mit steigendem Alter zunahm, aber die Differenzierung der Benevolenz keinen Alterseffekten unterlag. Mit anderen Worten: Die (mangelnde) Expertise zu erkennen, fiel jüngeren Personen (Klasse 5 und 7) schwerer als Jugendlichen der Klassenstufe 9 und Studierenden. Bei der Benevolenz ähnelten sich die Gruppen hingegen, sie konnten ähnlich gut bzw. schlecht unterscheiden, ob die Quelle gute Absichten hegte. Abel et al. (in Druck) stellten ebenfalls fest, dass es älteren Jugendlichen besser und schneller gelang, die Expertise und die Benevolenz bei gezielt in diesen Belangen experimentell manipulierten Social-Media-Posts korrekt einzuschätzen. Außerdem berichteten sie, dass die Benevolenzurteile den Personen schwerer fielen als die zur Expertise.

4) Zusammenfassung zu den produktbezogenen Studienergebnissen

Die Studienlage zum Beurteilen von Metadaten ist von einer hohen Menge an Empirie gekennzeichnet, was der intensiven Erforschung dieses Sourcing-Segments geschuldet ist. Darum wird an dieser Stelle noch einmal rekapituliert und gebündelt, was die empirische Essenz ist. Aus den bisherigen Darstellungen des Forschungsstands kristallisieren sich die folgenden großen Linien der Befunde heraus:

- *Leserinnen und Leser differenzieren Glaubwürdigkeit, Expertise und Vertrauenswürdigkeit – vor allem bei deutlicher Salienz.* In experimentellen Kontexten mit definierten Dokumentensets und gezielter Variation der Metadaten glückt es den Personen überwiegend, die Pole in den Glaubwürdigkeitskontinua korrekt einzuschätzen. Sie können also vertrauenswürdiger von weniger vertrauenswürdigen Dokumenten, Quellen, Metadaten und Aussagen unterscheiden und ihre Urteile teils darauf abstützen. Im Mittelfeld ist diese Trennschärfe so nicht mehr feststellbar.
- *Neutrale Informationen von epistemischen Autoritäten genießen mehr Vertrauen.* Je mehr Fachwissen der Quelle unterstellt wird und je neutraler sie informiert, desto mehr Vertrauen spiegeln die Urteile über die Dokumente bzw. über die Inhalte wider.
- *Die Glaubwürdigkeitsurteile und die Begründungen erfolgen mehrdimensional.* Personen nutzten mehrere Begründungsoptionen, darunter den Inhalt, die Verwertbarkeit (Relevanz), das Vorwissen und – nicht notwendigerweise stark ausgeprägt – das Sourcing. Bei vertrauteren Themen fungiert das Vorwissen eher als Benchmark, bei unvertrauteren Gegenständen nimmt die Bedeutsamkeit der Metadaten zu. Expertise der Person und Publikationsorgan scheinen aktuell als besonders wichtige Kriterien zu fungieren.

- *Die Zusammenhänge von Urteilen und Begründungen in variablenzentrierten Zugängen sind untererforscht und wirken komplex.* Auffälligerweise werden die Urteile bezüglich der Vertrauenswürdigkeit und die Begründungen zwar durchaus häufig erfasst, indem die Begründung dem Urteil in den Designs buchstäblich folgt. Dennoch werden diese beiden Indizes selten gemeinsam in der Analyse aufeinander bezogen. Eine auf variablenzentrierten Zugängen basierende Auswertung ist daher bislang nur dazu geeignet, die Unverbundenheit beider Urteile zu illustrieren. Anders gesagt: Die Frage, welche Art von Begründung für welches Urteil günstig oder ungünstig war, ist aktuell noch wenig geklärt. Möglicherweise sind dafür valenzbezogene Auswertungen besser geeignet, also eine (Re-)Fokussierung darauf, *wie* etwas beurteilt/begründet wurde statt alleinig *was* beurteilt wurde.
- *Es gibt Leistungsunterschiede in Personengruppen.* Besonders deutlich ist, dass Expertinnen und Experten die Dokumente präziser, homogener und trennschärfer in ihrer Glaubwürdigkeit zu beurteilen imstande sind. Zudem sind ältere Personen tendenziell ebenfalls mehr dazu in der Lage, teils auch Leistungsgruppen zugunsten der Personen mit mehr Vorwissen oder einer besseren Leistung in anderen kognitiven Maßen. Die Leistungsunterschiede sind nicht immer und in jeder empirisch fokussierten Variable gleich gut sichtbar.

3.2.4.2 Prozesse

Beurteilungsprozesse mit Fokus auf die Metadaten sind sowohl mit unspezifischen als auch mit spezifischen Maßen operationalisiert worden. Beide Perspektiven sind komplementär und werden hier aufgrund der eher geringen Anzahl an Studien in einem ersten Schritt en bloc betrachtet, wobei die unspezifischen Prozessmaße primär aus einer personenzentrierten Perspektive fokussiert werden. In einem zweiten Schritt werden die Zusammenhänge von Prozess- und Produktmaßen dargestellt.

1) Prozessperspektiven: Muster bei den evaluativen Prozessen

Die Studien mit Lautdenk-Protokollen haben die mündlichen Äußerungen, die parallel zur Aufgabenbearbeitung entstanden sind, auf unterschiedliche Weisen ausgewertet. Trotz der Heterogenität in den Operationalisierungen und den damit einhergehenden unterschiedlich differenzierten Befunden in den einzelnen Studien lassen sich diese studienübergreifenden Muster benennen:

- *Spezifisches Beurteilen erfolgt mehrdimensional.* Die Studien mit spezifischer Auswertung der verbalen Prozessdaten haben mit unterschiedlichen Granularitätsgrößen operiert, die nicht immer gleichartig waren und darum

nicht direkt miteinander zu vergleichen sind. Die teils sehr detaillierten Deskriptionen der Kodierschemata (Anmarkrud et al., 2013; Barzilai & Zohar, 2012; Cho et al., 2017; von der Mühlen et al., 2016) bzw. die Berichte der Auftretenshäufigkeiten einzelner Subkodes (Anmarkrud et al., 2014; Brand-Gruwel et al., 2017; Cho, Afflerbach & Han, 2018; Gerjets et al., 2011; Goldman et al., 2012; Kiili et al., 2008, 2019; McGrew, 2021b; Strømsø & Bråten, 2014; Ulyshen et al., 2015; Walraven et al., 2009; Wiley et al., 2009, Experiment 1) zeugen von einer großen Bandbreite. Freilich waren manche Kodes sehr selten bzw. kamen im Fall deduktiver Kodierraster empirisch nicht in den Experimenten vor, waren also mit den Daten nicht in Einklang zu bringen. Einzelne Studien betrachteten neben den Dimensionen noch die *Ebenen der Strategieanwendung*, indem oberflächen- bzw. tiefenorientierte Strategien entweder nach ihrer Funktion zu einer Kategorie gebündelt (Brand-Gruwel et al., 2017; Ulyshen et al., 2015) oder bezogen auf die Hauptkodes in ihrer Verstehensunterstützung als höherwertig bzw. weniger günstig kodiert wurden (Cho, Afflerbach & Han, 2018; Cho et al., 2017). Somit lassen sich Verarbeitungstiefen unterscheiden, die sich auf den Bezugspunkt und das Ziel des Verstehens beziehen lassen.

- *Spezifische Beurteilungsprozesse hängen teils mit anderen Verstehensprozessen zusammen, sind also Teil eines größeren Verbunds von Verstehensprozessen.* Sowohl Studien mit deskriptiven Darstellungen der evaluativen Prozesse (Anmarkrud et al., 2014; Kiili et al., 2008; Strømsø & Bråten, 2014; Walraven et al., 2009) als auch mit Korrelationsanalysen (Anmarkrud et al., 2013, 2014; Cho, Afflerbach & Han, 2018; Cho et al., 2017) weisen auf ein Neben- bzw. Miteinander evaluativer Prozesse hin. Die Studien mit Korrelationsanalysen ergeben: Ein stärker ausgeprägtes Sourcing im Sinne häufigerer quellenbezogener Evaluationen korrespondiert mit mehr Prozessen des Verstehens (inter-)textueller Informationen (Anmarkrud et al., 2013, 2014; Cho et al., 2017; Kammerer et al., 2021; Ulyshen et al., 2015) sowie metakognitiven Strategien der Überwachung und Steuerung (Anmarkrud et al., 2013; Cho et al., 2017). Mitunter ist die Auftretenshäufigkeit des Sourcings unabhängig von der anderer Prozesse (Cho, Afflerbach & Han, 2018; Kammerer et al., 2021).
- *Leistungsgruppen unterscheiden sich darin, wie zielgerichtet und elaboriert sie das Sourcing betreiben.* Gebildet wurden Leistungsgruppen über ihren Status als Expertinnen und Experten bzw. Laienpersonen resp. Novizinnen und Novizen (Brand-Gruwel et al., 2017; von der Mühlen et al., 2016; Wineburg, 1991) – und damit a priori – oder leistungsdatenbasiert über ihren Erfolg in der Korrektheit ihrer Aufgabenprodukte innerhalb der jeweiligen Studie (Goldman et al., 2012; Rott & Gavin, 2015; Wiley et al., 2009, Experiment 1) – also ex post. Das deutlich erkennbare Muster dabei: *Leistungsstärkere Personen bzw. solche mit hoher Expertise legen ein Verhalten an den Tag, das auf ein Erkennen von zuverlässigeren Dokumenten schließen lässt.*

Dieses Muster hat sich studienübergreifend bereits in den *unspezifischen Prozessdaten* gezeigt. So ermittelten Wiley et al. (2009, Experiment 1), Goldman et al. (2012) sowie Rott und Gavin (2015), dass erfolgreichere Personen dazu neigten, wiederholt bzw. länger die zuverlässigeren Websites zu konsultieren, während die weniger erfolgreichen Personen ein Gegenteiliges Muster aufwiesen. Sie lasen mehrfach die weniger zuverlässigen digitalen Dokumente. Im Vergleich von Expertinnen und Experten und weiteren Personen ergaben sich Differenzen zwischen den Gruppen insofern, als die Personen mit hoher Expertise länger Suchmaschinenergebnisse betrachteten, also mehr Zeit für die Analyse investierten, aber teils weniger lang auf Websites verblieben (Brand-Gruwel et al., 2017) und – im Falle von Faktencheckerinnen und -checkern – schneller die zu beurteilenden Seiten verließen und mit ihrem „lateralen Lesen“ begannen, also einem Recherchieren über die Quellen (Wineburg & McGrew, 2019).

Was sich in den unspezifischen Prozessindikatoren andeutete, setzt sich prinzipiell fort, wenn man sich den *spezifischen Prozessdaten* aus Lautdenk-Protokollen widmet, selbst wenn die Muster dort nicht so evident sind. Generell neigen Expertinnen und Experten zu erheblich mehr metadatenbasierten Einschätzungen (von der Mühlen et al., 2016; Wineburg, 1991). In dem Fall, in dem sich bei der reinen Anzahl der Äußerungen keine Differenzen ergaben, zeigten sie sich auf einer feinteiligeren Ebene der Auswertung bei Brand-Gruwel et al. (2017). Sie manifestierten sich darin, dass die für die Testleistung (vertrauenswürdiger Websites benennen) günstigeren spezifischen Evaluationen häufiger bei den Expertinnen und Experten erfolgten. Die Novizinnen und Novizen in derselben Studie neigten demgegenüber eher zu oberflächlichen Einschätzungen, die wiederum negativ mit den Evaluationsergebnissen im Test zusammenhingen. In diesen Differenzen offenbarten sich unterschiedliche Fokusse bei der Evaluation.

Andere Studien mit Fokus auf studentische Leistungsgruppen gelangten zu ähnlichen Befunden, nämlich dass sich die kodierten Äußerungen in den Lautdenk-Protokollen nicht durchgängig unterschieden (Goldman et al., 2012; Rott & Gavin, 2015; Wiley et al., 2009, Experiment 1). Leistungsstärkere Personen nutzten jedoch zum einen mehr das Sourcing mit Blick auf die Metadaten und zum anderen die inhaltliche Eignung stärker für die Einschätzung der Websites (Wiley et al., 2009, Experiment 1). Nicht immer erreichten die in der deskriptiven Statistik feststellbaren Differenzen zwischen den Leistungsgruppen statistische Signifikanz (Goldman et al., 2012; Rott & Gavin, 2015). Gleichwohl konturierten Goldman et al. (2012) in ihren feinteiliger vorgenommenen Kodierungen leicht andere Schwerpunktsetzungen bei den evaluativen Kommentaren der leistungsstärkeren Personen: In der Tendenz achteten diese Personen mehr auf die inhaltliche Gültigkeit und die Vertrauenswürdigkeit der Quellen als die leistungsschwächeren Personen.

Die prozessfokussierten Studien machen auf die Vielgestaltigkeit der Evaluationen beim Sourcing aufmerksam, nach denen diverse Metadaten den Gegenstand lesendenseitiger Einschätzungen werden (können), wobei die Verarbeitungstiefe variiert. Solche Prozesse erfolgen jedoch nicht losgelöst von anderen Prozessen der Aufgabenbearbeitung, sondern sind Ausdruck und Bestandteil eines Verbunds von kognitiven und teils metakognitiven Prozessen. Diese Komplexität impliziert, dass Personen unterschiedlich gut dazu in der Lage sind, derartige Prozesse anzuwenden. Sowohl unspezifische als auch spezifische Prozessindikatoren konturierten ein kohärentes Bild: Leistungsstärkere Personen und solche mit Expertise in dem jeweiligen Bereich wirken zielgerichteter, analytischer und präziser in ihrem Sourcing. Gleichwohl unterscheiden sich produktbasiert gebildete Leistungsgruppen in Studien mit in sich homogenen Stichproben (Studierende) nicht durchgängig und nicht in der Differenz stark ausgeprägt voneinander.

2) Prozess- und Produktperspektiven: Zusammenhänge von evaluativen Sourcingprozessen und Produktmaßen

Die Zusammenhänge von evaluativen Sourcingprozessen und Produkten sind ein Desiderat. Unterscheiden lassen sich in der bislang vorhandenen Empirie zwei Zugänge. Der erste *personenzentrierte Zugang* basiert darauf, anhand der Leistungsdaten Gruppen von erfolgreicheren und weniger erfolgreichen Personen zu bilden und deren Prozessdaten zu untersuchen. Diese Ergebnisse wurden bereits im Rahmen dieses Kapitels dargestellt. Die *variablenzentrierten Analysen*, bei denen über gesamte Stichproben die Zusammenhänge zwischen Prozess- und Produktmaßen fokussiert werden, sind in der Vergangenheit selten durchgeführt worden, was angesichts der Vielzahl an Untersuchungen mit Produktmaßen nur umso mehr auf das Desiderat verweist (s. aber die Studien zu den Diskurssynthesen und Sourcingprozessen in Teilkap. [5.2.2.2](#)).

Die punktuell vorliegenden Ergebnisse aus variablenzentrierten Analysen unterstreichen die Bedeutung des Evaluierens von Metadaten. Je mehr evaluative Kommentierungen die Personen während den jeweiligen Aufgabenbearbeitungen vorgenommen hatten, desto

- stärker stimmten die Einschätzungen der Glaubwürdigkeitsurteile mit einer Expert:innen-Norm überein, das galt vor allem für die Beurteilung von Metadaten (Brand-Gruwel et al., 2017; von der Mühlen et al., 2016);
- höher schätzten sie die Glaubwürdigkeit ausbalanciert berichtender Dokumente ein bzw. für desto weniger glaubwürdig hielten sie alarmistisch berichtende Dokumente (Anmarkrud et al., 2014);
- höher waren – teilweise – die Verstehensleistungen in offenen Aufgaben; diese Befunde gestalteten sich teils aber nicht als direkt linear (Cho, Han & Kucan,

2018), sondern wurden über andere Prozesse wie Informationsrecherche und Verstehensprozesse mediiert (Cho et al., 2017).

Es gibt demnach Hinweise darauf, dass evaluative Prozesse beim Lesen mit besseren Leistungen sowohl im Sourcing als auch bei weiteren anspruchsvollen Verstehensprodukten einhergehen. Gleichwohl sind gerade vor dem Hintergrund der umfassenden Empirie zu Produkten die vielen Desiderate zu den Zusammenhängen von Prozessen und Produkten unübersehbar, insbesondere mangelt es an Studien zu Prozessverbünden des Evaluierens in situ, die dabei helfen, die evaluativen Produktmaße in ihrer Genese zu verstehen.

3.3 Determinanten und interindividuelle Unterschiede

Die interindividuellen Unterschiede beim Sourcing sind in einer Vielzahl von Studien untersucht worden, allerdings gestalten sich die die Effekte individueller Merkmale nicht immer eindeutig (Anmarkrud et al., 2022). Dies ist im Großen und Ganzen auch das an dieser Stelle bereits antizipierte Ergebnis der umfassenden Sichtung von Studien, die in diesem Kapitel in fünf Schritten erfolgt. In einem ersten Schritt (3.3.1) sind Personengruppen Gegenstand der Betrachtung, die sich in ihrem Sourcingverhalten differenzieren lassen. Die sich anschließende Darstellung ist eher variablenzentriert und erfolgt zuvorderst tabellarisch kondensiert, wobei die Produktmaße stark dominieren. Bei den vier Konstrukten handelt es sich um das thematische Vorwissen (3.3.2), thematische Überzeugungen (3.3.3), epistemische Überzeugungen (3.3.4) sowie die Lesefähigkeiten (3.3.5).

3.3.1 Gruppen von Personen beim Sourcing (und weiteren Prozessen)

Sourcing als komplexer Prozessverbund erfordert für sein Gelingen, dass die Leserinnen und Leser ein aufgaben-, dokumenten- und kontextangemessenes Vorgehen wählen. Dabei ist allein wegen der schieren Zahl beteiligter Komponenten und Prozesse davon auszugehen, dass sich Personen im Vorgehen unterscheiden. Dies ist die Grundannahme des „Cognitive Affective Engagement Model of Multiple Source Use“ (CAEM; List & Alexander, 2017b, 2018a). Das mehrphasige Modell postuliert, dass Personen ausgehend von einer ersten Phase durch das Zusammenspiel von wahrgenommenen Aufgabenerfordernissen und individuellen motivationalen Merkmalen in einer zweiten Phase eine von vier Grundhaltungen bzgl. des Sourcings aufweisen bzw. diese aktivieren. Diese Grundhaltung speist sich aus zweierlei: motivationalen und kognitiven Merkmalen, die sich in ihrer jeweils hohen oder geringen Ausprägung miteinander kombinieren lassen. Je

nach Kombination ergeben sich dementsprechend vier Grundhaltungen: 1) *nicht-engagiert* mit geringer Motivation und geringen dispositionalen evaluativen (Sourcing-)Fähigkeiten, 2) *affektiv engagiert* mit hoher Motivation, aber geringen habituellen Fähigkeiten, 3) *evaluativ* mit geringer Motivation, aber ausgeprägten habituellen Fähigkeiten und 4) *kritisch-analytisch* mit hoher Motivation und hohen Fähigkeiten. Je nach Grundhaltung bearbeiten die Leserinnen und Leser dann in der dritten Phase die jeweilige Aufgabe auf unterschiedliche Art und Weise, wobei die Konfiguration aus Motivation und Fähigkeiten, aus Wollen und Können, den Ausschlag gibt, mit welchem Aufwand, welcher Verarbeitungstiefe und welchem Resultat die (Sourcing-)Prozesse ausgeführt werden. Nur wenige Studien haben bislang explizit das CAEM getestet und mit einem personenzentrierten Analysezugang geprüft (Strømsø et al., 2020; Tarchi, 2019). In einigen zusätzlichen Studien (Lawless & Kulikowich, 1996; Lawless et al., 2002) wurden teils sehr allgemeine Navigationsdaten genutzt, die aus Sicht des Sourcings nur als Hinweise zu werten sind, aber das Gesamtbild von Personengruppen zu komplettieren helfen.

Gemäß der angenommenen Kongruenz von Dispositionen bzw. Grundhaltungen und Verhaltensindikatoren lassen sich demnach Personen unterscheiden, und die Forschung hat dazu in verschiedenen Studien mit Viert- bis Sechstklässlern (Hautala et al., 2018; Lawless et al., 2002), älteren Sekundarschuljugendlichen (Kiili et al., 2008; Strømsø et al., 2020) sowie Studierenden (Barzilai et al., 2015; Lawless & Kulikowich, 1996; List & Alexander, 2017c; List et al., 2016a; Tarchi, 2019) Ergebnisse vorgelegt. Die Datenbasis für die ermittelte Gruppen- bzw. Clusterzugehörigkeit bildeten Logfile-Daten zum Navigieren in hypertextuell organisierten Lern-Apps (Lawless & Kulikowich, 1996; Lawless et al., 2002) bzw. Dokumenten-Bibliotheken (List & Alexander, 2017c; List et al., 2016a), Eye-Tracking-Daten zur Betrachtung von Suchmaschinenergebnissen (Hautala et al., 2018), retrospektiv berichtete Vorgehensweisen in aufgabenbezogenen Strategiefragebögen (Tarchi, 2019), Lautdenk-Protokolle (Barzilai et al., 2015) sowie Kombinationen aus Logfile-Daten und Lautdenk-Protokollen (Kiili et al., 2008). Zum Teil wurden noch motivational-affektive Merkmale bzw. Überzeugungen bei der Bildung der Cluster verwendet (Lawless et al., 2002; Tarchi, 2019) bzw. basierte die Zuweisung von Personen zu Gruppen gänzlich auf Individualmerkmalen und nicht auf Prozess- bzw. Produktindikatoren (Strømsø et al., 2020).

Die Studien nutzten als Vorgehen grosso modo die Clusteranalyse als statistisches Verfahren. Lediglich die Studie von Barzilai et al. (2015) stellt eine Ausnahme dar: In ihr wurde nicht per Clusteranalyse eine Typologie von Personen bestimmt. Vielmehr waren die in Lautdenk-Protokollen vier geäußerten Varianten des Sourcings ausschlaggebend. Die Gruppen von Personen (a) kein Sourcing, b) minimales Sourcing, c) geringes Sourcing und d) starkes Sourcing) ergaben sich daraus, ob es einen Mangel an dokumentenbasierten Sourcingaktivitäten (a) oder noch zusätzlich jeweils das Bemerken von Quellen (b), die Verbindung

von Quelle-Inhalt-Verbindungen (c) oder die Verbindungen von Quellen untereinander (d) gab. Die Gruppe mit starkem Sourcing wies also die umfassendsten Sourcingaktivitäten auf, nämlich alle drei Varianten des Sourcings, die Gruppen mit minimalem bzw. geringem Sourcing nur eine resp. zwei.

Einen Schwerpunkt bei der Suche nach in sich möglichst konsistenten Personengruppen bildeten überwiegend Prozessdaten, teils unter Rückgriff auf Individualmerkmale (wie die Einstellungen und das Interesse bei Tarchi, 2019, bzw. das Interesse in der Studie von Lawless et al., 2002). Die Studie von List et al. (2016a) ist eine Ausnahme in diesem Muster, denn sie stützte sich auf rein produktbezogene Maße beim Sourcing, nämlich auf die Gesamtzahl verwendeter Dokumente in den Antworten der Probandinnen und Probanden im Allgemeinen nebst dem Vorkommen von vier unterschiedlich vertrauenswürdigen Dokumenten im Besonderen. Eine weitere Ausnahme stellte das Vorgehen von Strømsø et al. (2020) dar, denn das Forschungsteam nutzte drei Individualmerkmale: Interesse und Einstellungen als motivationale Maße sowie die per Test gemessene Fähigkeit zum Sourcing. Damit ist diese Studie in ihrer Operationalisierung am dichtesten an der dispositional modellierten Logik der Grundhaltung gemäß CAEM (List & Alexander, 2017b, 2018a).

Ehe im Folgenden die Ergebnisse der Personengruppen und deren Unterschiede berichtet werden, seien die insgesamt nur wenigen Befunde zu den motivationalen Maßnahmen vorweggenommen: Das Interesse erwies sich mit Ausnahme der Studie von Strømsø et al. (2020) in den Studien als wenig geeignet, die Personengruppen voneinander zu unterscheiden (Lawless et al., 2002; Tarchi, 2019). In ihren Einstellungen zum Thema divergierten die Clustermitglieder hingegen deutlicher (Strømsø et al., 2020; Tarchi, 2019).

In der Reihung der normativen Wünschbarkeit des Sourcings gemäß CAEM werden die Profile im Folgenden empirisch unterfüttert und anhand der Studienergebnisse charakterisiert.

3.3.1.1 Gruppe 1: Personen, die kein Sourcing betreiben (nicht-engagierte Personen)

Die Studienergebnisse deuten insgesamt darauf hin, dass diese Personen einen oberflächlichen Zugang wählen, die Dokumente zügig und mit wenig erkennbarer Differenzierung zwischen glaubwürdigen und wenig glaubwürdigen Inhalten und Quellen zu rezipieren. Dies schlägt sich neben prozessualen Merkmalen ebenfalls in schriftlichen Produkten bzw. der Erfassung von Sourcingindikatoren in geschlossenen Aufgaben nieder.

Zu den Studien zunächst mit Blick auf *Prozessdaten* im Einzelnen: Gemessen an der rein dispositionalen Gruppenbildung waren solche Personen in der Studie von Strømsø et al. (2020) vorfindbar, sie hatten moderate Fähigkeiten im Sourcing, waren aber der Thematik gegenüber eher negativ eingestellt und äußerten

wenig thematisches Interesse. Solche explizit als „nicht-engagierte Personen“ bezeichneten Probandinnen und Probanden unterschieden sich mit Blick auf die fünf metrischen Prozessdaten wie Auswahl- und Lesezeit, Anzahl ausgewählter Texte und inhaltliche bzw. metadatenbasierte Begründungen der Auswahl in der statistischen Signifikanz auf den ersten Blick kaum von den beiden anderen Gruppen der Studie, den evaluativen und den kritisch-analytischen Personen. Allerdings kennzeichnete sie gemäß einer qualitativen Perspektive auf die gegebenen Antworten, dass ihre Begründungen für die Auswahl am oberflächlichsten ausfielen.

Personen, die kein erkennbares Sourcing betrieben, bildeten eine eigene Gruppe in der Studie von Barzilai et al. (2015). Geringes Sourcing ermittelten Kiili et al. (2008) in ihrer Gruppe der „unkritischen Leserinnen und Leser“. Das kennzeichnendste Merkmal dieser Gruppe war, dass sie einen erheblichen Teil der Lesezeit damit verbrachte, wenig glaubwürdige Websites zu lesen und wenig kritisch damit umzugehen. Mit Blick auf das Sourcing-Verhalten bei Suchmaschinen fanden Hautala et al. (2018) eine Gruppe, die gemäß Blickverhalten bei Suchmaschinen nur auf die Titel achteten. Diese Personen waren am schnellsten, sie machten aber auch die meisten Fehler in den Tests.

Ähnlich eher rasch und mit geringem Navigationsaufwand vorgehende Personen wurden in den Studien von Lawless und Kulikowich (1996) sowie Lawless et al. (2002) als „apathische Nutzerinnen und Nutzer“ bezeichnet. Sie verbrachten wenig Zeit mit der Lernsoftware und riefen nur wenige Hypertexte auf. List et al. (2016a) stellten anhand der Navigation innerhalb einer Bibliothek virtueller Dokumente fest, dass bei der Beantwortung einer komplexeren Frage fünf Cluster anhand der Konsultation von Dokumenten bestanden. Ein Cluster, das durch eine sehr geringe Anzahl von genereller Quellennutzung auffiel, wurde als „nicht-engagierte“ Personen bezeichnet. Diese Testpersonen nutzten das Set an Dokumenten kaum. Ganz ähnlich lässt sich das Cluster „limitierte Navigation“ aus einer weiteren Studie von List und Alexander (2017c) beschreiben, das zudem noch am wenigsten Metadaten ansteuerte. In beiden Studien fielen die beiden Cluster – nicht-engagierte Personen und solche mit limitierter Navigation – dadurch auf, dass ihre schriftlichen Antworten weniger Argumente (List & Alexander, 2017c; List et al., 2016a) und Begründungen (List & Alexander, 2017c) enthielten. In anderen Produktmaßen fanden sich jedoch keine überzufälligen Differenzen zwischen nicht-engagierten und weiteren Personen (List et al., 2016a).

Tarchi (2019) stützte sich in seiner Vierer-Typologie nicht auf Sourcing-Maße bei der Clusteranalyse, sondern nutzte neben den thematischen Überzeugungen vor allem zwei retrospektiv geschilderte Vorgehensweisen beim Lesen multipler Dokumente, die sich darin unterscheiden lassen, ob Personen mehr vom Erkennen und Sammeln von isolierten Informationen (Informationsakkumulation) bzw. vom Herstellen intertextueller Inferenzen (intertextuelle Elaboration) berichteten (s. dazu im Detail Teilkap. [4.2.1.1](#) und dort den Abschnitt b)

Selbstauskünfte zum intertextuellen Integrieren: Informationsakkumulation sowie intertextuelle Elaboration). Die Personen, die auf dieser Datenbasis als „nicht-engagierte Leserinnen und Leser“ klassifiziert wurden, zeichneten sich dadurch aus, dass sie relativ viele Informationen akkumulierten und am wenigsten von allen Gruppen intertextuelle Inferenzen herstellten. Sie fielen auch in ihren Produktmaßen dadurch auf, dass sie (im Vergleich zu den Personen mit den höchsten Ausprägungen vor allem bei der intertextuellen Elaboration und teils der der Informationsakkumulation) in ihren Urteilen zur Vertrauenswürdigkeit von Aussagen aus Dokumenten am wenigsten zwischen den Dokumenten differenzierten. Sie hielten subjektive und verzerrte Dokumente für valable Quellen und stuften vertrauenswürdige Dokumente als unzuverlässig ein. Zudem beurteilten sie die Wichtigkeit sämtlicher dokumentenbezogenen Metadaten als deutlich niedriger für ihre Glaubwürdigkeitsurteile.

3.3.1.2 Gruppe 2: Personen, die oberflächliches Sourcing betreiben (affektiv engagierte Personen)

Die Studien mit ihren heterogenen Fokussen und Befunden eint, dass Personen aus dieser Gruppe ein Vorgehen praktizieren, bei dem vor allem ein auf ein Dokument bzw. ein auf diskrete Teile abzielendes Lese- und Navigationsverhalten und die Präferenz für besonders zugängliche Dokumente auffällig waren. Es kommt dadurch zu einer gewissen Einseitigkeit und Unvollständigkeit in den Prozessen, partiell findet dies eine Entsprechung in eher affirmativ zu nennenden Mustern in den Produktdaten.

Die Studien bilden ein fragmentarisches Mosaik in ihren *prozessualen* Befunden, wie es die folgende Ergebnisdarstellung zeigen wird: Als „affektiv engagierte Leserinnen und Leser“ galten bei Tarchi (2019) jene Personen mit positiven Einstellungen zum Thema Impfstoffe und Impfungen, die zugleich die nahezu geringsten lesestrategischen Aktivitäten schilderten. Diese Personen fielen dadurch auf, dass sie besonders wenig nach intertextuellen Verbindungen suchten.

Rein auf das Sourcing bezogene Aktivitäten, die sich darin erschöpften, nur die Quellenmerkmale zu benennen, kennzeichneten das Vorgehen der Personen mit „minimalem Sourcing“ in der Untersuchung von Barzilai et al. (2015). Eine weitere Personengruppe aus derselben Studie zeichnete sich dadurch aus, dass sie Quelle-Inhalt-Verbindungen im Sinne von Interpretationen nur auf der Ebene einzelner Dokumente vornahm. Barzilai et al. (2015) nannten diese Personen bzw. deren Vorgehen „geringes Sourcing“, denn in dieser Gruppe fand Sourcing nur auf geringem Niveau statt.

Ein solches geringes Sourcing betrieben zudem gleich drei Gruppen bei Kiili et al. (2008): die „relevanzorientierten Leserinnen und Leser“, die „limitierten Evaluatorinnen und Evaluatoren“ und die „desorientierten Leserinnen und Leser“. Diese Personengruppen evaluierten wenig und kommentierten kaum die

Glaubwürdigkeit, aber sie lasen gleichwohl eher die glaubwürdigeren Quellen. Bei aller Unterschiedlichkeit in anderen Verhaltensindikatoren einte diese Gruppen ihr eher oberflächliches bis rudimentäres Sourcing.

In der Studie von List et al. (2016a) wurden analog zu jener von Kiili et al. (2008) ebenfalls drei Cluster ermittelt, deren Nutzungsverhalten eine gewisse Einseitigkeit aufwies. Kennzeichnend war, dass jede der drei Gruppen eine von vier Gruppen von Dokumenten ausschließlich privilegiert genutzt hatte. Das waren Personen, die sich entweder nahezu vollständig auf das unzuverlässigste Dokument, einen Blogbeitrag, stützten („unkritische Leserinnen und Leser“), auf die inhaltlich am leichtesten zugänglichen Dokumente wie Wikipedia-Einträge zugriffen („Nutzerinnen und Nutzer zugänglicher Quellen“) oder aber Presstexte besonders häufig frequentierten; diese letzte Gruppe rief zudem am meisten Texte auf und wurde darum als „engagierte Leserinnen und Leser“ bezeichnet. Das Muster, dass eines der Dokumente überproportional lang konsultiert wurde, nannten List und Alexander (2017c) in ihrer Untersuchung „vorrangige Navigation“. Dies schlug sich darin nieder, dass Personen dieser Gruppe ca. die Hälfte der gesamten Lesezeit auf einen Text verwendeten.

Personen mit einer eher oberflächlich zu nennenden Orientierung, nämlich die sogenannten „Erkunderinnen und Erkunder von technischen Features“, fanden Lawless et al. (2002) und Lawless und Kulikowich (1996) ebenfalls vor. Diese Kinder bzw. Erwachsenen wendeten sich eher den technischen Möglichkeiten wie Videos und Sounds sowie den Zusatzangeboten in der Lernsoftware zu, sie waren jedoch eher weniger auf den informationsreicheren Hypertextseiten aktiv.

Die *produktbezogenen* Auswertungen ergaben, dass die „unkritischen Leserinnen und Leser“ in der Studie von List et al. (2016a) am wenigsten argumentative Elemente in ihren schriftlichen Antworten einfließen ließen. Diese Differenz erwies sich als einziger Gruppenunterschied als signifikant. Die anderen Cluster ähnelten sich diesbezüglich, und in zwei weiteren Produktindikatoren gab es keine überzufälligen Differenzen zwischen allen Gruppen. Tarchi (2019) stellte fest, dass die „affektiv engagierten Leserinnen und Leser“ überzufällig jene Dokumente kritisch beurteilten, welche ihren eigenen Überzeugungen am stärksten widersprachen. Insgesamt kennzeichnet die Produkte damit eine Imbalance.

3.3.1.3 Gruppe 3: Personen, die ein unvollständiges Sourcing betreiben (evaluative Personen)

Testpersonen aus dieser Gruppe sind laut den empirischen Untersuchungen darum bemüht, die Dokumente hinsichtlich ihrer inhaltlichen Verbindungen und metadatenbezogenen Eignung zu durchdringen. Die Experimente mit *Prozessdaten* ergaben dieses Bild: In der Studie von Strømsø et al. (2020) fielen die evaluativen Personen (deskriptiv, nicht statistisch signifikant) dadurch auf, dass sie am wenigsten Zeit für die Auswahl und das Lesen der Texte verwendet hatten, dass

sie aber für die Begründung der Textauswahl am wenigsten inhaltliche Gründe zu Protokoll gaben. Im Vergleich zu den anderen beiden Personengruppen stützte sich ihre Auswahl in Relation zu den metadatenbezogenen Begründungen mehr auf die Einschätzung der Quellenmerkmale, was ihrem überdurchschnittlich hohen Wert im Sourcing gemäß vorgängigem Test entsprach. Die „evaluativen Leserinnen und Leser“ aus Tarchis Studie (2019) wiesen eine kritische Haltung gegenüber dem Thema auf, praktizierten aber am meisten Lesestrategien, vor allem der intertextuellen Elaboration. Dieses Vorgehen wandte die Gruppe am stärksten von allen insgesamt vier Gruppen an.

In der Studie von Hautala et al. (2018) ließ sich eine Gruppe von Jugendlichen eruieren, die zwei der drei Komponenten in Suchmaschinen-Resultaten häufig fixierten: Titel und Snippet, also die Kurzbeschreibung. Gelegentlich nutzten sie zudem die URLs. Anhand ihrer Navigationsdaten bestimmten List und Alexander (2017c) ein Cluster mit der Bezeichnung „verteilte Navigation“. Die Personen aus dieser Gruppe widmeten sich den Texten ungefähr gleich lang und riefen am zweithäufigsten im Vergleich aller sinnvoll bezeichnbaren Gruppen Metadaten auf.

Die *Produktdaten* in den Studien ergaben, dass evaluative Personen am meisten Referenzierungen im Vergleich zu anderen Profilen in schriftlichen Antworten vornahmen, allerdings auch zu deutlich kürzeren Textprodukten neigten (Strømsø et al., 2020). Solche Personen wiesen bei Tests zu Suchmaschinenergebnissen bessere Resultate auf als jene mit ausschließlicher Zuwendung zu den Titeln der Ergebnisse; zusätzlich waren ihre Leistungen vergleichbar mit jenen Personen mit dem aufwändigsten Sourcing (Hautala et al., 2018). Außerdem waren sie am kritischsten gegenüber offiziellen Dokumenten und schätzten teils die Vertrauenswürdigkeit unzuverlässiger Quellen höher ein, als andere Cluster es taten (Tarchi, 2019).

3.3.1.4 Gruppe 4: Personen, die ein umfassendes Sourcing betreiben (kritisch-analytische Personen)

Eine breite und zielbezogene Nutzung von Strategien und eine vertiefte Zuwendung zu einzelnen Dokumenten kennzeichnet das Vorgehen dieser Personengruppe. Ihre Produkte des Verstehens sind überwiegend die elaboriertesten bzw. erzielten die höchsten Leistungen im Spektrum.

Zunächst ein Blick auf die *Prozesse*: Strømsø et al. (2020) fanden eine Gruppe kritisch-analytischer Personen in ihrem Sample vor. Diese Personen verbrachten die meiste Zeit mit dem Lesen und der Auswahl der Dokumente und nutzten zudem die höchste Anzahl an Dokumenten. Personen mit positiven Einschätzungen zum Gegenstand der zu lesenden Texten und einer stärkeren Anwendung von sowohl intratextuellen Strategien (Informationsakkumulation) als auch

intertextuellen Vorgehensweisen (intertextuelle Elaboration) galten bei Tarchi (2019) als „kritische Leserinnen und Leser“.

Barzilai et al. (2015) ermittelten in ihrem Korpus ein Vorgehen bzw. eine Personengruppe mit starken Sourcingaktivitäten. Kennzeichnend war, dass diese Personen nicht nur die meisten Sourcingaktivitäten praktizierten, sondern überdies die einzigen waren, welche die Metadaten und deren Evaluationen dokumentenübergreifend verwendeten. Sie verglichen damit die Perspektiven der Dokumente. Kiili et al. (2008) konnten ebenfalls eine Gruppe besonders stark evaluativ vorgehender Personen extrahieren. Diese „vielseitigen Evaluatorinnen und Evaluatoren“ nutzten die längste Zeit zum Lesen, konzentrierten sich dabei auf die zuverlässigsten Websites und betrieben die meisten glaubwürdigkeitsbezogenen Kommentierungen und evaluierten auf der breitesten Basis.

Hautala et al. (2018) fanden einen Personentypus vor, der die Suchmaschinenergebnisse sowohl in Titel als auch bei den Snippets als auch die URLs durchgängig fokussierte („Leserinnen und Leser aller Komponenten“). Diese Personen verwendeten damit die verwertbaren Informationen gemäß den Eye-Tracking-Daten am ausgiebigsten und zeitlich mit am längsten. „Informationssuchende“ wurden die Testpersonen in den Studien von Lawless und Kulikowich (1996) sowie Lawless et al. (2002) genannt, die im Vergleich zu den anderen beiden vorgefundenen Personengruppen am meisten Zeit mit dem Lesen und der zielgerichtet wirkenden Konsultation jener Hypertextelemente verbrachten, welche die meisten verwertbaren Informationen enthielten.

Bei List et al. (2016a) wurde dasjenige Cluster als „kritisch-analytische Quellennutzerinnen und -nutzer“ bezeichnet, das in der zur Verfügung gestellten Dokumentenbibliothek ausschließlich die glaubwürdigsten Dokumente (Buchkapitel bzw. einen Fachartikel) konsultiert hatte. Ein ähnliches Cluster ermittelten List und Alexander (2017c) in ihrem Datensatz: die „kritischen Leserinnen und Leser“. Diese Gruppe las am längsten, konsultierte Metadaten zu den Dokumenten am häufigsten und widmete sich verschiedenen Dokumenten unterschiedlich lang, wobei sie einzelne Dokumente auffällig lang und anderen auffällig kurz lasen, sie differenzierten also in ihrer zeitlichen Zuwendung zugunsten einzelner Dokumente in der digitalen Bibliothek.

Mit Blick auf *Produktmaße* sind Personen mit hohem Sourcinganteil bzw. ausführlicherer Konsultation von vertrauenswürdigen Dokumenten anderen Personengruppen überlegen. Sie zeigen bessere Erinnerungsleistungen (Lawless & Kulikowich, 1996) und können Suchmaschinen-Resultate überwiegend korrekter einschätzen (Hautala et al., 2018). Die geschriebenen Texte auf der Basis mehrerer Dokumente sind von kritisch-analytischen Personen länger sowie thematisch umfassender und enthalten mehr Integrationsleistungen in der Form, dass Personen häufiger zwischen den verarbeiteten Dokumenten wechselten (Switches) (Strømsø et al., 2020). Tarchi (2019) zufolge waren die Einschätzungen der glaubwürdigsten Dokumente am prononciertesten ausgeprägt, außerdem fielen diese

Personen – analog zur Gruppe der evaluativen Leserinnen und Leser – dadurch auf, metadatenbezogene Kriterien für besonders wichtig für ihre Einschätzungen der Dokumente resp. deren Aussagen in der Vertrauenswürdigkeit zu halten.

3.3.1.5 Zusammenschau

Die hier berichteten Ergebnislinien aus neun Studien ergeben bei aller Unterschiedlichkeit in den Prozess- und Produktdaten und bei großer Divergenz in der Nähe und Distanz zum CAEM als Rahmen doch erste Muster. Erstens lassen sich die Personen unterscheiden. Je nach Studie wurden zwischen drei und sechs Gruppen gefunden. Diese Gruppen passen – zweitens – mehr oder minder gut zu den Profilierungen gemäß CAEM. Gleichwohl lassen sich die Gruppen danach differenzieren, wie stark sie verarbeitungsintensive Prozesse betreiben. Diese Prozesse betreffen die quantitative Nutzung von und Zuwendung zu Texten bzw. Dokumenten sowie qualitativ das Ausmaß an Integrations- und Sourcingprozessen. Die Produkte divergieren ebenfalls, doch sind hier die Differenzen nicht so groß wie bei den Prozessen, was sich teils erst in Extremgruppendifferenzen im Spektrum gezeigt hat.

3.3.2 Thematisches Vorwissen (Produktperspektiven)

Das Vorwissen wird aktuell in Modellen aus der Multiple-Documents-Literacy als Ressource für den Erfolg bei komplexen Leseaufgaben betrachtet (Britt et al., 2018; List & Alexander, 2019; Rouet & Britt, 2011; Stadtler & Bromme, 2014). So benennen List und Alexander (2019) das Vorwissen als eine Basis für eine erfolgreiche Strategienutzung, darunter im Sourcing. Stadtler und Bromme (2014) beschreiben ebenfalls den Stellenwert des Vorwissens als implizite Bedingung für einen angemessenen Umgang mit konfligierenden Informationen aus Dokumenten, um das Zustandekommen des Konflikts auf verschiedene Quellen zurückzuführen bzw. Perspektiven zu gewichten. Wie aus der Tabelle 12 hervorgeht, sind ausgerechnet Prozessmaße bestenfalls spärlich erforscht.

Die Studienlage lässt sich folgendermaßen bündeln:

- Ein erstes Muster aus den zusammengetragenen Studien zu den Vorwissenseffekten in Tabelle 12 ist daran zu erkennen, dass Vorwissenseffekte beim Beachten bzw. Benennen und bei der Verwendung von Metadaten noch kaum erforscht sind. Wo solchen Zusammenhängen empirisch nachgegangen wurden, zeigten sie sich nicht bzw. nicht eindeutig.
- Das zweite Muster, welches sich im unteren Teil der Tabelle 12, abzeichnet, betrifft die Divergenz der Befunde im deutlich stärker erforschten Evaluieren. Hier konnten im Verhältnis deutlich mehr Studien keinen Zusammenhang

von Vorwissen und produktbezogenen Evaluationsleistungen belegen als umgekehrt.

Dieser Widerspruch von hoher theoretischer Bedeutsamkeit und empirischer Bedeutung des Vorwissens wirft die Frage nach dem Zustandekommen der Diskrepanz auf. Anmarkrud et al. (2022) haben in ihrer primär auf Studien mit digitalen Leseanlässen basierenden Überblicksarbeit von individuellen Merkmalen und deren Zusammenhängen mit Sourcingleistungen ein ähnliches Befundmuster vorggefunden. In ihrer tiefergehenden Analyse stellten sie fest, dass positive Zusammenhänge in Studien vorfindbar waren, a) in denen das Vorwissen über Multiple-Choice-Tests oder mittels offener Fragen erfasst wurde, b) in denen das Vorwissen der Personen recht hoch war und c) in denen Gesundheitsthemen behandelt wurden. Dies illustriert, dass es moderierende Variablen zu geben scheint, welche die Heterogenität der Ergebnisse erklären könnten. Dazu gehört ebenfalls, dass die Operationalisierungen der Sourcingmaße erheblich divergieren und damit studienübergreifende Muster derzeit schwer erkennbar bzw. auf Studienmerkmale zurückführbar sind.

Tabelle 12: Zusammenhänge von Sourcingmaßen und Vorwissen

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativ-effekte
Metadaten beachten bzw. benennen	• Produkte: <i>korrekte Antworten zur Identifikation von Metadaten als Teil eines umfassenderen Tests</i> (Kiili et al., in Druck) • Prozesse: –	Produkte: <i>korrekte Angabe von explizit erfragten Metadaten</i> (Gottschling & Kammerer, 2021; Kammerer, Meier & Stahl, 2016) Prozesse • <i>Logfile-Daten: Auswahl von Metadaten beim proaktiven Sourcing</i> (List et al., 2017) • <i>Logfile-Daten: Lesezeit von verschieden vertrauenswürdigen Dokumenten</i> (Mason, Junyent & Tornatora, 2014) • <i>Sourcingaktivitäten in Lautdenk-Protokollen</i> (Gesamtwert: Barzilai et al., 2015)
Metadaten nutzen	• Produkte: – • Prozesse: –	• Produkte: <i>Belegen: Anzahl Referenzierungen</i> (Salmerón, Gil & Bräten, 2018a) • Prozesse: –

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativ-effekte
Metadaten evaluieren	Produkte • <i>Gesamtscore im Evaluationstest</i> (Forzani, 2018; Korrelation; Hämäläinen et al., 2021; Strømso et al., 2020) bzw. bei dem <i>Elaborationsgrad in den Begründungen</i> (Kiili et al., 2022) • <i>Beurteilung der Vertrauenswürdigkeit von Informationen von gelesenen Dokumenten</i> (Tarchi, 2019, partiell) • <i>Korrekte Antworten zur metadatenbasierten Begründung</i> als Teil eines umfassenderen Tests (Kiili et al., in Druck) • <i>Begründungen für Reihungen von Dokumenten</i> (Braasch, Bräten, Strømso & Anmarkrud, 2014) bzw. Anzahl von Quellennennungen in den Begründungen von Rangreihungen (Kessler et al., 2021) • <i>Wichtigkeit einzelner Beurteilungskriterien für ein holistisches Glaubwürdigkeitsurteil von Dokumenten</i> (Tarchi, 2019, partiell) Prozesse: <i>Anzahl evaluativer Kommentare in Lautdenk-Protokollen</i> (Kammerer et al., 2021)	Produkte • <i>Quellenevaluation</i> (latentes Maß: Barzilai & Eshet-Alkalai, 2015) • <i>Rangreihung der Dokumente nach Glaubwürdigkeit</i> (Braasch et al., 2023; Braasch, Bräten, Strømso & Anmarkrud, 2014; Kessler et al., 2021; Mason, Junyent & Tornatora, 2014; Mason, Scrimin, Tornatora et al., 2018) • <i>Beurteilung der Dokumentenglaubwürdigkeit</i> (Kammerer, Meier & Stahl, 2016; List et al., 2017; Londra & Saux, 2023, Experimente 1 und 2; Tarchi, 2021b) • <i>Beurteilung der Vertrauenswürdigkeit von Informationen von gelesenen Dokumenten</i> (Bräten et al., 2009 = Strømso et al., 2011; Tarchi, 2021a) • <i>Wichtigkeit einzelner Beurteilungskriterien für ein holistisches Glaubwürdigkeitsurteil von Dokumenten</i> (Bräten et al., 2009) • <i>Anzahl genutzter Kriterien für Begründungen der Glaubwürdigkeit von Dokumenten</i> (Mason, Junyent & Tornatora, 2014) • <i>Gesamtwert im Test für Begründungen beim Sourcing</i> (Kiili et al., 2021) • <i>Anzahl Begründungen für Auswahl von Dokumenten</i> (Bräten, Brante & Strømso, 2018) • <i>Beurteilung einzelner Metadaten</i> (Kanniainen et al., 2019; Kiili, Leu, Marttunen et al., 2018) • <i>Wichtigkeit einzelner Metadaten für Glaubwürdigkeitseinschätzung</i> (Strømso et al., 2011; Tarchi, 2021a, überwiegend) Prozesse: <i>Anzahl von Evaluationen von Website-Informationen in Lautdenk-Protokollen</i> (Ulyshen et al., 2015)

3.3.3 Thematische Überzeugungen (Produktperspektiven)

Thematische Überzeugungen werden definiert als subjektive Standpunkte gegenüber dem Wahrheitsgehalt von Aussagen (Bråten & Strømsø, 2020), also beispielsweise ob eine Person das Impfen für gesundheitlich riskant oder angemessen hält. Je nach Ausprägung dieser oftmals impliziten Überzeugungen fungieren diese als unbewusst wirkende Filter der Verarbeitung von Aussagen und können unkorrigiert den Aufbau umfassender und vor allem ausgewogener mentaler Repräsentationen komplexer Gegenstände unterminieren (Richter & Maier, 2018, s. dazu ausführlicher das Teilkap. 4.3.2). Die Leseforschung hat sich wegen der sowohl protektiven als auch unterminierenden Effekte der Überzeugungen auf den Kohärenzaufbau primär auf integrative Verstehensprozesse in Zusammenspiel mit dem „Validieren“ befasset und erste Theorien dazu vorgelegt (O’Brien & Cook, 2016b; Richter & Maier, 2017). Die Effekte auf das Sourcing sind hingegen noch nicht Gegenstand einer umfassenden Theoriearbeit geworden. Zum gegenwärtigen Stand wird aber angenommen, dass das Sourcing extensiver erfolgt, wenn textuelle Aussagen den eigenen Überzeugungen widersprechen und wenn weitere internale und externale Variablen eine Überwindung des sogenannten „Text-Überzeugungskonsistenz-Effekts“ ermöglichen bzw. einfordern (Richter & Maier, 2017).

Die Empirie zu den Effekten epistemischer Überzeugungen ist in bestimmten Aspekten des Sourcings wahlweise noch kaum vorhanden (Prozessmaße zum einen und Beachtung bzw. Benennung sowie Nutzung von Metadaten zum anderen, s. Tabelle 13) oder aber stark uneinheitlich. Dies wird in der Tabelle 13 im unteren Teil beim Evaluieren deutlich: Nicht nur sind sowohl positive als auch keine Zusammenhänge dokumentiert. Im Fall der positiven Zusammenhänge treten diese teils nur vereinzelt in den jeweiligen Studien auf, was in der Tabelle über die Kennzeichnung „partiell“ verdeutlicht wird.

Dass die Befundlage so heterogen ist, mag mit der Operationalisierung des Sourcings zu tun haben. Denn sehr globale Maße, die beispielsweise die inhaltliche Positionierung von Dokumenten in Relation zu den lesendenseitigen Überzeugungen verwischen, wirken per se weniger sensitiv als Maße, die hierauf Rücksicht nehmen. Das mag erklären, dass beispielsweise Tarchi (2019, 2021a) ermittelte, dass positive Überzeugungen mit höheren Glaubwürdigkeitseinschätzungen der Informationen von neutralen Quellen einhergehen. Negative Zusammenhänge von Einstellungen und Glaubwürdigkeitseinschätzungen der Informationen zeigten sich mitunter bei epistemisch unsicheren Quellen (Tarchi, 2019). Teils sind nur einzelne Facetten von inhaltlichen Überzeugungen mit Sourcingmaßen assoziiert (Braasch et al., 2023), teils sind sie je nach Ausprägung positiv oder negativ mit Glaubwürdigkeitsurteilen korreliert (Braasch et al., 2023; Kessler et al., 2021; van Strien et al., 2016). Das bedeutet – wie so häufig bei der aktuellen Forschungslage zu den Zusammenhängen von Sourcingleistungen und

Individualmerkmalen wie den thematischen Überzeugungen (Anmarkrud et al., 2022) –, dass Operationalisierung und Auswertung fürderhin einen genaueren Blick verdienen.

Tabelle 13: Zusammenhänge von Sourcingmaßen und thematischen Überzeugungen

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativ-effekte
Metadaten beachten bzw. benennen	• Produkte: – • Prozesse: –	• Produkte: <i>Erinnerung an Metadaten:</i> geschlossenes Frageformat (Salmerón, Sampietro & Delgado, 2020) bzw. offenes Frageformat (Bräten, Salmerón & Strømsø, 2016; Gottschling & Kammerer, 2021) • Prozesse: <i>Fixationszeiten</i> (van Strien et al., 2016, überwiegend)
Metadaten nutzen	• Produkte: – • Prozesse: –	• Produkte: Anzahl Zitationen (Salmerón, Sampietro & Delgado, 2020) • Prozesse: –
Metadaten evaluieren	Produkte • <i>Rangreihung von Dokumenten nach Glaubwürdigkeit</i> (Braasch et al., 2023, partiell; Kessler et al., 2021) • <i>Beurteilung der Glaubwürdigkeit ganzer Dokumente</i> (Pilotek et al., in Druck; Tarchi, 2021b, überwiegend; van Strien et al., 2016, partiell) • <i>Beurteilung der Vertrauenswürdigkeit von Informationen in gelesenen Dokumenten</i> (partiell: Tarchi, 2019, 2021a) • <i>Beurteilung einzelner Metadaten:</i> Expertise (Kiili et al., 2023, partiell), Benevolenz (Kiili et al., 2023, partiell) • <i>Begründungen für Rangreihungen</i> (Braasch, Bräten, Strømsø & Anmarkrud, 2014, partiell) bzw. Nutzung von Metadaten in den Rangreihungen (Kessler et al., 2021, partiell) • <i>Wichtigkeit einzelner Metadaten für die Vertrauenswürdigkeits-einschätzung von Informationen aus Dokumenten</i> (Tarchi, 2021a, partiell) Prozesse: –	Produkte • <i>Gesamttestwert für Evaluation von Websites</i> (Hämäläinen et al., 2021) • <i>Beurteilung der Glaubwürdigkeit ganzer Dokumente</i> (Braasch, Bräten, Strømsø & Anmarkrud, 2014; Kammerer, Meier & Stahl, 2016) • <i>Rangreihung von Dokumenten nach Glaubwürdigkeit</i> (Kiili et al., 2023) • <i>Beurteilung der Schlussfolgerungen in Dokumenten</i> (Bräten, Salmerón & Strømsø, 2016; Strømsø & Bräten, 2017; Strømsø et al., 2017) • <i>Valenz der Begründungen für die Rangreihung</i> (Braasch et al., 2023) • <i>Wichtigkeit einzelner Beurteilungskriterien für ein holistisches Glaubwürdigkeitsurteil von Dokumenten</i> (Tarchi, 2019) Prozesse: –

3.3.4 Epistemische Überzeugungen (Prozess- und Produktperspektiven)

Die epistemischen Überzeugungen sind ein Konstrukt, das sich auf die subjektiven Annahmen von Personen zu Merkmalen des Wissens und dessen Herstellung bezieht (Bråten, Strømsø & Ferguson, 2016). In der Leseforschung sind die epistemischen Überzeugungen primär, aber nicht ausschließlich bei den Integrationsprozessen fokussiert worden und ihre Zusammenhänge mit dem Aufbau komplexer Repräsentationen aus multiplen Dokumenten zu mehrheitlich kontrovers dargestellten Sachverhalten theoretisch modelliert (Bråten, Britt et al., 2011) und erforscht worden (s. dazu detailliert das Teilkap. 4.3.3). Da die epistemischen Überzeugungen im Teilkapitel 4.3.3 ausführlich behandelt werden, werden an dieser Stelle nur die wesentlichsten Punkte skizziert.

Epistemische Überzeugungen betreffen unterschiedliche Ebenen des Wissens (generisch, domänen-, themenspezifisch), werden in der Regel als mehrdimensional und damit bei Personen eher als System von subjektiven Theorien angenommen und werden in verschiedenen taxonomischen Ansätzen auf je eigene Arten klassifiziert. Etabliert haben sich bei den Taxonomien grundsätzlich zwei Zugänge bzw. Ordnungssysteme: Das erste, auf Dimensionen bezogene System (beschrieben in Teilkap. 4.3.3.1) beschreibt epistemische Überzeugungen als subjektive Theorien zum einen hinsichtlich der Merkmale von Wissen und zum anderen bezüglich der Genese des Wissens, postuliert für diese beiden Hauptdimensionen noch zwei Teildimensionen (*Merkmale des Wissens*: a) Wissen als einfach vs. komplex bzw. b) Wissen als bestimmt und fix vs. veränderlich; *Genese des Wissens*: c) Begründung über verschiedene Quellen sowie d) Herkunft des Wissens) und nimmt innerhalb dieser Teildimensionen Kontinua an. Das zweite System folgt einer Entwicklungsperspektive (4.3.3.2) und beschreibt in seinem Ursprung eine Drei-Phasen-Logik (Absolutismus, Multiplismus, Evaluatismus), in der stark vereinfachende Überzeugungen komplexeren Überzeugungen weichen, welche umfassende evaluative Bestandteile beinhalten.

Die epistemischen Überzeugungen sind demnach ein mehrdimensionales Konstrukt, es gibt also nicht eine einzige Form von wissensbezogenen Überzeugungen, und es existieren unterschiedliche Modellierungs- und Taxonomieansätze. Doch wie lässt sich dies auf das Sourcing beziehen? Mehrere theoretische Modelle aus der Multiple-Documents-Literacy-Forschung nehmen an, dass solche epistemischen Überzeugungen je nach Ausprägung als günstig bzw. ungünstig für das Sourcing gelten (Bråten, Britt et al., 2011; List & Alexander, 2019). Dabei unterscheiden sich Modelle darin, wie sie die epistemischen Überzeugungen in ihren Effekten auf das Sourcing beschreiben. List und Alexander (2019) konzeptualisieren die epistemischen Überzeugungen als eher distal, indem sie sie als Einflussfaktoren auf die Grundhaltungen von Personen beim Umgang mit multiplen Dokumenten behandeln. Je nach Ausprägung bei den Überzeugungen

primär bei der Genese des Wissens weisen Personen eine grundsätzliche Affinität auf, Sourcing oberflächen- oder tiefenorientiert zu betreiben, was sich in den entsprechenden Prozessen und Produkten negativ bzw. positiv niederschlägt. Demgegenüber verorten Bräten, Britt et al. (2011) die epistemischen Überzeugungen proximaler bei den konkret ablaufenden Prozessen und Produkten beim Aufbau eines Dokumentenmodells. Sie beschreiben, welche Ausprägung welcher Teildimension günstiger bzw. weniger günstig ist und akzentuieren ebenfalls die Hauptdimension der Genese des Wissens und die Teildimension der Bestimmtheit des Wissens stärker. Zusammengefasst erfolgt in beiden Zugängen dann mehr Sourcing, wenn Personen die Veränderlichkeit von Wissen annehmen (und darum qua Sourcing die verschiedenen Perspektiven eruieren wollen) und sie Wissen auf der Basis von Expert:innen-Aussagen und epistemisch belastbaren Begründungen konzeptualisieren.

Die Empirie dazu ist in der Tabelle 14 zusammengestellt, in ihr sind die Ergebnisse aus den Studien mit den dimensional Systemen zuoberst dargestellt, es folgen – in geringer Zahl – die Befunde aus den Studien mit Skalen aus der Entwicklungsmodell-Logik. Die Essenz der Studien lautet:

- *Epistemische Überzeugungen, die auf die Akzeptanz von Komplexität verweisen, sind tendenziell mit besseren Leistungen assoziiert.* Dies schlägt sich bei den Studien aus der *Tradition der dimensional Zugänge* in der Tabelle 14 darin nieder, dass Personen mit der Anerkennung von Wissen als komplexem Gegenstand und der Abstützungsnotwendigkeit auf multiple Quellen und epistemische Autoritäten ein stärker selektives Vorgehen in der Zuwendung, Inspektion und kritischen Evaluation von unterschiedlich glaubwürdigen Dokumenten in den Prozessen demonstrieren. Dazu passt die Tendenz, bessere Leistungen in umfassenden Tests zu erbringen, die evaluative Leistungen zum Gegenstand hatten. Bei den Studien mit Skalen aus der *entwicklungsbezogenen Tradition* ist neben einem Mangel an Empirie zu konstatieren, dass sich Unterschiede teils nur bei einzelnen Skalen zeigten. So waren Personen mit hohen Ausprägungen bei einem evaluativistischen Verständnis zu besseren Leistungen bei dem sogenannten Verstehen der Autor:innenperspektive in der Lage, während hohe Ausprägungen in Absolutismus und Multiplismus hinderlich waren (Barzilai & Eshet-Alkalai, 2015). Partiiell negative Zusammenhänge ergaben sich zudem bei den Prozessen, nämlich bei der Skala Multiplismus (Barzilai et al., 2015). Das bedeutet: Personen, die Wissen als Nebeneinander von personellen Perspektiven konzeptualisieren (Multiplismus) oder aber als dogmatisch starr erachten (Absolutismus), sind unterlegen. Demgegenüber sind – nur bezogen auf die Produkte – Personen in ihrer Performanz besser, die einen analytischen Zugang zu Perspektiven und deren Berechtigung aufweisen (Evaluatismus).

Tabelle 14: Zusammenhänge von Sourcingmaßen und epistemischen Überzeugungen

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativ-effekte
Metadaten beachten bzw. benennen	Produkte: – Prozesse • <i>Allgemeine internetbezogene epistemische Überzeugungen:</i> Lesezeiten von objektiven Websites (Kammerer et al., 2015) • <i>Komplexität des Wissens:</i> Fixationszeiten bei Dokument mit mittlerer Glaubwürdigkeit (Mason, Pluchino & Ariasi, 2014) • <i>Begründung des Wissens über eigene Person:</i> Lesezeiten von kommerziellen Websites (Kammerer et al., 2015) • <i>Begründung des Wissens über multiple Quellen:</i> Lesezeiten von objektiven Websites (Kammerer et al., 2015)	Produkte: – Prozesse • <i>Allgemeine internetbezogene epistemische Überzeugungen:</i> Lesezeiten von kommerziellen und subjektiven Websites (Kammerer et al., 2015) • <i>Begründung des Wissens über eigene Person:</i> Lesezeiten von objektiven und subjektiven Websites (Kammerer et al., 2015) • <i>Begründung des Wissens über multiple Quellen:</i> Lesezeiten von kommerziellen und subjektiven Websites (Kammerer et al., 2015)
Metadaten nutzen	• Produkte: <i>allgemeine internetbezogene epistemische Überzeugungen:</i> Auswahl objektiver Websites (Kammerer & Gerjets, 2012a) • Prozesse: –	Produkte: – Prozesse • <i>Allgemeine internetbezogene epistemische Überzeugungen:</i> Dauer der Fixation von Metadaten (Kammerer et al., 2013), Anzahl evaluative Kommentare (Kammerer et al., 2013) • <i>Begründung des Wissens (generisch):</i> Dauer Fixation Metadaten (Kammerer et al., 2013) • <i>Absolutismus:</i> Repräsentation von Metadaten, Quelle-Inhalt-Verbindungen und Quelle-Quelle-Verbindungen in Lautdenk-Protokollen (Barzilai et al., 2015) • <i>Multiplismus:</i> Repräsentation von Metadaten, Quelle-Inhalt-Verbindungen (negative Korrelation) und Quelle-Quelle-Verbindungen (negative Korrelation) in Lautdenk-Protokollen (Barzilai et al., 2015) • <i>Evaluatismus:</i> Repräsentation von Metadaten, Quelle-Inhalt-Verbindungen und Quelle-Quelle-Verbindungen in Lautdenk-Protokollen (Barzilai et al., 2015)

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativ-effekte
Metadaten evaluieren	Produkte • <i>Einfachheit des Wissens</i>: Beurteilung der Vertrauenswürdigkeit von Informationen aus zwei Dokumenten (Strømsø et al., 2011, partiell) • <i>Begründung des Wissens</i>: Beurteilung der Vertrauenswürdigkeit von Informationen aus zwei Dokumenten sowie – partiell – Wichtigkeit der Einschätzungskriterien für Glaubwürdigkeitsurteile bei zwei Dokumenten (Strømsø et al., 2011) • <i>Begründung des Wissens über epistemische Autoritäten</i>: Gesamtwert für Evaluation von Websites (Hämäläinen et al., 2021) und Elaborationsgrad in den Begründungen (Kiili et al., 2022) • <i>Begründung des Wissens über multiple Quellen</i>: Gesamtwert für Evaluation von Websites (Hämäläinen et al., 2021) und Elaborationsgrad in den Begründungen (Kiili et al., 2022) • <i>Ursprung des Wissens</i>: Beurteilung der Vertrauenswürdigkeit von Informationen aus zwei Dokumenten und Wichtigkeit der Einschätzungskriterien für Glaubwürdigkeitsurteile bei zwei Dokumenten (Strømsø et al., 2011) • <i>Evaluatismus</i>: Quellenevaluation (latentes Maß, Barzilai & Eshet-Alkalai, 2015) Prozesse • <i>Allgemeine epistemische Überzeugungen</i>: Anzahl evaluativer Kommentare zu Website-Informationen (Ulyshen et al., 2015) • <i>Begründung des Wissens über epistemische Autoritäten</i>: Anzahl evaluative Kommentare (Kammerer et al., 2021)	Produkte • <i>Einfachheit des Wissens</i>: Wichtigkeit der Einschätzungskriterien für Glaubwürdigkeitsurteile bei zwei Dokumenten (Strømsø et al., 2011) • <i>Bestimmtheit des Wissens</i>: Beurteilung der Vertrauenswürdigkeit von Informationen aus zwei Dokumenten und Wichtigkeit der Einschätzungskriterien für Glaubwürdigkeitsurteile bei zwei Dokumenten (Strømsø et al., 2011) • <i>Begründung des Wissens über eigene Person</i>: Gesamtwert für Evaluation von Websites (Hämäläinen et al., 2021) und Elaborationsgrad in den Begründungen (Kiili et al., 2022) • <i>Absolutismus</i>: Quellenevaluation (latentes Maß, negativer Prädiktor, Barzilai & Eshet-Alkalai, 2015), Glaubwürdigkeitsurteile ganzer Dokumente und Wichtigkeit einzelner Kriterien für Glaubwürdigkeitsurteile (Pieschl & Sivyver, 2021) • <i>Multiplismus</i>: Quellenevaluation (latentes Maß, negativer Prädiktor, Barzilai & Eshet-Alkalai, 2015), Glaubwürdigkeitsurteile ganzer Dokumente und Wichtigkeit einzelner Kriterien für Glaubwürdigkeitsurteile (Pieschl & Sivyver, 2021) • <i>Evaluatismus</i>: Glaubwürdigkeitsurteile ganzer Dokumente und Wichtigkeit einzelner Kriterien für Glaubwürdigkeitsurteile (Pieschl & Sivyver, 2021) Prozesse • <i>Allgemeine internetbezogene epistemische Überzeugungen</i>: Anzahl evaluativer Kommentare (Kammerer et al., 2013) • <i>Begründung des Wissens (generisch)</i>: Anzahl evaluativer Kommentare (Kammerer et al., 2013) • <i>Begründung des Wissens über eigene Person</i>: Anzahl evaluativer Kommentare (Kammerer et al., 2021) • <i>Begründung des Wissens über multiple Quellen</i>: Anzahl evaluativer Kommentare (Kammerer et al., 2021)

- *Die Zusammenhänge wirken komplexer, wenn Analysemethoden und Bezugspunkte in den Blick geraten.* In den Studien wurden hin und wieder in Strukturgleichungs- (Barzilai & Eshet-Alkalai, 2015; Härmäläinen et al., 2021) sowie Regressionsanalysen (Kammerer et al., 2015; Kammerer et al., 2021; Kiili et al., 2022; Strømsø et al., 2011) verschiedene Facetten der epistemischen Überzeugungen simultan als Prädiktoren verwendet, wodurch Nettoeffekte einzelner Teildimensionen besser abschätzbar sind. In diesen Studien erweisen sich nur einzelne Teildimensionen als prädiktiv. Ähnliches gilt für jene Studien, in denen einzelne Dokumente (Strømsø et al., 2011) oder Dokumentengruppen (Kammerer et al., 2015; Kammerer & Gerjets, 2012a) gesondert betrachtet wurden. Anmarkrud et al. (2022) berichteten in diesem Zusammenhang davon, dass sie in den heterogenen Befunden in ihrem Überblick feststellten, dass themenspezifische Operationalisierungen der epistemischen Überzeugungen häufiger in erwarteter Weise mit Sourcingmaßen korrelierten als eher abstrakter erfasste epistemische Überzeugungen. Möglicherweise sind also insgesamt beim Sourcing differenzielle und komplexere Muster anzunehmen (und theoretisch zu plausibilisieren), als es die Studienlage und die bislang vorgelegten Analysen derzeit nahelegen. Hierfür benötigt es neben weiterer Theoriearbeit auch vertiefender Studien mit Moderations- und/oder Mediationsanalysen, um die direkten und indirekten Effekte von Überzeugungen noch besser zu verstehen.

Die Zusammenhänge von epistemischen Überzeugungen sind also insgesamt von zweierlei gekennzeichnet: Es mangelt zum einen noch an Studien und Studienergebnissen, die sich zu allgemeinen konsistenten Befundmustern aggregieren lassen. Dies ist ein empirisches Desiderat. Zum anderen fehlen Theorien bzw. Modelle mit spezifischen Wirkannahmen und Wirkmechanismen.

3.3.5 Lesefähigkeiten (Produktperspektiven)

Sourcing wird in verschiedenen Modellen des Lesens als ein Prozessverbund betrachtet, der dabei hilft, mit widersprüchlichen Informationen oder bei wenig Vorwissen im Umgang mit epistemisch unsicheren Informationen zu kohärenten Repräsentationen bzw. zu begründeten Auswahlen zu gelangen (Philipp, 2023). Darum lässt sich eine Nähe von Verstehensprozessen und Sourcing annehmen, was sich ebenfalls in dem Referat an Theorien bei Anmarkrud et al. (2022) niederschlägt. Wie in dem systematischen Forschungsüberblick von Anmarkrud

et al. (2022) zeigt sich ebenfalls in den 20 Studien aus Tabelle 15 kein klares Muster bei den primär auf Produktmaßen basierenden Leseleistungen. Anders gesagt: Dass Personen ein besseres oder schlechteres Leseverstehen, bessere oder weniger gute Inferenzleistungen oder eine höhere bzw. geringere Leseflüssigkeit aufweisen, geht mit Blick auf die reine Verteilung der Studien nicht systematisch mit besserem Sourcing einher.

Allerdings haben Anmarkrud et al. (2022) nach vertiefenden Analysen in ihrem Set aus 16 Studien festgehalten, dass positive Zusammenhänge von Leseleistungen und Sourcing eher dann bestehen, wenn die Lesetests umfassendere Leseverstehensleistungen erfragten, also die Wort- und teils die propositionale Repräsentationsebene gezielt hinter sich ließen, um Situationsmodelle zu fokussieren (s. dazu Teilkapitel [4.1.1.1](#)). Je kognitiv anspruchsvoller die zu erbringenden Leistungen im Leseverstehen und je fähiger Personen diesbezüglich sind, desto besser sind Sourcingleistungen. Damit stellt sich die bislang in der Theorie nur unzureichend geklärte Frage, ob es Schwellenwerte gibt oder ob sich differenzielle Effekte unterschiedlicher Leseverstehensleistungen begründbar postulieren lassen. Diese Frage, die sich beispielsweise beim Vorwissen mit Blick auf schwächere und stärkere Leserinnen und Leser durchaus beantworten lässt (Smith et al., 2021), ist für den Nexus von Lesefähigkeiten und Sourcing noch unausgelotet.

Für die Effekte von unterschiedlichen Lesefähigkeiten auf das Sourcing sprechen Studienbefunde, die auf mehreren Lesetests basieren. Wenn nämlich die Lesefähigkeiten über multiple Tests (jenseits der zeitökonomisch durchführbaren basalen Lesetests) erfasst und statistisch in Bezug zu den Sourcingleistungen gesetzt werden, ist eine Prüfung einzelner Prädiktoren möglich. Tatsächlich ergeben sich aus derartigen Studien erste Hinweise, zum Beispiel zu differenziellen Beiträgen in Regressionsanalysen: So wurde etwa das proaktive Sourcing nur vom Leseverstehen vorhergesagt, aber nicht von Leseflüssigkeit (Hahnel et al., 2021). Das betraf ebenfalls die Fähigkeit, Metadaten zu beurteilen, in der Studie von Kiili et al. (in Druck). Ähnliches wurde teils ebenso bei Korrelationsanalysen (Dyoniziak et al., 2023) festgestellt. Auffällig ist außerdem an den Ergebnissen aus Tabelle 15, dass bei den umfassenden Tests zum Sourcing (erster Eintrag bei der Evaluation von Metadaten) eine markante Klumpung positiver Resultate besteht. All dies unterstreicht, dass Lesefähigkeiten vermutlich komplexer mit den Sourcingleistungen zusammenhängen, als es die heterogenen Befunde insinuierten.

Tabelle 15: Zusammenhänge von Sourcingmaßen und Lesefähigkeiten

Bereich des Sourcings	Positive statistische Effekte	Statistische Null- oder Negativeffekte
Metadaten beachten bzw. benennen	Produkte • <i>Erinnerung an Metadaten</i> (Kammerer, Meier & Stahl, 2016; Salmerón, Gil & Bräten, 2018a; Salmerón, Sampietro & Delgado, 2020) • <i>Korrekte Antworten zur Identifikation von Metadaten</i> als Teil eines umfassenderen Tests (Kiili et al., in Druck: Leseverstehen und Worterkennung) • <i>Korrektes Benennen unter Verfügbarkeit der Dokumente</i> (Paul et al., 2018, Studie 1; Potocki et al., 2020) Prozesse • <i>Konsultation von Metadaten</i> (Merk et al., 2017) • <i>Proaktives Sourcing</i> (Leseverstehen: Hahnel et al., 2021)	• Produkte: <i>Erinnerung an Metadaten</i> (Merk et al., 2017) • Prozesse: <i>proaktives Sourcing</i> (Leseflüssigkeit: Hahnel et al., 2021)
Metadaten nutzen	• Produkte: <i>Zitationen</i> (Anzahl: Kammerer & Gerjets, 2014a; binär kodiert: Potocki et al., 2020) • Prozesse: –	• Produkte: <i>Zitationen</i> (Anzahl: Salmerón, Sampietro & Delgado, 2020; binär kodiert: Macedo-Rouet, Salmerón et al., 2019) • Prozesse: –
Metadaten evaluieren	Produkte • <i>Gesamtscore im Evaluationstest</i> (Dyoniziak et al., 2023: Leseverstehen; Forzani, 2018: Regressionsanalyse; Hahnel et al., 2016, 2018; Hämmäläinen et al., 2021; Macedo-Rouet et al., 2013, Experiment 1) • <i>Beurteilung der Dokumentenglaubwürdigkeit</i> (Kammerer, Meier & Stahl, 2016, nur bei einem Teil der Dokumente; Kobayashi, 2014) • <i>Diskriminierung vertrauenswürdiger Websites bei der Verwendbarkeit</i> (Macedo-Rouet, Salmerón et al., 2019) • <i>Beurteilung von einzelnen Facetten der Glaubwürdigkeit verschiedener Quellen</i> (Kanninen et al., 2019, Leseverstehen), speziell bei der <i>Expertise</i> (Paul et al., 2018, Studie 1; Potocki et al., 2020) • <i>Korrekte Antworten zur metadatenbasierten Begründung</i> als Teil eines umfassenderen Tests (Kiili et al., in Druck: Leseverstehen) Prozesse: –	Produkte • <i>Gesamtergebnis Evaluation von Websites</i> (Dyoniziak et al., 2023: Dekodieren und Inferenzen) • <i>Rangreihung von Dokumenten nach Glaubwürdigkeit</i> (Braasch, Bräten, Strømsø & Anmarkrud, 2014; Mason, Scrimin, Tornatora et al., 2018) • <i>Beurteilung der Verwendbarkeit einzelner, definierter Websites</i> (Macedo-Rouet, Salmerón et al., 2019) • <i>Anzahl Begründungen für Auswahl von Dokumenten</i> (Bräten, Brante & Strømsø, 2018) • <i>Einschätzung von Metadaten</i> (Macedo-Rouet, Salmerón et al., 2019) • <i>Begründungen für Rangreihungen</i> (Braasch, Bräten, Strømsø & Anmarkrud, 2014 – als Gesamtwert: Kiili et al., 2021) • <i>Beurteilung von einzelnen Facetten der Glaubwürdigkeit verschiedener Quellen</i> (Kanninen et al., 2019, Leseflüssigkeit) • <i>Korrekte Antworten zur metadatenbasierten Begründung</i> als Teil eines umfassenderen Tests (Kiili et al., in Druck: Worterkennung) Prozesse: –

3.4 Probleme

Das Sourcing ist ein Prozessverbund, der nicht nur aus drei einzelnen Prozessen besteht, sondern innerhalb dieser Prozesse selbst sind noch Binnendifferenzierungen möglich, wie es die Darstellung insbesondere im Teilkapitel [3.2](#) impliziert. Allein aufgrund der großen Zahl der Prozesse, die zudem noch interagieren und von der lesenden Person gesteuert werden müssen, sind Problemlagen erwartbar (Bråten et al., 2017). Tatsächlich ergeben sich aus Studien in Bezug auf das Sourcing gleich mehrere gut dokumentierte Probleme, nämlich dergestalt dass Leserinnen und Leser Metadaten erstens nicht finden, sie zweitens falsch nutzen, drittens nicht mit Informationen aus einzelnen oder multiplen Dokumenten in Verbindung bringen und viertens nicht evaluieren (s. dazu im Überblick Philipp, 2020a, S. 206–216).

In diesem Teilkapitel werden aus Gründen der Redundanzvermeidung mit bereits publizierten Mustern nur zwei ausgewählte Problemfelder vorgestellt. Diese Problemfelder sind komplementär. Das erste betrifft generell ungenügende Leistungen in der metadatenbasierten adäquaten Einschätzung ([3.4.1](#)). Das zweite Problemfeld ist eingegrenzter auf das laterale Lesen als ein prototypischer und normativ wünschenswerter Umgang mit Internetdokumenten, den Faktencheckerinnen und Faktenchecker zeigen und der sich fördern lässt (McGrew, 2024). Die Studien zum lateralen Lesen bzw. die diesem Zugang anschlussfähige Empirie machen auf die substanziellen Defizite der Heranwachsenden und Erwachsenen aufmerksam ([3.4.2](#)).

3.4.1 Problemfeld 1: ungenügende Beurteilungen und Beurteilungsleistungen von Metadaten

Das erste Problemfeld, welches sich in einer Vielzahl von Studien niederschlägt, besteht darin, dass es Leserinnen und Lesern schwerfällt, ausreichend treffsicher die Glaubwürdigkeit von Dokumenten einzuschätzen. Dieses Phänomen ist sowohl in Produktmaßen ([3.4.1.1](#)) als auch schon in den vorgelagerten Prozessen erkennbar ([3.4.1.2](#)).

3.4.1.1 Produktperspektiven

Die Studien zu den evaluativen Outcomes haben Befunde produziert, welche sich zu vier Problemen aggregieren lassen:

- *Studien, die Gesamtwerte für anspruchsvolle Sourcingleistung auf einer Skala ausweisen, kommen zu Befunden, die auf geringe bzw. stark ausbaufähige evaluative Leistungen verweisen.* In einigen Studien wurden umfassende

Tests erstellt, in denen diverse Teilleistungen des Sourcings separat erfasst, aber als Gesamtwert berichtet wurden. Solche Tests erfassten das Erkennen und Beurteilen von Metadaten bzw. ganzen Dokumenten (Britt & Aglinskaskas, 2002, Experiment 1) bzw. prüften auf Skalen das Zusammenkommen von verschiedenen Einzelleistungen auf einer Ordinalskala (Marttunen et al., 2021), teils waren die evaluativen Leistungen erst am oberen Ende des Spektrums der Skala enthalten (Kiili et al., 2021). Gemäß dieser Logik der Testwertbildung entsprechen hohe Werte auf der Skala mit einem a priori definierten Spektrum wünschenswert hohen Leistungen. Die Ergebnisse in diesen Studien variieren stark und fallen teils sehr niedrig aus, da die Personen maximal ein Viertel der möglichen Punkte erreichten (Barzilai & Eshet-Alkalai, 2015; Barzilai, Mor-Hagani et al., 2020, Pretest; Britt & Aglinskaskas, 2002, Experiment 1; Dyoniziak et al., 2023; Mason, Scrimin, Tornatora et al., 2018). In anderen Untersuchungen lagen die Mittelwerte höher, nämlich bei einem Drittel (Hahnel et al., 2016) bzw. ca. 40 Prozent (Forzani, 2018; Kiili et al., 2021). In wenigen Fällen erzielten Stichproben aber auch hoch wirkende Werte von rund 60 Prozent (Hämäläinen et al., 2021; Marttunen et al., 2021). Selbst in Studien mit hohen Mittelweltergebnissen der Testpersonen besteht jedoch mit Blick auf die Deskriptoren der Skalen kein Anlass zur Entwarnung, da die vermeintlich hohen Testergebnisse immer noch unvollständigen Evaluationen entsprechen und aus der Sicht des Evaluierens eher rudimentäre Leistungen festgestellt wurden.

- *Die Evaluation von Dokumenten gelingt nicht ausreichend trennscharf bzw. schließt fehl.* Ein Phänomen, das unter diesen Punkt fällt, ist die teils hohe Ungenauigkeit bei der Beurteilung von unterschiedlich glaubwürdigen Websites. Dies zeigt sich darin, dass beispielsweise insbesondere Jugendliche nicht ausreichend dazu in der Lage sind, vertrauenswürdige von wenig vertrauenswürdigen Dokumenten zu unterscheiden (Abel et al., in Druck; Braasch et al., 2009; Kammerer & Gerjets, 2012b, 2014a; Nygren & Guath, 2019). Außerdem sind teilweise die Fehlerquoten in Studien sehr hoch, das heißt: Personen beurteilen fehlerhafte bzw. verzerrte Quellen als belastbarer, als sie es wirklich wären (Lucassen & Schraagen, 2011; Mason, Junyent & Tornatora, 2014; Nygren & Guath, 2019; von der Mühlen et al., 2016). Dazu passt, dass sie Schwierigkeiten haben, gerade die weniger vertrauenswürdigen Quellen kalibriert einzuschätzen (Anttonen et al., in Druck; Kiili, Leu, Marttunen et al., 2018) bzw. diese weniger in ihren eigenen Antworten zu nutzen als solche von epistemisch sichereren Quellen (Kiili, Brante et al., 2020).
- *Die Begründungen für Einschätzungen sind inadäquat bzw. unterkomplexer, als es sachdienlich bzw. nötig wäre.* Dieses Muster ergibt sich in Studien mit differenzierteren Analysen der offenen Antworten. So erfolgen die Begründungen kaum mit Blick auf die Metadaten (Braasch et al., 2013, Kontrollgruppe; Bråten, Strømsø & Andreassen, 2016; Delgado et al., 2019,

Baseline-Testung; Macedo-Rouet, Potocki et al., 2019, Experiment 1; Wiley et al., 2009, Experiment 1) bzw. stützen sich auf eher wenige (Kiili et al., 2022) resp. unverbundene Kriterien (Marttunen et al., 2021) oder fallen gerade bei den in den Ergebnissen fehleranfälligeren Dokumenten stärker inhaltsgetrieben aus, als es angemessen wäre (Kiili, Leu, Marttunen et al., 2018).

- *Fehlleistungen und misslingendes Begründen treten zusammen auf.* Einschätzungen zu einzelnen Metadaten bzw. Dokumenten und die Begründungen sind, zumal in der Kombination beider Variablen, ein empirisch rege beforschter Gegenstand (s. Teilkap. [3.2.1](#) und insbesondere [3.2.4.1](#)). In aller Regel werden dabei die Begründungen mit offenen Antworten erfasst und in variierender Korngröße kodiert. Die Auswertung geht danach auf die Auftretenshäufigkeit einzelner Kategorien ein. Einen anderen Zugang wählte eine Studie mit Jugendlichen der siebten Klassenstufe aus den USA (Coiro et al., 2015). Die Jugendlichen sollten im Rahmen einer Testung unter anderem Fragen zur Expertise der Autorin bzw. des Autors der Website-Inhalte und zur Zuverlässigkeit der Informationen des Website-Texts einschätzen und begründen. In der Auswertung wurden Korrektheit und Begründungen nach Richtigkeit bzw. Inadäquanz ausgewertet. Darum war es möglich, die Herausforderungen und Problemlagen genauer in den Blick zu nehmen und zu systematisieren. In Bezug auf die Korrektheit auf beide Fragen waren über zwei Drittel der Jugendlichen nicht in der Lage, zutreffend Expertise und Zuverlässigkeit einzuschätzen. In den Begründungen fiel auf, dass den Testpersonen belastbare oder klare Kriterien fehlten, sie keine Trennschärfe bei den Kriterien geäußert hatten (indem sie inhaltliche Relevanz mit Expertise gleichsetzten) oder ihre Begründungen unpräzise formulierten. Solche Studien mit genauerer Analyse der Fehlanwendungen zeigen also durchaus häufige und musterhaft auftretende Schwierigkeiten bei der angemessenen Beurteilung von Metadaten und Dokumenten.

Die Probleme beim Evaluieren betreffen damit insgesamt nicht nur die Leistungen beim Einschätzen, sondern auch deren Grundlage in Form der Begründungen. Auf den allgemeinsten Nenner gebracht mangelt es an kalibrierten und trennscharfen Bemessungsgrundlagen, die beim Sourcing mehrdimensional und konzertiert in Anschlag gebracht werden.

3.4.1.2 Prozessperspektiven

In den Studien mit den Prozessmaßen sind an dieser Stelle jene von Interesse, in denen die Häufigkeiten von evaluativen Kommentierungen der Metadaten hinsichtlich Glaubwürdigkeit, Vertrauenswürdigkeit und/oder Expertise in Lautdenk-Protokollen ausgewertet haben. Das konsistent wirkende Resultat dieser Studien lautet: *Personen (Laien und Novizinnen und Novizen) evaluieren*

Metadaten kaum hinsichtlich ihrer Aussagekraft für die Glaubwürdigkeit der Dokumente und Informationen. Das war besonders deutlich in der Studie von Walraven et al. (2009) das Ergebnis, wobei diese Studie ein dezidiert auffälliges Negativbeispiel für evaluative Sourcingprozesse war. Laut den Lautdenk-Protokollen kamen bei den untersuchten Jugendlichen evaluative Prozesse mit Fokus auf die Informationen zum einen und zur Einschätzung der Quelle zum anderen marginal bis überhaupt nicht vor. Doch selbst in Studien, in denen – gemessen von den ohnehin wenigen evaluativen Sourcingaktivitäten im Allgemeinen als Benchmark (Barzilai et al., 2015; Cho, Woodward & Li, 2018; Goldman et al., 2012; Kammerer, Kalbfell & Gerjets, 2016, Experiment 1) – relativ viel Sourcing betrieben worden zu sein scheint (Anmarkrud et al., 2014; Barzilai & Zohar, 2012; Brand-Gruwel et al., 2017; Cho, Han & Kucan, 2018; Gerjets et al., 2011; Jay, 2021; Kammerer et al., 2013, 2021; Kiili et al., 2008; Rott & Gavin, 2015; Strømsø et al., 2013; Wineburg, 1991), beschränkt sich das Sourcing auf wenige Male pro Person bei der Aufgabenbearbeitung.

Natürlich ist eine alleinige Quantifizierung der Auftretenshäufigkeit evaluativer Prozesse noch nicht dazu geeignet, ein Problemfeld zu benennen, zumal die Häufigkeit allein noch nichts über die Eignung, Adäquanz oder den Erfolg aussagt. Gleichwohl ist die über diverse Studien hinweg festgestellte empirisch geringe Beobachtbarkeit bei den evaluativen Prozessen ein deutliches Signal dafür, dass diese Facette des Sourcing im Ensemble der Verstehensprozesse unterrepräsentiert ist. Prozessdaten aus dem Kontext des lateralen Lesens unterstreichen jedoch, dass Personen eher oberflächliche Einschätzungen von Metadaten, Quellen und Dokumenten vornehmen (s. Teilkap. [3.4.2.2](#)).

3.4.2 Problemfeld 2: schwache Heuristiken bei der Beurteilung

Aus dem Kontext jener Studien zum ökologisch validen (lateralen) Lesen im Internet – d. h. ohne Simulation von Suchmaschinen oder fix vorgegebene Dokumentensets – stammen kritische Stimmen, welche monieren, dass ein Teil des alltäglichen Lesens und der Ressourcen des Lesens im Internet exkludiert bleiben (s. dazu Teilkap. [3.1.3.3](#)). Denn indem Recherchemöglichkeiten über die Quellen von Aussagen verwehrt bleiben und Personen sich auf fehleranfällige Urteile nur anhand der experimentell vorgegebenen Metadaten verlassen können, besteht allein durch den Versuchsaufbau der Experimente die Gefahr, sogenannte „schwache Heuristiken“ zu bemühen (McGrew, 2021a). Diesen Ausdruck verwendet McGrew (2021a) in dezidiert Abgrenzung zu den effektiven Vorgehensweisen von Faktencheckerinnen und Faktencheckern in ökologisch validen Settings (McGrew, 2021b, 2022; Wineburg & McGrew, 2019), um oberflächliche und ungeeignete Vorgehensweisen zu benennen. Die Bezugspunkte der Beurteilungen werden

dabei von McGrew (2021a) auch im Hinblick auf den problematischen Einsatz statischer Checklisten (Breakstone et al., 2018) problematisiert und lassen sich zudem als besser bewusst zu ignorierende Informationen werten (Kozyreva et al., 2023). Der Grund dafür ist, dass die Einschätzbarkeit von Websites fehleranfällig ist (s. Teilkap. [3.1.3.1](#)), da die Beurteilungsgrundlage

- entweder leicht von den Urheberinnen und Urhebern manipulierbar ist (wie das Aussehen der Websites, die Gestaltung des Textes oder die Präsentation in „Über uns“-Texten) – zumal in Zeiten fortschreitender KI-Einsatzgebiete,
- keine direkte Beurteilung der Vertrauenswürdigkeit (wie die URLs oder Social-Media-Engagements wie Likes, Bewertungen, Follower) zulässt
- oder aber ohne hohes Vorwissen kaum adäquat gelingen kann (Inhalt der Texte).

Vorgehensweisen, deren Bezeichnung als „schwache Heuristik“ (McGrew, 2021a), „Noviz:innenpraktiken“ (Kohnen et al., 2021) oder als „vertikales Lesen“ (McGrew & Chinoy, 2022) bezeichnet werden, eint ein eher oberflächenorientiertes Vorgehen. Dieses Vorgehen ist ein Phänomen, welches sich in einer Vielzahl von Studien in Prozess- und Produktmaßen niederschlägt, auf welche die weitere Darstellung nun eingeht.

3.4.2.1 Produktperspektiven

Die Studienlage bei den Produkten, bei denen regelmäßig die Einschätzungen von Websites mit in der Regel offenen Aufgaben erfasst und inhaltsanalytisch kodiert wurden, ergibt sich aus drei unterschiedlichen Gruppen von Studien, von denen die ersten beiden unmittelbar auf das laterale Lesen beziehbar sind und die dritte nur mittelbar:

- *Große Anteile von Personen mit geringen Testleistungen in Studien zum lateralen Lesen:* Studien aus der sozialwissenschaftlich orientierten Grundlagenforschung zeigen immer wieder, dass große Teile von unterschiedlichen Personengruppen – Sekundarschuljugendliche (Breakstone, Smith, Wineburg et al., 2021; McGrew et al., 2018), aber auch Studierende (Breakstone et al., 2022; Dissen et al., 2022; McGrew, 2022; McGrew et al., 2018, 2019; Wineburg & McGrew, 2019) – zu großen Teilen *fehlerhafte oder unzureichende Urteile* abgeben, die sich mit schwachen Heuristiken direkt oder indirekt in Verbindung bringen lassen. Ein besonders deutliches und auf einer großen Stichprobe basierendes Exempel hierfür ist im Kasten „Illustrierendes Beispiel 15“ enthalten.

- *Hohe Ausgangswerte bei und hohe Stabilität von schwachen Heuristiken in Leistungsdaten bei Studien der angewandten Grundlagenforschung:* Selbst aus *Leistungsdaten aus Interventionsstudien zum lateralen Lesen* (Breakstone et al., 2024; Brodsky, Brooks, Scimeca, Todorova et al., 2021; Kohnen et al., 2020; McGrew, 2020; McGrew & Breakstone, 2023; McGrew & Byrne, 2021; McGrew et al., 2019; Wineburg et al., 2022) geht hervor, dass die schwachen Heuristiken persistent auftreten. So sind mitunter relativ große Anteile der *geförderten* Personen aus den Interventionsgruppen gemäß ihren Antworten selbst nach absolvierter Förderung nicht dazu in der Lage, tiefergehende Prüfungen mithilfe weiterer Internetdokumente und -recherchen in ihren Antworten zu realisieren.
- *Geringe Anteile von Begründungen mit Hinweis auf tiefergehende und vergleichende Zugänge bei der Quellenevaluation:* Diese Ergebnisse sind bereits an anderem Ort in diesem Kapitel Gegenstand gewesen: Aus der Perspektive schwacher Heuristiken wirken bereits die Gewichtungen von *Begründungen* von Urteilen aus Studien, die in Teilkapitel [3.2.4.1](#) und dort in der Tabelle 11 versammelt wurden, problematisch. Denn das tiefere Prüfen und eine Analyse des Inhalts bzw. das Verwenden anderer Dokumente für die Urteile ist dort – wenn überhaupt – nur selten vorgekommen. Zu diesem Ergebnis kommen andere Studien ebenfalls, in denen Begründungen nach dem kognitiven Anspruchsgrad analysiert worden und denen zufolge die Begründungen eher oberflächlich erfolgten (Artmann et al., 2023, Kontrollgruppe) oder nur selten auf textvergleichenden Analysen fußten (Brand-Gruwel et al., 2017; Hämäläinen et al., 2021; Kiili, Leu, Marttunen et al., 2018; Mason, Moè et al., 2022; Mason, Scrimin, Tornatora et al., 2018).

Leistungsdaten, die unmittelbar oder mittelbar auf die Fähigkeit verweisen, Dokumente mittels anderer recherchierter Dokumente zu beurteilen, verdeutlichen konsistent, dass ohne und teils selbst mit Förderung diese Leistungen von einem (großen) Teil der Testpersonen nicht realisiert werden.

Illustrierendes Beispiel 15

Die knappe Ein-Prozent-Marke: Jugendliche und ihre Fähigkeiten, Internetinformationen kritisch zu evaluieren

A) Was war das Erkenntnisinteresse?

Die Studie verfolgte das Ziel, US-weit Aussagen darüber zu treffen, wie fähig Jugendliche darin sind, Internetdokumente zu bewerten (Breakstone, Smith, Wineburg et al., 2021). Diese Bewertung wurde in der Studie als „Civic Online Reasoning“ bezeichnet und inkludiert die Fähigkeiten, effektiv nach sozialen und politischen Informationen im Internet zu suchen, diese zu bewerten und zu verifizieren.

B) Wer wurde untersucht?

Die Stichprobe bestand aus 3.446 US-amerikanischen Jugendlichen der Klassenstufen 9–12. Bei diesem repräsentativen Sample, das aus einer größeren Stichprobe stammte, achteten die Autorinnen und Autoren darauf, dass es in seiner Zusammensetzung so weit wie möglich der nationalen, aber auch regionalen Zusammensetzung der Schülerinnen und Schüler entsprach, also möglichst gut die Grundgesamtheit repräsentierte.

C) Welche Instrumente kamen zum Einsatz?

Das Instrument bestand aus sechs Aufgaben, welche die Jugendlichen mit einem Computer lösten, der mit dem Internet verbunden war. Diese Aufgaben machten es z. B. erforderlich, die Glaubwürdigkeit von lobbyierten Internetseiten, von Postings in Social Media zum Thema Wahlbetrug, die Eignung von Websites als Startpunkt zum Thema Waffenkontrolle etc. zu beurteilen und schriftlich zu begründen. Die Antworten wurden kodiert und jeweils einer Kategorie zugeordnet: a) „Fehlantworten“ (falsche oder irrelevante Antworten), b) „tendenziell richtige Antworten“ (partiell inkorrekte Antworten, darunter solche mit unvollständigen Begründungen) und c) „korrekte Antworten“ (korrekte Beurteilungen, Analysen und Vergleiche von Websites). Für den Großteil der untersuchten Jugendlichen wurde dann ein Mittelwert über alle Antworten gebildet, der es zulässt, aus den numerischen Werten die prozentualen Anteile auszuweisen, welche Leistungen die Personen erbracht hatten.

D) Was waren die Hauptbefunde?

Die Zugehörigkeit der Jugendlichen zu drei Gruppen gemäß den o.g. Antwortkategorien enthält das Kreisdiagramm in Abbildung 9. Es unterstreicht, dass nur ein äußerst geringer Teil der Jugendlichen die Höchstleistungen erbracht hatte, nämlich in absoluten Zahlen 23 Personen bzw. prozentual: 0,7 Prozent. Weitere 13 Prozent demonstrierten Leistungen im Mittelfeld der Skala, aber fünf von sechs Jugendlichen (86 Prozent) wiesen problematische Antwortmuster auf. Diese überwiegende Mehrheit konnte die Aufgaben nicht korrekt lösen. Der rechte Teil des Diagramms schlüsselt dies nochmals auf und hat das Ergebnis, dass mit 58 Prozent mehr als jede zweite Person bei keiner einzigen Aufgabe eine auch nur tendenziell korrekte Antwort gegeben hatte. Abgesehen davon erbrachten Regressionsanalysen zusätzliche Befunde. Nach diesen Ergebnissen hatten jene Personen, die jünger

und schulisch leistungsschwächer waren, sowie Heranwachsende aus sozial schwächer gestellten Familien (geringerer Bildungsabschluss der Mütter und Teilnahmeberechtigung an vergünstigter Schulverpflegung) schlechtere Leistungen.

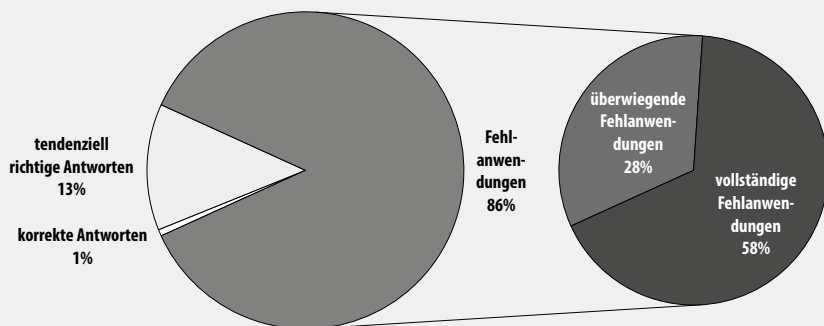


Abbildung 9: Verteilung der Probandinnen und Probanden zu Leistungsgruppen (eigene Darstellung, basierend auf Breakstone, Smith, Wineburg et al., 2021, Appendix F, N = 3.402)

E) Diskussion

Das Verdienst der Studie besteht in einer großen Stichprobe mit einem deutlichen Bemühen, die Grundgesamtheit möglichst adäquat zu repräsentieren. Außerdem ist eine Stärke darin zu sehen, dass die Untersuchung mittels offener Aufgaben und hoher ökologischer Validität das Konstrukt „Civic Online Reasoning“ erfasste. Vor diesem Hintergrund sind die Ergebnisse, dass knapp drei von fünf Jugendlichen in den Aufgaben durchgängig schlechte Leistungen erzielten, als kritisch einzustufen, da diese Personen einen unkritischen, naiv wirkenden Umgang mit verschiedenen Informationen und Informationsquellen im Internet demonstriert hatten. Nur einem Bruchteil gelang es hingegen, die Aufgaben bravourös zu absolvieren.

Hinzu kommt, dass diese Studie wie so viele vor ihr in der Domäne Lesen den Effekt von individuellen Merkmalen wie Alter und schulischen Leistungen (operationalisiert über Noten) und vor allem des familialen Hintergrunds feststellte. Anders gesagt: Die Fähigkeiten zum versierten Umgang mit gesellschaftswissenschaftlich relevanten Informationen hängen von der Herkunft ab, und diese Fähigkeiten sind außerdem nur schwach entwickelt.

3.4.2.2 Prozessperspektiven

Die Studienergebnisse zu den Prozessen lassen sich zu einem Muster abstrahieren: *Laterales Lesen ist bei Leserinnen und Lesern ohne spezifische Schulung ein höchstens selten vorkommendes Vorgehen, um Websites einzuschätzen.* Dies ergibt sich aus Prozessdaten in der Grundlagenforschung (Kohnen et al., 2021; McGrew, 2021b, 2022; Wineburg & McGrew, 2019) ebenso wie aus den Daten in Kontrollgruppen bzw. Pretestdaten in Interventionsstudien (Breakstone, Smith, Connors et al., 2021; Brodsky et al., 2023; Brodsky, Brooks, Scimeca, Galati et al.,

2021; McGrew, 2020; McGrew & Chinoy, 2022; Panizza et al., 2022, Experimente 1 und 2; Walsh-Moorman & Pytash, 2022). In diesen Studien ist das Vorkommen des lateralen Lesens entweder nur bei einem geringen Prozentsatz der Testpersonen oder sogar gar nicht beobachtet worden, die Auftretenshäufigkeit lag bei rund einem Zehntel der Fälle, maximal bei einem knappen Fünftel (McGrew & Chinoy, 2022). Ein empirisch besonders eindrucksvolles Beispiel hierfür ist die Studie von Brodsky et al. (2023). In wurde anhand von schriftlichen Antworten von über 2.200 kanadischen Jugendlichen der Klassenstufen 7–12 rekonstruiert, dass bei den vier gestellten Aufgaben zwischen 9 bis 15 Prozent der Personen von einer Vorgehensweise berichteten, die auf ein laterales Lesen Rückschluss gab.

Hinzu kommt, dass selbst dort, wo aktiv Strategien des lateralen Lesens in Interventionsstudien vermittelt wurden, laut Unterrichtsbeobachtungen (McGrew, 2021a; McGrew & Byrne, 2022; McGrew et al., in Druck) und anderen Prozessindikatoren (Brodsky et al., 2023; Brodsky, Brooks, Scimeca, Galati et al., 2021; Kohnen et al., 2020; McGrew, 2020) die Persistenz der schwachen Heuristiken besteht. Während der Förderung bzw. im Anschluss daran glückt es also geförderten Personen nicht ausreichend, die zu lernenden Strategien angemessen anzuwenden. Dieses Muster spiegelt wider, was bereits bei den Produktmaßen aus dem Teilkapitel [3.4.2.1](#) berichtet wurde.

3.5 Fazit

Das Sourcing ist ein komplexer und mehrdimensionaler Prozess, dessen Bezugspunkt Metadaten sind. Sourcing bezeichnet das Erkennen, Verwenden und Beurteilen von Metadaten. Solche Informationen über die textuellen Informationen, die zusammen aus Metadaten und Text Dokumente formieren, werden in der Leseforschung zusehends als wichtige Bestandteile umfassender Verstehensleistungen, aber auch begründeter Auswahl- und Gewichtungsprozesse erachtet. Der Hintergrund dafür ist, dass über die Verbindung von Aussagen und ihrem Ursprung bzw. Kontext interpretatorische und heuristische Anreicherungen möglich werden, die eine bedeutungsmodifizierende bzw. -erweiternde Funktion haben. Dadurch ist es der lesenden Person idealerweise möglich, über verschiedene Kategorien wie zugeschriebene Expertise, Aktualität, Adäquanz in der Genese von Aussagen, Wohlwollen etc. zu Glaubwürdigkeitsurteilen zu gelangen. Aktuell werden in der Theorie die Expertise einer Quelle und deren Vertrauenswürdigkeit als kleinster gemeinsamer Nenner und als Konstituenten für die übergeordneten Glaubwürdigkeitseinschätzungen angenommen.

Diese Glaubwürdigkeitsurteile – und das muss an dieser Stelle stark betont werden – sind nicht an spezifische Phasen im Gesamtprozess der Auseinandersetzung mit (digitalen) Dokumenten gebunden. Sie können vielmehr vor dem Lesen auftreten (z. B. bei der Analyse von Suchmaschinenergebnissen anhand

der Snippets in der Vorschau), beim Lesen (z. B. bei dem Versuch, Dissensursachen zu attribuieren oder Fake News zu identifizieren) vorkommen und sich zusätzlich nach dem Lesen ereignen, wenn begründete Positionierungen zu einem Sachverhalt nötig werden. Hierin deutet sich die große Bandbreite des Sourcing als Multifunktionsverbund kognitiver Prozesse des Erkennens, Verwendens und Beurteilens von Metadaten an. Dies wird besonders bei einem neueren Phänomen des Sourcing beim Lesen evident: dem lateralen Lesen. Dieses von professionellen Faktencheckerinnen und Faktencheckern genutzte Recherchieren über die Quellen von Aussagen bzw. von Dokumenten mithilfe des Internets beinhaltet diverse Facetten des Sourcing. Das laterale Lesen verweist darauf, dass gezielte Recherchen über und mithilfe diverser Quellen dabei helfen, sich einem ausreichend belastbar wirkenden Glaubwürdigkeitsurteil durch einen dynamischen Prozess anzunähern.

Die experimentelle und Grundlagen-Forschung haben das laterale Lesen weniger stark in den Blick genommen als solche Settings, in denen die Dokumente und Quellen gezielt variiert und als Stimuli bereitgestellt wurden. Die Sourcingmaße als (abhängige) Variablen sind dominant produktbezogen erforscht worden, Studien zu den Prozessen sind im Vergleich rarer. Entlang der drei Teilprozesse des Sourcing – Beachten der Metadaten, Verwenden der Metadaten und Evaluieren der Metadaten – wurden die vorliegenden Studien und deren Befunde systematisiert. Die Studienlage gestaltet sich dabei nicht nur aus der Zuwendung zu den Produkten bzw. den Prozessen ungleich, es ist ferner auffällig, dass die metadatenbasierten Beurteilungsleistungen besonders stark erforscht wurden. Die Ergebnisse seien an dieser Stelle gebündelt:

- *Metadaten beachten bzw. benennen*: Die *Prozessmaße*, die sich in eher unspezifische Verhaltensmaße und spezifischere Benennungen (in Studien mit Lautdenk-Protokollen) untergliedern lassen, zeichnen ein differenziertes Bild. So lassen sich bereits aus den *unspezifischen Indikatoren* verschiedene Funktionen des Erkennens und Benennens anhand der Auswertungsmethoden abstrahieren: ein vorgängiges proaktives Sourcing, ein gezieltes, dem relevanzbasierten Lesen ähnelndes Suchen nach Metadaten, wo sie für die Aufgabenbearbeitung nötig sind, und eine wiederholte Konsultation. Ebenfalls lassen sich über die Sequenz der Konsultation von Metadaten Verhaltensindizes bilden, die auf eine gezielte Zuwendung zu den glaubwürdigeren Dokumenten hinweisen und ein absichtsvoll wirkendes Sourcing implizieren. Solche Sourcingmaße korrespondieren häufig, aber nicht immer mit besseren Leistungen bzw. kamen in Personengruppen mit besseren Testleistungen häufiger vor und sind durchaus als Bestandteil eines gesamthaft positiven Vorgehens wertbar, da sie mit anderen Prozessdaten korrelieren. Die *spezifischeren Verhaltensindikatoren* ergeben ein komplementäres Bild: Sie illustrieren erstens, dass der Ex- und Implizitheitsgrad variiert, Personen

also entweder deutlich einzelne Metadaten und Quellen erkennen oder aber diese ungenau benennen. Zweitens erfolgt das Sourcing je nach Dokument und nach Art des Metadatum unterschiedlich: Leserinnen und Leser gehen demnach nicht uniform vor. Die sich hierin andeutende Sensibilität ist allerdings mit Außenkriterien im Sinne einer externen Validierung noch nicht gut abgesichert. Zum einen fehlen Studien, die andere Verstehensprozesse mit dem Sourcing in Beziehung setzen, zum anderen ist auch der Zusammenhang mit Produktmaßen nicht durchgängig und damit konsistent belegt. Die schon angesprochenen *Produktindikatoren* wurden ebenfalls mit mehreren Zugängen erfasst. Die Produktmaße illustrieren, dass die Frage nach dem Erkennen und Benennen sowie Erinnern nicht losgelöst von der Art der Erfassung beantwortbar ist. Bei gedächtnisbasierten Erfassungen der produktbezogenen Sourcingmaße liegen beispielsweise Hinweise vor, dass geschlossene Tests dazu führen, dass Personen eine auffällig bessere Leistung erbringen. Offene Fragen führen demgegenüber zu deutlich geringeren Quoten in den korrekten Testleistungen.

- *Metadaten nutzen:* Die Verwendung von Metadaten ist ein noch untererforschter Bereich, doch lassen sich anhand der Prozess- und Produktdaten selbst aus der geringen Menge an Studien diverse Funktionen extrahieren: Sie reichen von der Nutzung der Metadaten bei der Auswahl von Dokumenten, dem Vorhersagen der Inhalte, dem Verbinden mit den Aussagen (produktbezogen in Form von Referenzierungen) sowie dem Vergleichen von Perspektiven der Quellen und Dokumente über das Gewichten von Aussagen hin zur (potenziellen) Nutzung der epistemisch belastbar wirkenden Quellen. Dieses Spektrum, welches sich hierin andeutet, unterstreicht die Mannigfaltigkeit des Sourcings in der Kategorie und verdient zukünftig eine systematischere Zuwendung der Empirie.
- *Metadaten evaluieren:* Die Beurteilung der Metadaten ist im Vergleich mit den anderen beiden Prozessen des Sourcings am umfassendsten erforscht. Ähnlich wie bei dem Erkennen und dem Nutzen von Metadaten gibt es beim Evaluieren zwei Prozessgruppen, wobei sich dies eher indirekt aus den Produktmaßen ergibt: das Beurteilen und das Begründen. Die Studien differenzieren hierbei diverse Bezugseinheiten der Beurteilungen (Dokumente, einzelne Informationen und teils einzelne Metadaten) und haben die Wichtigkeit bzw. Häufigkeit einzelner Dimensionen der Begründungen mittels geschlossener und offener Fragen erfasst. Demnach ist das Evaluieren ein mehrdimensionales Mehrebenen-Konstrukt. Die Beurteilung von Dokumenten und Informationen erfolgt präziser an den äußeren Rändern von in ihrer Glaubwürdigkeit divergierenden Sets an Dokumenten, und neutrale Informationen von Quellen mit hoher epistemischer Autorität (z. B. Expertise) werden in der Regel als glaubwürdiger eingeschätzt. Die Begründungen für die Urteile erstrecken sich auf eine Vielzahl von Kriterien, von denen nur ein

Teil die Metadaten touchiert. Inhaltliche und relevanzbezogene Urteile dominieren demgegenüber. Außerdem spielt die Vertrautheit des Gegenstands in den Dokumenten eine Rolle: Sourcing wird dann wichtiger, wenn das eigene Vorwissen gering ist.

Bereits in den drei Prozessen des Sourcings sind immer wieder *Personengruppenunterschiede* untersucht und festgestellt worden. Ältere Personen und Personen mit mehr Vorwissen bzw. Expertise sind zu besseren Sourcingleistungen imstande, die sich teils in den Prozessdaten niederschlagen. Zudem lassen sich Personen danach trennen, wie sie – gemäß der theoriebasierten Taxonomie des „Cognitive Affective Engagement Model of Multiple Source Use“ – beim Sourcing vorgehen. Demnach existieren Gruppen, die minimales bis kein Sourcing betreiben, während andere einen kritisch-evaluativen Umgang mit den Metadaten und den Dokumenten demonstrieren. Dies wirft die Frage danach auf, welche personenbezogenen Merkmale mit den Prozess- und Produktmaßen systematisch korrespondieren.

Bestehende Forschungsüberblicke haben bereits auf eine starke Heterogenität bei den Zusammenhängen von Individualmerkmalen und Sourcingleistungen hingewiesen (Anmarkrud et al., 2022). Dies deckt sich mit dem, was in diesem Kapitel anhand eines eigenen primär produktbezogenen und vor allem das Evaluieren betreffenden Forschungsüberblicks mit einer differenzierteren Perspektive auf die verschiedenen Bereiche des Sourcings ermittelt wurde. Mit aller Vorsicht beim Benennen von Mustern erweisen sich thematisches Vorwissen und thematische Überzeugungen als uneinheitlicher Faktor für bessere oder schlechtere Leistungen. Das Leseverstehen scheint vor allem dann positiv mit Leistungen zu korrespondieren, wenn sowohl beim Lesetest als auch beim Sourcing komplexere kognitive Leistungen nötig sind. Bei den epistemischen Überzeugungen gibt es Hinweise darauf, dass komplexere – und normativ gewendete: angemessenere – Annahmen zu Wissensmerkmalen und -produktion in Einklang mit besseren Leistungen stehen. Alles in allem ist aber die Unklarheit der Befundlage ein frappantes Ergebnis, welches auf die dringliche Notwendigkeit differenzierterer Arbeit in Theorie und Empirie verweist.

Die reichhaltige Forschung vor allem zum Evaluieren als Teil des Sourcings bringt es mit sich, dass Fehlanwendungen und Problemlagen erkennbar werden. Diese Fehlleistungen und Fehlanwendungen schlagen sich sowohl in Prozess- als auch in den Produktmaßen nieder. Sie betreffen auf der einen Seite die primär auf internalen Ressourcen fußenden Evaluationen bei vorgegebenen Sets an Dokumenten, die in der Regel a priori auf Differenzen in der Glaubwürdigkeit geprüft und entsprechend kompiliert wurden. Auf der anderen Seite sind solche Maße in einem weiteren Forschungszweig als Variablen verwendet worden, welcher es den Testpersonen beim „lateralen Lesen“ überlässt, über Internetrecherchen selbstständig zu Urteilen zu gelangen, welche kontextuell grundiert sind, sich also

auf externe Ressourcen stützen. Die Forschungsschwerpunkte bzw. der grundsätzliche Ansatz, evaluative Leistungen zu erfassen, mögen sich unterscheiden. Die Ergebnisse aber konvergieren. Die Testpersonen neigen in Prozessen und Produkten zu einer eher oberflächlichen Bearbeitung, die zugleich auf die hohen kognitiven Anforderungen erfolgreicher Evaluationen von Dokumenten, Quellen und Aussagen verweist, zumal wenn es sich um epistemisch unsichere Informationen handelt. Die prozessbezogenen Maße ergeben studienübergreifend das Muster, dass evaluative Anteile im Gesamtprozess gering sind und das laterale Lesen spärlich oder unvollständig vorkommt. Diese als „schwache Heuristiken“ benennbaren Vorgehensweisen bilden sich in den Produktmaßen ab. Die verschiedenen Testergebnisse eint, dass Begründungen unvollständig ausfallen und bei der adäquaten Einschätzung des zu beurteilenden Gegenstands teils große Unsicherheiten bestehen. Zur Wahrheit gehört bei alledem, dass dies ein Muster ist, dass sich in Interventionsstudien selbst bei den aktiv geförderten Schülerinnen und Schülern sowie Studierenden noch deutlich aus den Daten abzeichnet.

4 Komponente 3: Informationen verstehen und integrieren

Darum geht es

Wenn Personen Texte und Bilder in singulärer Form oder im Auftreten von mehreren Dokumenten lesen, dann ist das Herstellen einer kohärenten internalen Repräsentation eines von mehreren möglichen Lesezielen. Verstehensprozesse werden in der Leseforschung als dynamisch und komplex, als Mischung von passiven Aktivierungs-, Resonanz- sowie Validierungsprozessen und willentlich beeinflussbaren Prozessen beschrieben. Einer der Prozesse, der für das Verstehen als elementar gilt, wird „Integrieren“ genannt. Es ist ein Prozess der Aktivierung und gleichzeitig der Reduktion von Propositionen, mit welchem Leserinnen und Leser auf lokaler, globaler und intertextueller Ebene auf der Basis von Schlussfolgerungen (Inferenzen) Kohärenz stiften. Darum wird das Integrieren im ersten Teilkapitel ([4.1](#)) aus Produkt- und Prozessperspektive beleuchtet und vor allem der Aufbau von Kohärenz fokussiert. Integrieren zum Zweck des Kohärenzaufbaus wird hierbei zunächst als allgemeiner Bestandteil von Leseverstehenstheorien betrachtet, denn dieser Prozess ist unverzichtbar. Zugleich werden zwei Felder mit Besonderheiten des Integrierens genauer fokussiert. Es handelt sich erstens um das Integrieren von schriftlichen und nicht-schriftlichen Inhalten, also Kombinationen von Text und (Bewegt-)Bild, welche in multimedialen digitalen Kontexten gang und gäbe sind. Zweitens wird der Umgang mit multiplen Dokumenten behandelt, also mit Kombinationen von Informationen aus Texten/Bildern und Metadaten über diese Informationen. Das Lesen multipler, sprich: mehrerer, Dokumente ist mit dem Internet zu einem neuen Normalfall des Lesens avanciert, und die Besonderheiten dieses Lesens gilt es theoretisch zu fassen, da die Anforderungen bezüglich Kohärenzaufbau an die lesenden Personen durch diese technologisch vorangetriebene Entwicklung einer Veränderung unterworfen sind.

Den Prozessen und Produkten des Integrierens geht das Teilkapitel [4.2](#) nach. Dessen Fundament bilden intertextuelle Inferenzen, also Kombinationen von Aussagen aus mindestens zwei Texten über die Grenzen dieser einzelnen Texte hinweg. Inferenzen gelten bereits beim Verstehen singulärer Texte als zentral, ihre Bedeutung für das darauf aufbauende intertextuelle Leseverstehen wird ebenfalls hoch veranschlagt. Das gilt insbesondere für einen Ausschnitt, nämlich das Integrieren von konfligierenden Informationen. Dazu ist es nötig, solche Konflikte als solche zu erkennen. Um trotzdem zu einer kohärenten Repräsentation zu gelangen, die auf der inhaltlich konfligierenden Darstellung basiert, sind zusätzliche Prozesse der Kohärenzweiterherstellung nötig. Sie werden in der Forschung unterschiedlich modelliert und in Modellen spezifisch postuliert, aber sie lassen sich zwei Gruppen zuordnen. Erste-Hand-Einschätzungen fokussieren auf evaluative Prozesse, um die Plausibilität einzuschätzen, und betreffen damit den Geltungsanspruch von Aussagen, wahr zu sein. Zweite-Hand-Einschätzungen dienen

der Vertrauenswürdigkeitsbeurteilung. Anhand von Metadaten ist es möglich, Konflikte auf Quellen zurückzuführen, sie zu erklären, aber auch über Quellen zu gewichten. Die Zweite-Hand-Einschätzungen bilden die besser untersuchte Variante.

Das intertextuelle Integrieren ist eine Fähigkeit, welche hinsichtlich ihrer Zusammenhänge mit individuellen Merkmalen in Verbindung gebracht wird (4.3). Drei von ihnen sind das thematische Vorwissen, die thematischen Überzeugungen und die epistemischen Überzeugungen. Das Vorwissen ist aus empirischer Sicht positiv mit Verstehensleistungen verbunden: Je stärker ausgeprägt es ist, desto besser verstehen Personen den Inhalt multipler Dokumente. Bei den Prozessmaßen zeigen sich keine durchgängigen Muster. Die thematischen Überzeugungen als subjektive Annahmen zum Wahrheitsgehalt von Positionen bei einem Thema sind ebenfalls ein durchgängig Produktmaße betreffender Faktor. Sie fungieren als Filter, welcher die Inhalte aus Texten, die überzeugungskonsistent sind, leichter passieren lässt. Dieser „Text-Überzeugungskonsistenz-Effekt“ manifestiert sich darin, dass Leseprozesse unterbrochen bzw. verzögert werden, wenn Personen auf Informationen treffen, die ihren Überzeugungen widersprechen. Auf der Produktseite ist ein deutliches Muster, dass vor allem die Inferenzleistungen betroffen sind: Personen kombinieren Informationen effektiver, wenn sie aus Texten stammen, die ihren Überzeugungen entsprechen. Die dritte und in sich komplexeste Variable sind die epistemischen Überzeugungen, also subjektive und mehrdimensionale Theorien zu Merkmalen des Wissens und dessen Produktion. Dieser in der jüngeren Vergangenheit stark empirisch bearbeitete Bereich verdeutlicht, dass es insbesondere komplexe Überzeugungen sind, die mit besserem Integrieren aus einer Produktwarte einhergehen.

Zwei Problembereiche werden am Ende des Kapitels beleuchtet (4.4). Der erste betrifft die Leistungen von Personen, umfassende mentale Repräsentationen zu entwickeln. Prototypisch dafür ist der relativ geringe Anteil von Studierenden, die Dokumentenmodelle in ihren Produktmaßen erreicht haben. Ein zweiter Problembereich bezieht sich eher auf die misslingenden bzw. suboptimalen Kohärenz- und Integrationsleistungen, welche aus den zu verarbeitenden multimedialen Materialien stammen bzw. aus deren suboptimaler Gestaltung. Die Forschung zum Multimedia-Lernen hat ausgehend von Theorien wie jener zur kognitiven Belastung Prinzipien beschrieben, wie die Arbeitsgedächtnisbelastung reduziert werden kann. Drei Verletzungen dieser Prinzipien – Kohärenzprinzip (Verzicht auf verführerische Details), Signalisierungsprinzip (Struktur und Informationshierarchie hervorheben) sowie Kontiguitätsprinzip (zeitliche bzw. räumliche Nähe von zu integrierenden Informationen) – werden auf der Basis metaanalytischer Ergebnisse dargestellt.

4.1 Funktionen

Der prototypische Kern des Lesens und der Lesekompetenz bezieht sich darauf, dass Personen schriftlich kodierte externe Inhalte eigenständig in internale bedeutungsvolle Sinnzusammenhänge überführen. Die kognitionspsychologische und psycholinguistische Theorie hat sich intensiv damit befasst, diesen Vorgang zu modellieren. So ist aus Sicht der Produkte, dem Leseverstehen, plausibel

annehmbar und empirisch nachgewiesen, dass Personen Textinhalte mit unterschiedlichen Transformationsgraden mental repräsentieren. Ein solches Leseverstehen wird gegenwärtig lediglich als Vorstufe angesehen, um ein sogenanntes tiefes Leseverstehen aufzubauen. Dieser Thematik widmet sich das Teilkapitel [4.1.1](#). Es stellt damit eine Vorleistung für das sich anschließende Teilkapitel [4.1.2](#) dar, welches sich dem Aufbau von widerspruchsfreien und in sich konsistenten Strukturen zentraler Inhalte aus schriftlichen (deskriptionalen) und bildlichen (depiktionalen) Texten verschreibt. Kohärenzleistungen auf der Basis von der lesenden Person selbstständig realisierten Verknüpfungen von Inhalten untereinander und mit Nutzung des eigenen Vorwissens (Inferenzen) bilden das Rückgrat des Leseverstehens. Deshalb verdienen diese Prozesse im Rahmen dieses Teilkapitels besondere Beachtung und entsprechenden Raum.

4.1.1 Leseverstehen – Produktperspektiven

Leseverstehen ist das Ziel des Lesens, auch wenn die Forschung zur Relevanz unterstreicht, dass nicht alle Aufgaben diesem Ziel dienen müssen (s. Teilkap. [2.1.3](#) und [2.2.2](#)). Das Verstehen beim Lesen wird in der Forschung als ein kognitives Produkt beschrieben, das als mentale Repräsentation der lesenden Person auf der Basis textueller und vorwissensbasierter Informationen entsteht und auf der glückenden Interaktion einer Vielzahl von kognitiven Prozessen fußt. Leseverstehen ist kein einheitliches Konstrukt, sondern wird als Mehrebenen-Produkt beschrieben (s. Teilkap. [4.1.1.1](#)) und zunehmend unter dem Schlagwort des „tiefen Leseverstehens“ als normative Leitidee diskutiert ([4.1.1.2](#)).

4.1.1.1 Ebenen mentaler Repräsentationen beim Leseverstehen

Der Ausdruck „Leseverstehen“ ist ein weit gefasster Begriff, dessen alltagssprachlicher Gebrauch die wissenschaftliche Präzisierung erfordert. Üblicherweise und mit breitem wissenschaftlichem Konsens (McNamara & Magliano, 2009) wird mit diesem Terminus das *Produkt* der kognitiven Auseinandersetzung einer lesenden Person mit einem Text bzw. mehreren Texten bezeichnet, das sich auf eine innerhalb des Gedächtnisses der lesenden Person befindliche interne Repräsentation des in einem oder mehreren Texten dargestellten Sachverhalts bezieht. *Prozessuale* Verständnisse betonen hingegen eine Sichtweise als simultane Extraktion und Konstruktion der Bedeutung seitens einer lesenden Person in einer interaktiven kognitiven Auseinandersetzung mit dem Text bzw. den Texten (RAND Reading Study Group, 2002).

Zu den Prozessen und ihren Gewichtungen wurden verschiedene theoretische Modellierungen vorgelegt. Aus Platzgründen sei deshalb an dieser Stelle auf einen Abriss von allgemeinen Theorien des verstehenden Lesens verzichtet,

zumal es bislang keine unitarische (Meta-)Theorie gibt, welche der Vielzahl von interagierenden Prozessen beim Lesen gerecht würde (Sadoski & Paivio, 2007; einen exzellenten und kompakten Überblick über sieben einflussreiche Modelle und Theorien des Leseverstehens geben McNamara und Magliano, 2009, einen weiteren lesenswerten Überblick liefern Graesser et al., 1997). Stattdessen wird ein Modell fokussiert, das wie kein anderes eine Wirkmacht entfaltet hat, nämlich das „Construction-Integration-Modell“ von Kintsch (1998), zumal dieses Modell die Grundlage für Weiterentwicklungen bildete, um das Lesen von multiplen Dokumenten und Bild-Text-Kombinationen zu erklären. Das kognitionspsychologische Construction-Integration-Modell hat Prozesse und Produkte des verstehenden Lesens in ihrer Verschränktheit beschrieben. Der dem Modell den Namen gebende Prozess des Integrierens wird an anderer Stelle (Teilkap. [4.1.2](#)) vertieft, hier soll zunächst das Produkt des Leseverstehens beleuchtet werden.

Zentral für das Leseverstehen ist, welche interne kognitive Repräsentation eine lesende Person auf der Basis des externen Textes entwickelt. Das Construction-Integration-Modell differenziert hierbei drei Repräsentationsebenen bzw. -formen aus: 1) die wortwörtliche Repräsentation, 2) die propositionale Repräsentation und 3) das abstrakt-symbolische Situationsmodell (Kintsch, 1998, 2018), welches in anderen Theorien als „mentales Modell“ bezeichnet wird, wenn sich die Repräsentation auf referenzielle Zusammenhänge erstreckt, z. B. bei räumlichen Modellen (Johnson-Laird, 1983). Diese Repräsentationsebenen unterscheiden sich darin, wie nah oder fern sie am genauen Wortlaut des Textes zu verorten sind, wie sehr individuelle Wissensbestände die Repräsentation steuern und beeinflussen und damit auch wie sehr Inhalte aus Texten transformiert, verarbeitet und in ihrer tieferen Bedeutung verstanden wurden. Die drei Ebenen sind die folgenden:

- 1) *Oberflächenrepräsentation*: Die wortwörtliche Repräsentation bezieht sich auf eine buchstäblich allein die Abfolge von Wörtern und Sätzen, syntaktische Merkmale und genauen Wortlaut betreffende Erinnerung an Textoberflächenmerkmale. Diese Repräsentation wird nicht als Textverstehen erachtet, weil sie keiner eigenständigen Sinnkonstruktion des Gelesenen bedarf. Zudem gilt sie als flüchtig, weil sich in Experimentalstudien hat nachweisen lassen, dass sich Leserinnen und Leser, zumal geübte, schon nach kurzer Zeit nicht mehr an die exakten Ausdrücke in Texten erinnern können, sondern bereits tiefere Verarbeitungsprozesse absolviert hatten, deren Ergebnisse in ihr Langzeitgedächtnis überführt worden waren.
- 2) *Repräsentation der Textbasis*: Die Textbasis bezeichnet das Geflecht von untereinander verbundenen, explizit im Text enthaltenen Informationseinheiten, welche in der Terminologie des Construction-Integration-Modells als „Propositionen“ bezeichnet werden. Propositionen bilden eine abgeschlossene Einheit, eine Aussage, und sie können in ihrem Gesamt ein komplexes Gefüge von über- und untergeordneten, globalen und lokalen Einheiten

darstellen, weshalb man in Makro- und Mikropropositionen unterscheidet. Erstgenannte bilden Hauptideen eines Textes, letztere eher Details. In ihren Zusammenhängen werden sie als Makro- bzw. Mikrostruktur bezeichnet. Zusammen formieren beide Strukturen die Textbasis, also die hierarchisch anordenbaren, explizit gegebenen Informationen untereinander, deren Rekonstruktion beim Lesen Verstehensprozesse erfordert, die eine bereits reduzierte und transformierte Repräsentation des Textinhalts zur Folge haben.

- 3) *Repräsentation als mentales Modell*: Diese Repräsentationsebene stellt nach herrschender Meinung in der Leseforschung die eigentliche inhaltliche Repräsentation des Gelesenen dar. Diese Repräsentation basiert zum einen auf dem propositionalen Netzwerk, aber es handelt sich überdies um eine erheblich durch das Vorwissen angereicherte und teils restrukturierte Vorstellung davon, worum es in dem Text geht. Dafür bedarf es diverser Inferenzen, welche text- und vorwissensbasiert sind und dabei helfen, eine sich vom Wortlaut des Texts gänzlich lösende, nicht einmal zwingend sprachliche Repräsentation aufzubauen.

Soweit ein kursorischer Gang durch die drei Repräsentationsebenen des Construction-Integration-Modells. Diese drei Ebenen der Repräsentation gelten als Kern dessen, was diverse Modelle des Leseverstehens beinhalten (Graesser et al., 1997; McNamara & Magliano, 2009). Insbesondere zeigt sich an diesen Ebenen, wie sehr sich von der Oberflächenrepräsentation hin zum mentalen Modell die unmittelbaren Oberflächenmerkmale des Texts inklusive expliziter Inhalte in ihrer Bedeutsamkeit verringern, während die Nutzung der eigenen Vorwissensbestände an Bedeutung gewinnt. Beker et al. (2017) haben im Zusammenhang mit dem textbasierten Lernen den Begriff „Dekontextualisierung“ verwendet, um diese stärkere Nutzung des Vorwissens zu betonen, die freilich auch dazu dienen soll, textuelle Informationen dauerhaft ins Gedächtnis zu übernehmen, wo sie dann als Ressource für weitere Lesezyklen fungiert.

Die drei Repräsentationsebenen sensu Construction-Integration-Modell bilden nicht nur eine Möglichkeit, Verstehensleistungen nach Transformationsaufwand zu differenzieren und analytisch zu trennen. Vielmehr sind sie auch mitunter die Basis für eine in der Leseforschung zunehmende Fokussierung auf ein Konstrukt, welches als tiefes Leseverstehen bezeichnet wird und welches als Basis für den umfassenden Wissenserwerb gilt (Beker et al., 2017).

4.1.1.2 Tiefes Leseverstehen – zu den Konturen eines Hochwert-Konzepts

Dem Begriff „tiefes Leseverstehen“ haftet aus normativer Perspektive eine positive Konnotation an, die sich allein darin manifestiert, dass die höchsten Kompetenzstufen der kontinuierlichen Skalen in Schulleistungstudien das Konstrukt

widerspiegeln (Jambor-Fahlen & Philipp, 2022), sodass die Personen besonders kompetent gelten, welche diese Stufen empirisch zu erreichen in der Lage sind. Teil der normativen Perspektive ist auch, wenn bspw. oberflächenorientiertes bzw. sogar oberflächliches Lesen als Begründung bemüht wird, um Mediumseffekte im Leseverstehen, also ein schlechteres Leseverstehen beim Lesen am Bildschirm, zu erklären (Delgado et al., 2018). Dabei ist tiefenorientiertes Lesen bzw. Lernen empirisch genauso wenig per se superior, wie es im Gegenzug auf die vermeintliche Inferiorität der oberflächenorientierten Zugangsweisen zu Texten und Lerngegenständen zutrifft (Dinsmore, 2017a). Tatsächlich wirft die dichotomisierende Zuordnung immer noch diverse konzeptionell-definitorische Fragen auf, zudem sind individuelle und kontextuelle Merkmale als theoretisch einzubeziehende Einflussgrößen ebenfalls von Belang, und auch die Frage der Operationalisierung von Prozessen und Produkten des oberflächen- resp. tiefenorientierten Lesens stellt sich vehement in der Forschung (Dinsmore & Alexander, 2012). Aus dieser Unschärfe heraus bedarf es einer Konturierung dessen, was mit tiefem Leseverstehen sowohl produkt- als auch prozessbezogen gemeint ist (s. dazu bspw. den Sammelband von Millis et al., 2018a).

Das Konstrukt tiefes Leseverstehen ist wie angedeutet als *Produkt* nicht einheitlich definiert. Einige Autorinnen und Autoren binden tiefes Leseverstehen explizit an höhere Repräsentationsebenen des Gelesenen an, indem sie es für mentale Modelle bzw. Situationsmodelle reservieren (Allen & McNamara, 2020; Graesser et al., 1996; Mayer, 2002b; McNamara et al., 2016; Schnotz, 2022). Tiefes Leseverstehen ist nach dieser Auffassung ein Leseverstehen, welches im „Construction-Integration-Modell“ (Kintsch, 2018) die höchste zu erreichende Stufe darstellt und besonders hohe Verarbeitungsleistungen seitens der lesenden Person im Sinne der vorwissensbasierten Kohärenzbildung voraussetzt. Andere Autorinnen und Autoren nehmen diesen Grundgedanken auf, indem sie tiefes Leseverstehen auf die erbrachten Leistungen auf den höchsten Kompetenzstufen in empirischen Leseleistungsstudien engführen (Millis et al., 2018b). Das heißt: Tiefes Leseverstehen ist hier partiell auf Kompetenzstrukturmodelle und deren Operationalisierung gestützt; der Ausdruck verweist damit auf den Anforderungscharakter der Leistungen auf höheren Kompetenzstufen in Leseleistungsstudien. Weitere Leseforscher beschreiben einen Idealtypus eines tiefen Leseverstehens ebenfalls eher personengebunden, indem sie eine lesende Person mit hohen Kohärenzstandards, strategischen Inferenzbildungsprozessen und Mehrebenen-Repräsentationen des Gelesenen skizzieren, die ein tiefes Leseverstehen entwickelt (Graesser, 2015).

Die Gleichsetzung tiefen Leseverstehens mit umfassender Kohärenzbildung der Inhalte unter Nutzung der eigenen Wissensbestände in den produktbezogenen Perspektiven findet sich komplementär in *prozessualen Perspektiven* wieder. Dabei werden Prozesse des tiefen resp. tiefenorientierten Verstehens mit jenen des oberflächenorientierten Vorgehens kontrastiert. Diese Prozesse werden in

der Forschung nicht allein auf das Lesen beschränkt, sondern breiter gefasst und als zwei grundsätzliche Herangehensweisen des Lernens (mit Texten) als stabile Eigenschaften von Personen oder als situative Vorgehensweisen modelliert (Dinsmore, 2017a). Ihr Wert, aber auch ihre Gefahr liegt in der heuristischen Gegenüberstellung von Oberflächen- und Tiefenorientierung. Der Wert besteht darin, zwei Modi des Lernens mit spezifischen Fokussen, Prozessen und Nähe zu Ergebnissen zu beschreiben, die Gefahr darin, dass sie gegeneinander ausgespielt werden.

Tiefen- und oberflächenorientierte Prozesse sind bezogen auf Gedächtnisstrategien (Craik, 2002), Lernprozesse (Anderson et al., 2001), musterhafte und eher stabile Herangehensweisen beim Lernen (Marton & Säljö, 1976), als konsequente und gleichwohl interaktive Abfolge von Aneignung und Konsolidierung von Wissen (Hattie & Donoghue, 2016) oder als Passung von (selbstregulatorischen) Strategien und angestrebtem Lernerfolg (Dinsmore & Hattan, 2020; Winne, 2018b) in der kognitionspsychologisch orientierten Forschung vielerorts herausgearbeitet und diskutiert worden. Gemeinsamer Nenner der Gegenüberstellung von oberflächen- und tiefenorientierten Verarbeitungsprozessen in der Forschungsliteratur ist,

- dass *oberflächenorientierte (eher automatisierte) Prozesse* darauf abzielen, explizit gegebene, einzelne Informationen zu erkennen, zu sammeln und zu memorieren,
- während *tiefenorientierte (strategische) Prozesse* auf Transformation, Kombination und Anreicherung und Verstehen des Gelesenen und darüber hinaus auf einen stabilen Wissensaufbau fokussieren.

In dieser Hinsicht besteht eine Komplementarität zur obigen Produktwarte: Kognitive Prozesse, die zur Bildung mentaler Modelle nötig sind (nicht nur als Verstehensprozesse, sondern auch als lernstrategische Vorgehensweisen), werden mit tiefem Leseverstehen etikettiert bzw. unterstützen als Lese-/Lernstrategien, die über den eigentlichen kognitiven Leseverstehensprozesse ablaufen, eben diese Verstehensprozesse.

Einen Merkmalskatalog des tiefen Leseverstehens, der sich aus der Frage nach der Operationalisierung des Konstrukts speist, haben O'Reilly et al. (2018) vorgelegt. Sie gehen von sieben Merkmalen bzw. von Prämissen aus. Zugleich machen sie explizit darauf aufmerksam, dass der Merkmalskatalog weder abgeschlossen noch ohne textuelle, kontextuelle und individuelle Klärungen bestimmbar sei, woran sich Tiefe beim Leseverstehen bemisst. Demnach modellieren sie tiefes Leseverstehen als ein relatives Konstrukt. Die Merkmale sind nicht mehr allein Prozessen und Produkten ohne weiteres zuzuordnen, sondern tragen beides in sich, wie es an der nachstehenden Merkmalsliste deutlich wird:

- 1) *Tiefes Leseverstehen bemisst sich an Zwecken und Ziel des Lesens.* Mit dieser Prämisse wird tiefes Leseverstehen auch als durch Aufgaben ansteuerbares Ergebnis konzeptualisiert, das als Passung zum Auftrag des Lesens angesehen wird. Explizit bemühen O'Reilly et al. (2018) die Kohärenzstandards (s. Teilkap. [2.1.2](#)) als theoretischen Anknüpfungspunkt. Tiefes Leseverstehen wird demnach in Abhängigkeit von der angestrebten Kohärenz beim Lesen bestimmt, entsprechend sind Leseziele mit der Notwendigkeit, Sachverhalte umfassend und vollständig zu repräsentieren, etwa beim Lernen, Aspiranten für ein sich durch intensivere Leseprozesse einstellendes umfassenderes Leseverstehen.
- 2) *Tiefes Leseverstehen basiert auf Inferenzen, vor allem hinsichtlich impliziter Inhalte.* Hierin deutet sich die weiter oben schon genannte Fähigkeit an, nicht nur explizit gegebene Informationen in Texten zu erkennen und zu repräsentieren, sondern auch Verbindungen zwischen Informationen herzustellen, welche der Text anlegt, aber nicht expliziert. Schlussfolgerndes Denken beim Lesen wird damit als Bedingung für tiefes Leseverstehen angenommen.
- 3) *Tiefes Leseverstehen bezieht sich auf Strukturen von Informationen.* Dies ähnelt einerseits dem zweiten Merkmal in der Aufzählung, geht aber darüber hinaus, weil es um die Strukturen von Informationen auf globaler Textebene geht und vor allem auf die kausalen Zusammenhänge zwischen größeren Informationseinheiten. Diese sind, je nach Text, verarbeitungsintensiver als lokale Inferenzen.
- 4) *Tiefes Leseverstehen beinhaltet das dynamische Stellen von gehaltvollen Fragen an den Text.* Statt auf Bedeutungsrekonstruktion zielt dieses Merkmal darauf ab, dass Personen in der Auseinandersetzung mit Fragen an Texte herantreten, neue Fragen beim Lesen bilden und somit eine Quasi-Befragung des Textes und ihres sich entwickelnden Leseverstehens vornehmen. Fragen nach dem Warum beispielsweise gelten als Beispiel für tiefergehendes Verstehen, andere Fragen beinhalten die kritische Überprüfung des eigenen Verstehens, etwa durch Selbsterklärungen. Wichtig ist, dass Fragen dynamisch genutzt werden, also nicht nur vor oder nach dem Lesen, sondern auch während der Lektüre.
- 5) *Tiefes Leseverstehen besteht darin, die Personen hinter den Texten und ihre Aussagen zu evaluieren.* Dieses Merkmal greift auf, was im Kapitel [3](#) umfassend behandelt wurde. Die Motive von Autorinnen und Autoren zu interpretieren, ihre grundsätzliche Eignung einzuschätzen, sich qualifiziert zu Sachverhalten zu äußern, und die Güte von Aussagen zu bemessen – dies sind Fähigkeiten, die vom digitalen Lesen forciert eingefordert werden. Solche Fähigkeiten erfordern zudem umfassende Wissensbestände der lesenden Person, da kritisches Lesen sich in einer aktiven Beurteilung mittels meist textexterner Kriterien zeigt.

- 6) *Tiefes Leseverstehen bemisst sich nicht nur an individuellen Merkmalen, sondern an Disziplinen und Diskursen.* Damit ist gemeint, dass Texte nicht akontextuell vorliegen, sie sind vielmehr ihrerseits in Diskursen zu situieren, die ihre Bedeutung modifizieren. Diese Kontexte, die als übergeordnete Diskurse, als fachgemeinschaftliche Artefakte und als disziplinär gerahmte Aussagen mit Spezifika der jeweiligen Wissensgemeinschaften zu denken sind, systematisch zu berücksichtigen, um adäquate Bedeutungskonstruktionen vorzunehmen, erhöht nochmals die Anforderungen. Wer Sekundär- und Primärquellen im Fach Geschichte nicht unterscheidet, wer in den Naturwissenschaften nicht weiß, wie dort Wissen hergestellt wird, und wer im Internet nicht Texte miteinander vergleicht, hat demnach wenig Chancen auf tiefes bzw. adäquates Leseverstehen.
- 7) *Tiefes Leseverstehen beinhaltet auch die Anwendung des Gelesenen in Kontexten, die nicht der Intention des Texts entsprechen.* Nach diesem instrumentellen Verständnis besteht tiefes Leseverstehen darin, mittels der prozessierten Informationen einen Anwendungsbezug (etwa beim Lösen von offenen Fragen, bei der Entscheidung bei Kontroversen oder im Transfer auf neue Situationen) herzustellen, in welchem sich die Adäquanz des Leseverstehens manifestiert.

Dieser Katalog von Merkmalen fokussiert also auf die Adäquanz von Informationsnutzung (1 und 7), auf Verknüpfung von Informationen (2 und 3), auf kritisches Interagieren mit dem Text (4 und 5) und die Berücksichtigung von Kontexten (6). Deutlich ist, wie sehr das Nutzen von eigenen Wissensbeständen, von Analysen und regelgeleitetem Vorgehen bei der (kritischen) Nutzung von textuellen Informationen das Konstrukt tiefes Leseverstehen konturieren und profilieren.

4.1.2 Kohärenz aufbauen: allgemein und mit multiplen Repräsentationen und Dokumenten – Prozessperspektiven

4.1.2.1 Kohärenz herstellen: der zentrale Prozess des Integrierens

Ein zentrales Ziel des Lesens ist – unter anderem – das Verstehen von in Texten beschriebenen Sachverhalten als mentale Repräsentation. Hierbei gilt der Aufbau einer kohärenten Repräsentation als bedeutsam. *Kohärenz* im kognitionspsychologischen Sinne meint *eine vernetzte, stabile Verbindung von Aussagen aus einzelnen Textteilen in einer nachvollziehbaren Weise*, und wie Kohärenz von Leserinnen und Lesern hergestellt wird, ist zentraler Gegenstand diverser

Lesetheorien (McNamara & Magliano, 2009). Für die Herstellung von Kohärenz ist ein Prozess verantwortlich, der als „Integrieren“ bezeichnet wird, in dessen Zentrum sich das Inferieren als Prozess schlussfolgernden Denkens befindet.

Integrationsprozesse als Bindeglied von textuellen Informationen und Vorwissen der lesenden Person

Was ist mit dem Ausdruck „Integrieren“ gemeint? Integrieren ist ein Prozess, der ganz besondere Prominenz im einflussreichen Construction-Integration-Modell von Kintsch (2018) hat und daher namensgebend ist. Verstehensprozesse, darunter im Lesen, werden im Construction-Integration-Modell als dynamisches, interagierendes Zusammenspiel von zwei Prozessen beschrieben. Diese beiden Prozesse haben eine komplementäre Funktion im Aufbau verschiedener kognitiver Repräsentationen beim Lesen:

- Die *Konstruktions*prozesse sind automatisch ablaufende, konnektionistische Aktivierungen verschiedener Informationen (Propositionen), welche sowohl aus dem Text als auch aus dem Gedächtnis stammen und im Arbeitsgedächtnis präsent gehalten werden und als lose verknüpft, durchaus inkohärentes Netzwerk von Informationen vorliegen. Diese Verknüpfungen sind bottom-up getrieben, ergeben sich also als teils automatisierte und gedächtnisbasierte Auseinandersetzung der lesenden Person, die diese dynamischen Prozesse nicht steuert.
- Dem *Integrieren* kommt eine restringierende und kohärenzstiftende Funktion zu: Aufgrund von genutztem Vorwissen und semantischen Hinweisen aus dem Kontext des Gelesenen werden nur die plausibel wirkenden Beziehungen zwischen Informationen aufrechterhalten. Das Netzwerk potenzieller Bedeutungszusammenhänge wird also einerseits ausgedünnt, andererseits verstärkt, und hier kommen die Inferenzen als vorwissensbasierte Schlussfolgerungen zum Einsatz: Sie verbinden, wo es sinnvoll und für das Verstehen nötig ist, Propositionen untereinander. Die Logik hierbei ist eine eher top-down ablaufende: Die Repräsentation enthält vor allem jene Propositionen, welche stark mit anderen Propositionen verbunden sind, während periphere Propositionen aus dem Netzwerk verschwinden.

Dieses Zusammenspiel von Konstruktion und Integration erfolgt dynamisch während des gesamten Leseprozesses, bis ein stabiles Netzwerk von Propositionen vorliegt, welches sowohl textuelle Anteile enthält als auch vorwissensbasierte Bestandteile (Richter & Singer, 2017).

Aus einer übergreifenden Perspektive auf diverse theoretische Lesemodelle hinweg extrahierten McNamara und Magliano (2009, S. 340) eine *allgemeine Definition des Integrierens*. Sie bezeichnen es dezidiert als *einen Prozess, in*

welchem Informationen aus verschiedenen Quellen wie Texten und eigenem Gedächtnisinhalt in einem iterativen Prozess integriert, also aufeinander mit Ziel der Kohärenzbildung bezogen werden, wobei dieser Prozess so lange fortschreitet, bis das Netzwerk kaum noch Veränderungen zwischen den Iterationen aufweist. Das Netzwerk stabilisiert sich, wenn die Aktivierung der Verbindungen zwischen den Knoten des Netzwerks ausreichend abgeschlossen ist.

Das *Produkt des Integrierens* bezieht sich auf verschiedene Repräsentationen, die mit der Informationsmenge und der Art von mentaler Repräsentationen des Gelesenen zu tun haben und miteinander kombinierbar sind, d. h. man kann auf lokaler Ebene mentale Modelle entwickeln oder auf globaler Ebene eine propositionale Repräsentation der Textbasis erlangen (Kintsch, 2018; McNamara & Magliano, 2009):

- Bei der *Textmenge* werden im Construction-Integration-Modell eine lokale, also kleinteiligere Textteile betreffende Kohärenz und die globale Kohärenz unterschieden, letztgenannte bezieht sich auf größere Textteile, ganze Texte oder im Falle multipler Dokumente auf mehrere Texte.
- Unterschieden wird zudem die Art der Repräsentation, in der Forschung auch als *Ebene des Verstehens* bezeichnet (s. Teilkap. 4.1.1.1). Für das Verstehen sind propositionale Repräsentationen nötig, also die Repräsentation von einzelnen Aussagen (Propositionen) in ihrer Verbindung untereinander, die als Textbasis bezeichnet wird. Eine weitere, erheblich stärker auf Vorwissensleistungen und die inferenzbasierte Elaboration der Textbasis angewiesene Repräsentation, ist das mentale Modell bzw. das Situationsmodell. Diese Repräsentation ist eine lesendenseitig stärker angereicherte Variante der Textbasis und enthält die impliziten Verbindungen von Propositionen.

Der kohärenzstiftende Prozess des Integrierens ist auf das Vorwissen wegen des Fokus' auf die Inferenzleistungen als *conditio sine qua non* angewiesen. Darum wird angenommen, dass kohärente Repräsentationen darauf basieren, dass Personen Schlussfolgerungen ziehen und textuelle Informationen mit dem eigenen Vorwissen automatisiert oder willentlich gesteuert aufeinander beziehen (Cook & O'Brien, 2017). Diese Zusammenhänge hat Goldman (2004) in Form einer Vierer-Typologie dargestellt, welche in Abbildung 10 wiedergegeben ist. Die Differenzierung verläuft danach, wie stark eine Person ihr eigenes Vorwissen nutzt und wie vernetzt die repräsentierte Textbasis als Resultat des Integrierens ausfällt. Unten links im Quadranten ist eine *fragmentarische Repräsentation* dargestellt, welche sich durch isolierte Informationen im Gedächtnis der Person auszeichnet. Rechts davon ist die *eingekapselte Repräsentation* im Quadranten unten enthalten. Eine Person mit dieser Repräsentation hat einen Text korrekt verstanden, aber ihn nicht mit den Vorwissensinhalten verbunden. Im oberen Teil des Diagramms kommt das Vorwissen zum Tragen. Links, als *assimilierte*

Variante, werden Textinhalte nur mit dem Vorwissen, aber nicht untereinander verbunden. Erst in der *integrierten kohärenten Variante* sind Vorwissen und Inhalte aus Texten aufeinander bezogen. Goldman (2004) hält nur in dieser Form echtes Lernen durch das Lesen möglich, und aktuelle Befunde sprechen dafür, dass mangelndes Vorwissen tatsächlich dazu führt, dass höhere Repräsentationsebenen über die Textbasis hinaus dann gefährdet sind, wenn das Vorwissen fehlt (Smith et al., 2021).

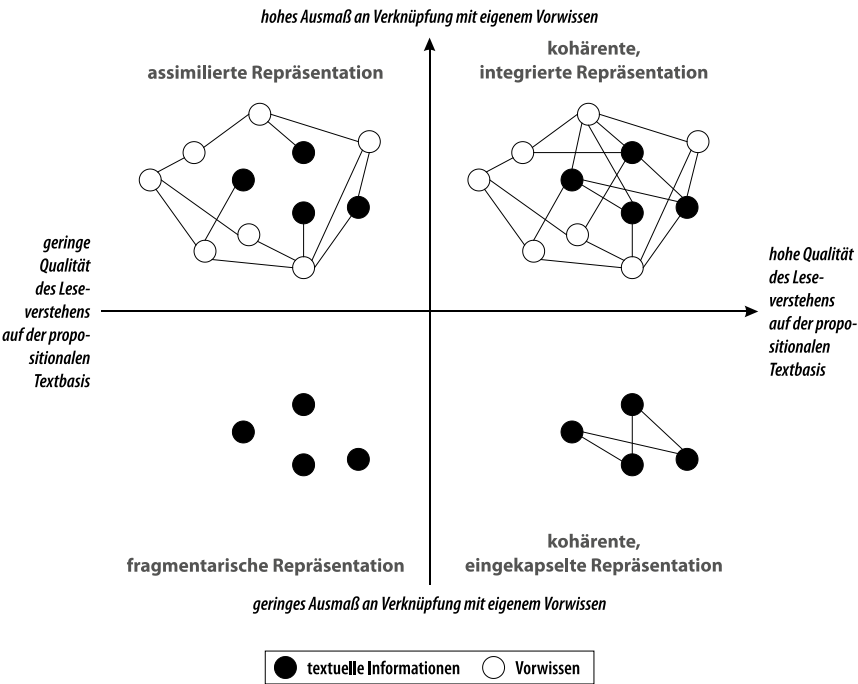


Abbildung 10: Vier Arten der mentalen Repräsentation der Textbasis in Abhängigkeit von Vorwissensnutzung (vertikale Achse) und der Qualität der Vernetzung der Propositionen (horizontale Achse; Quelle der Darstellung: Goldman, 2004, S. 325)

Der hohe Stellenwert der Inferenzen beim verstehenden Lesen

Die Kohärenzherstellung ist eine zentrale kognitive Leistung der lesenden Person, um Texte zu verstehen. Als konstitutiv dafür gilt das schlussfolgernde Kombinieren von Informationen, welches als Inferenzbildung bezeichnet wird. Wenn Leserinnen und Leser beispielsweise aus dem Satz zuvor das Adverb „dafür“ mit dem Substantiv „Kohärenzherstellung“ verbinden und daraus schließen, dass Inferenzbildung die Basis für Kohärenzherstellung ist, dann ist dies ein Beispiel für das, was mit Inferenzbildung gemeint ist.

Eine ganz allgemeine prozessuale Definition des Inferierens geben McNamara und Magliano (2009, S. 302), indem sie es als Prozess beschreiben, in welchem *Informationen aus der Umgebung* (z. B. der aktuelle Satz eines Textes) *mit Informationen verknüpft werden, die sich nicht in der aktuellen Umgebung befinden* (z. B. Informationen in der anschließenden Darstellung des Textes, Weltwissen oder episodisches Wissen über vergangene Ereignisse wie andere Texte oder frühere Erfahrungen). Es geht nach dieser Definition also um *Schlussfolgerungen seitens der lesenden Person, die mindestens zwei Informationen schlüssig kombiniert*.

Wann Leserinnen und Leser Inferenzen beim Lesen bilden, ist in der Lese-forschung sehr unterschiedlich beschrieben worden. Zwei Positionen, die sich kontrastieren lassen, sind die minimalistische Hypothese, die McKoon und Ratcliff (1992) vorgelegt haben, und die konstruktionistische Hypothese von Graesser et al. (1994). Diese beiden Perspektiven unterschieden sich darin, ob sie lokale und globale Kohärenz fokussieren und ob sie die Ziele der lesenden Person berücksichtigen:

- Die *minimalistische Hypothese* (McKoon & Ratcliff, 1992) geht auf den Unterschied von globaler und lokaler Kohärenz ein und fokussiert auf die lokale Kohärenz, wobei sie dies auf eine Textmenge von zwei Sätzen limitiert. Die Hypothese besagt, dass Leserinnen und Leser nur dann Informationen miteinander verbinden, wenn die dafür nötigen Informationen leicht verfügbar sind, weil sie im Arbeitsgedächtnis vorliegen (woraus sich die geringe Textmenge erklärt, welche die Hypothese fokussiert) oder weil die Wissensbestände der lesenden Person sie leicht aktivierbar machen und sie damit der lesenden Person schnell zur Verfügung stehen. Diese Hypothese postuliert demnach, dass nur das Minimum von lokalen Inferenzen gebildet wird, was sich im Namen der Hypothese niederschlägt. Alle weiteren Inferenzen, die darüber hinausgehen, wären lesestrategischer Natur. Mit Richter und Schnotz (2018) lässt sich annehmen, dass bei dieser Art des Inferierens lediglich eine Textbasis-Repräsentation erzielen lässt, die freilich zu fragmentarisch bleibt, um von einem tiefen Leseverstehen zu sprechen.
- Die *konstruktionistische Hypothese* (Graesser et al., 1994) beansprucht demgegenüber einen größeren Geltungsbereich, sie stammt jedoch ursprünglich aus dem Kontext narrativen Textverstehens. Sie besagt, dass Personen beim Lesen aufgrund dreier Annahmen Inferenzen bilden. Erstens tun sie es, um sowohl auf globaler als auch lokaler Ebene Kohärenz aufzubauen, zweitens fokussieren sie auf die Kausalität von Zusammenhängen zwischen lesend verarbeiteten Propositionen und drittens spielen die mit dem Lesen angestrebten Zwecke eine Rolle, wie viel Aufwand Leserinnen und Leser beim Herstellen von Kohärenz mit dem Inferieren betreiben. Im Vergleich zur minimalistischen Hypothese sind mehr Inferenzen erwartbar, vor allem implizite Zusammenhänge betreffende.

Die beiden Annahmen wirken zunächst antagonistisch, doch Graesser et al. (1997) spezifizieren die Situationen, in denen beide Hypothesen angemessen zum Einsatz kommen können. Ein Lesen gemäß der minimalistischen Hypothese ist erwartbar, wenn a) schnelles Lesen, b) global wenig kohärente Texte bzw. c) geringes Vorwissen der lesenden Person die Lesesituation kennzeichnen. Im Vergleich dazu tritt das Lesen im Sinne der konstruktionistischen Hypothese bei a) langsameren Lesen, b) global kohärenten Texten, c) höherem Wissen der lesenden Person sowie d) dem Ziel des Verstehens des gesamten Texts auf. Explizit bringen Richter und Schnotz (2018) mit dem Lesen gemäß der konstruktionistischen Hypothese die Kohärenzstandards (s. Teilkap. 2.1.2) in Verbindung: Hoch ausgeprägte Kohärenzstandards lassen mehr Inferenzen erwarten. Diese Gegenüberstellung von minimalistischer und konstruktionistischer Hypothese spiegelt nicht zuletzt wider, dass Verstehensprozesse, in diesem Fall: Inferenzen, entweder bottom-up geleitete Resonanzprozesse oder top-down, von Zielen, Schemata und Strategien beeinflusste Verstehensprozesse bilden (Cook & O'Brien, 2017; León & Escudero, 2017).

Spätestens an dieser Stelle muss die Frage gestellt werden, welche *Arten von Inferenzen* es gibt. Inferenzen sind ein Sammelbegriff für verschiedene Schlussfolgerungen. Entsprechend sind in der Leseforschung verschiedene Taxonomien vorgelegt worden, die Inferenzen klassifizieren (Kendeou, 2015). Grob unterschieden werden in Bezug auf einzelne Texte text- und vorwissensbasierte Inferenzen, die sich dadurch auszeichnen, dass textinterne, d. h. vorgängige Informationen vor dem aktuell gelesenen Inhalt mit dem aktuellen Inhalt verknüpft werden (textbasierte Inferenzen), oder außerhalb des Texts befindliche Informationen, welche mit den aktuellen Inhalten verknüpft werden (León & Escudero, 2017; McNamara & Magliano, 2009; van den Broek et al., 2015). Dabei wird eine Form von Inferenzen als besonders wichtig erachtet, nämlich die *Kausalinferenz* (León & Escudero, 2017) bzw. die kausal-logische Inferenz (van den Broek et al., 2015). Diese Form kennzeichnet, dass sie zwei Informationen als Ursache und Folge miteinander verbindet, ergo ein Bedingungsgefüge beinhaltet, welches wiederum für den Kohärenzaufbau dienlich ist. Kausalinferenzen lassen sich auf verschiedene Arten klassifizieren, ein Ausschnitt daraus ist in der Tabelle 16 dargestellt. Die darin enthaltenen Arten von Inferenzen unterscheiden sich darin, ob sich die Inferenzen auf Vorheriges oder Zukünftiges gemessen am fokalen Punkt der Aufmerksamkeit beim Lesen beziehen und ob sie eher automatisch und textbasiert erfolgen oder von der lesenden Person durchaus im Sinne einer Lesestrategie aktiviert werden, um das eigene Vorwissen gezielt zu nutzen.

Tabelle 16: Arten von Kausalinferenzen

Kategorie	Art (mit Zeitpunkt und Steuerung von Text bzw. lesender Person)	Kurzcharakterisierung
Rückwärts gerichtete Kausalinferenz	<i>Brücken-/Verbindungsinferenz</i> (parallel zum Lesen, textgesteuert)	Kausaler Vorgänger ist noch vom vergangenen Lesezyklus im Arbeitsgedächtnis
	<i>Wiederherstellungsinferenzen</i> (parallel zum Lesen, textgesteuert)	Kausaler Vorgänger ist nicht mehr im Arbeitsgedächtnis verfügbar, wird aber wieder gelesen oder reaktiviert
	<i>Erklärungsinferenzen</i> (parallel zum Lesen, lesendenseitig gesteuert)	Kausaler Vorgänger ist nicht im Arbeitsgedächtnis verfügbar, wird aber über Vorwissen aktiviert
Vorwärts gerichtete Kausalinferenz	<i>Prädiktive Inferenzen</i> (nach dem Lesen, lesendenseitig gesteuert)	Konsequenz lässt sich aufgrund von Vorwissen und zuvor gelesenen Text vorhersagen
Andere	<i>Elaborationen</i> (nach dem Lesen, lesendenseitig gesteuert)	Verknüpfte Information, die nicht für eine Erklärung oder die aktuelle logische Struktur zwingend gebraucht wird, aber die Kohärenz in weiteren Lesezyklen beeinflussen kann
	<i>Inferenzen bei multiplen Dokumenten/multimedialen Darstellungen/multiplen externen Repräsentationen</i>	Jedwede o.g. Inferenz, die sich auf Darstellungen in multiplen Dokumenten oder multimedialen Darstellungen bezieht

(Darstellung basierend auf van den Broek et al., 2015, S. 101, und León & Escudero, 2017, S. 68; Ergänzung multimedialer Darstellungen und multipler externer Repräsentationen in der letzten Art von Inferenzen am Tabellenende auf der Basis von Alexander & Disciplined Reading and Learning Research Laboratory, 2020, sowie Butcher & Davies, 2015)

Solche Inferenzen, wie sie in Tabelle 16 in aller Kürze charakterisiert wurden, sind für den Kohärenzaufbau beim Lesen nötig, indem sie die Informationen in den singulären oder multiplen (multimedialen) Texten und darüber hinaus im Vorwissen entweder automatisch ablaufend oder gezielt gesteuert zusammenfügen. Diese verbindende Funktion erklärt, weshalb Inferenzen sowohl als Produkte als auch der Prozesse des aktiven Inferierens für das Leseverstehen als hoch veranschlagt werden (McNamara & Magliano, 2009).

4.1.2.2 Multiple Repräsentationen integrieren (Bild-Text-Integration)

Verstehendes Lesen und Lernen im digitalen Kontext, aber nicht nur dort wird zunehmend als ein Lesen von mehreren Texten und verschiedenen Modalitäten aufgefasst. Hierfür beginnt sich der Ausdruck „multiple externe Repräsentation“ einzubürgern. Mit ihm ist eine *Kombination verschiedener Symbolsysteme in externalen Repräsentationen gemeint, die zusammen einen Sachverhalt beinhalten und dabei die Vorzüge der jeweiligen Symbolsysteme* – (Schrift-)Sprache als Variante, komplexe und abstrakte Inhalte in Texten zu transportieren, und

bildliche Darstellung zur effizienten Vermittlung von Strukturen und Analogie-Relationen – *gezielt vereinen* (Rau, 2020). Daraus ergeben sich Vorteile, doch zugleich stellen sich damit ebenfalls Anforderungen an lesende Personen, die aus mehreren externalen Repräsentationen, also stellvertretenden Darstellungen von Sachverhalten, Gegenständen, Ereignissen in Texten und Bildern, eigene interne Repräsentationen im Sinne des Verstehens aufbauen. Dieser Thematik widmet sich dieses Teilkapitel, indem es zunächst die Erweiterungsnotwendigkeit des Lesekompetenz-Konstrukts touchiert, um dann zwei einflussreiche Theorien zu präsentieren, von denen die eine ein Rahmenmodell ist, während die zweite genauer spezifiziert, wie Leserinnen und Leser aus mehreren Symbolsystemen kognitive Repräsentationen aufbauen.

Die Erweiterung des Lesekompetenzbegriffs

Die bisherigen Ausführungen waren allgemein gehalten und spiegeln die Forschungstradition wider, nach der eine lesende Person auf einen singulären schriftsprachlichen Text trifft und diesen zu verstehen versucht. Dieser Zugang ist jedoch zu limitiert, um den realweltlichen Anforderungen, an denen sich ein Kompetenzkonstrukt immer wieder zu orientieren und normativ zu legitimieren hat, vollends zu genügen. Darum plädierten Schnotz und Dutke (2004, S. 63) für einen *weiter gefassten Begriff der Lesekompetenz*, den sie als Fähigkeit beschreiben,

„schriftliche Dokumente zu verstehen, in denen sowohl verbale Informationen in Form von Schriftzeichen (graphemisch) als auch piktoriale Informationen in Form von Bildzeichen (graphisch) enthalten sind. Lesekompetenz entspricht also der Fähigkeit zum Verstehen von multiplen Darstellungen in schriftlichen Dokumenten, die Texte, Bilder, Diagramme, Tabellen oder andere Arten externer Repräsentationen enthalten können.“

An dieser Definition ist zweierlei hervorzuheben: erstens die Ausweitung auf auch grafische Texte mit piktorialen Informationen und zweitens der Plural bei den Texten selbst, d. h. auf mehrere Texte bzw. Texte mit mehreren Bestandteilen, die es kohärent aufeinander zu beziehen gilt. Deshalb differenzieren Schnotz und Dutke (2004) *intra- und interrepräsentationale Kohärenzbildung*. Erstgenannte Form beinhaltet, dass man die Inhalte jeder einzelnen externalen Repräsentation, also jeden einzelnen Text und jedes Bild, einzeln mittels kognitiven Integrationsprozessen in eine propositionale Repräsentation bzw. ein mentales Modell überführt und jeweils für sich kohärent intern repräsentiert. Die zweite Variante geht darüber hinaus, weil hier die mindestens zwei externen Repräsentationen in ihrem Gesamt aufeinander zu beziehen sind, wobei die Inhalte sich wechselseitig beeinflussen, also im günstigsten Fall den Aufbau einer interrepräsentationalen

Kohärenzbildung erleichtern, doch sie können auch einander widersprechen, komplementäre Informationen liefern, die Bedeutung der jeweils anderen externen Repräsentation einschränken und tiefere Lernprozesse gezielt stimulieren (Ainsworth, 1999, 2006, 2022).

Grundsätzlich wird jedoch in der Lehr-Lernforschung ein positives, sich durch verschiedene Funktionen von externen Repräsentationen erklärbares Verhältnis von verbalen und piktorialen Informationen im Verbund angenommen (Ainsworth, 2022), wie es die reichhaltige Forschungstradition zum multimedialen Lernen immer wieder theoretisch und empirisch betont (Mayer, 2019a, 2021). Dies konnte jüngst speziell für das Leseverstehen metaanalytisch nachgewiesen werden. Das Beispiel aus dem Kasten „Illustrierendes Beispiel 16“ unterstreicht, dass statische bildliche Ergänzungen von Texten positive Wirkung entfalten. Dieses Bemühen um die metaanalytische Klärung von möglichst lernförderlichen Effekten des Verstehens von Texten, denen gezielt Bilder hinzugefügt werden, beschäftigt die Forschung seit Längerem (Carney & Levin, 2002; Levin et al., 1987).

Illustrierendes Beispiel 16

Verstehen Personen Texte besser, wenn diese gezielt bebildert wurden?

A) Was war das Erkenntnisinteresse?

Gegenstand der Metaanalyse (Guo et al., 2020) war eine zweigeteilte Fragestellung zu den Effekten der Text-Bild-Kombinationen auf das Leseverstehen. Erstens wollte das Forschungsteam die allgemeinen Effekte von Studien bündeln, in denen gezielt Bilder verschiedener Art den zu lesenden Texten beigelegt waren oder fehlten. Zweitens bestand ein Erkenntnisinteresse daran, das Zustandekommen dieser Effekte auf Moderatorvariablen wie Art der Bilder, Alter der Teilnehmenden, Textsorten (narrativ/expositorisch) und Testform zurückzuführen.

B) Wer wurde untersucht?

Den Analysepool bildeten 39 kriteriumsorientiert ausgewählte Studien, in denen 2.103 Personen untersucht worden waren. Diese Studien wurden im Zeitraum 1985–2018 veröffentlicht. Bei den soziodemografischen Merkmalen sind lediglich die Altersgruppen in der Metaanalyse ausgewiesen: Elf der Studien wurden mit Primarschulkindern (KI. 1–6) durchgeführt, sieben mit Sekundarschuljugendlichen (KI. 7–12), die restlichen 21 mit Erwachsenen, mehrheitlich Studierenden.

C) Welche Instrumente kamen zum Einsatz?

Entscheidendes Kriterium war, dass in den Primärstudien Experimente durchgeführt worden waren, bei denen eine Gruppe einen Text ohne Bild erhielt, während eine weitere Gruppe denselben Text mit Bild(ern) bekam. Diese Variation in den unabhängigen Variablen wurde darauf geprüft, ob sie sich in den ausgewerteten Studien auf das dort zum

Einsatz gelangte Testinstrument zum Leseverstehen auswirkte, also ob Personen mit zusätzlicher Textbebilderung einen Vorteil im Leseverstehen hatten. Diese Effekte wurden mittels gewichteter Effektstärken (Hedges g) bestimmt und über alle 39 Studien gemittelt. Für die zweite Fragestellung wurden Merkmale der Studien kodiert. Neben der schon erwähnten Altersgruppendifferenzierung wurden weitere Merkmale erfasst: Textgenre (narrative Texte, expositorische Texte, Studien mit beiden Genres), Art des Bildes und Art der Testung. Bei der Art des Bildes wurden vier Varianten unterschieden: a) realistische Illustrationen, b) schematische Darstellungen, c) Flussdiagramme und d) Kombinationen mehrerer Bildvarianten. Die Tests wurden am feinteiligsten kodiert: a) Falsch-Richtig-Antworten, b) Multiple-Choice-Fragen, c) offene kurze Antworten (mündlich oder schriftlich), d) Kombinationen verschiedener Test und – als Restkategorie – e) sonstige Tests. Alter, Textgenres, Bildarten und Testvarianten waren potenzielle Moderatoren der Effektstärken. Die Moderatoranalysen wurden in dieser Studie als Metaregression durchgeführt, allerdings ohne die Textgenres einzubeziehen. Der Grund für den Ausschluss war, dass Bildarten und Textgenres zusammenhingen: Realistische Illustrationen waren fast ausschließlich bei den narrativen Texten vorfindbar. Die Wahl für die Metaregression begründet das Forschungsteam damit, gleichzeitig mehrere Moderatorvariablen statistisch zu berücksichtigen und damit Nettoeffekte auszuweisen.

D) Was waren die Hauptbefunde?

Der Haupteffekt und damit die Antwort auf die erste Fragestellung der Metaanalyse lautet: Personen, die Text-Bild-Kombination lasen, schnitten in den Tests gesamthaft leicht besser ab ($g = .39$). Der Zusatz von Bildern wirkt also verstehensförderlich.

Die Moderatoranalysen erbrachten keinen Effekt der Altersgruppe. Bei den eingesetzten Bildern erwiesen sich die einzelnen Bildtypen als gleichwertig, aber es gab in einem Fall einen statistisch signifikanten Vorsprung: Illustrierende Bilder waren Kombinationen mehrerer Bildarten überlegen. Auch das Testformat erwies sich als Moderator: Tests, die entweder offene Antworten verlangten oder sich auf mehrere Frageformate stützten, zeigten stärkere Vorteile als Tests, die mit Richtig-falsch-Format operierten. Komplexere Tests erwiesen sich also als sensibler für die positiven Effekte von Bildzusätzen bei Texten, während der Zusatz verschiedener Bilderarten diesen Vorteil zumindest im Vergleich mit illustrierenden realistischen Bildern nivelliert.

E) Diskussion

Wenn Texte mit zusätzlichen Bildern – in diesem Fall: in experimentellen Studien mit optimal konstruierten Bildern – ausgestattet werden, verstehen Leserinnen und Leser Textinhalte besser. Auffälligerweise war das nicht vom Alter der Personen abhängig. Die Metaanalyse sagt jedoch gerade bei den Moderatoranalysen relativ wenig darüber aus, welche Art von Verstehensleistungen davon besonders profitieren bzw. nicht profitieren. Offene Aufgabenformate bei der Erfassung des Leseverstehens ergaben deutlichere Vorsprünge der Personen, die Bild-Text-Kombinationen rezipierten, was sich als Hinweis

deuten lässt, dass kognitiv anspruchsvollere Verstehensleistungen eher von unterstützen- den Bildern profitieren können. Dies wäre aber noch mit weiteren Studien zu unterfüttern, die das Verhältnis von eingesetzten Bildern und Testungen fürderhin fokussieren sollten.

Multimediales Lernen und Bild-Text-Integration

Das Verstehen von Inhalten, die aus einer Kombination von sprachlichen und grafischen Inhalten stammen, geht über die typischen Anforderungen des Lesens von sprachlichen, kontinuierlichen Texten hinaus bzw. erfordert teils andere Prozesse. Damit wird zugleich die Frage aufgeworfen, was die Unterschiede im Verstehen von schriftbasierten Texten und grafisch kodierten Inhalten sind. Ausschlaggebend ist die Art der Repräsentation, wobei externe Repräsentationen als Kodierungen von verschiedenen Zeichensystemen den Ausgangspunkt bilden (Schnotz & Dutke, 2004):

- *Schriftliche Texte* fußen auf Symbolzeichen: Buchstaben (Grapheme) repräsentieren Laute (Phone), Wörter repräsentieren einzelne Konzepte etc. – wichtig ist hierbei, dass ein Symbolzeichen und der von ihm repräsentierte Inhalt nur durch eine linguistische Konvention verknüpft sind, aber Bezeichnetes und Bezeichnendes sonst keine Verbindung aufweisen. Den Inhalt repräsentieren sprachliche Texte auf eine deskriptionale Weise: Symbolzeichen transportieren ihren (propositionalen) Zusammenhang in den Konventionen der (Schrift-)Sprache. Ein Text wie dieser über die Verschiedenartigkeit von schriftlichen und bildlichen Texten benennt zum Beispiel Merkmale beider Arten sprachlichen und grafischen externen Repräsentationen und vergleicht sie.
- Bei *grafisch kodierten Inhalten* ist der Fall anders gelagert, und zwar hinsichtlich der Art der Repräsentation und der dafür verwendeten Zeichen. Es handelt sich um ikonische Zeichen, nicht Symbolzeichen, und wesenseigen ist diesen Zeichen, dass sie durch eine Ähnlichkeitsrelation mit dem, was sie repräsentieren, verbunden sind. Das gilt selbst für die unterschiedlichen Arten von Bildern. Grafische Darstellungen bzw. Bilder lassen sich in zwei Gruppen unterteilen, die spezifische Anforderungen an die lesende Person stellen. Zum einen sind „realistische“ Bilder anzuführen, also Gemälde, Zeichnungen und Fotografien, zum anderen „logische“ Bilder, ergo Diagramme. Bei realistischen Bildern ist die Ähnlichkeit ausschlaggebend, bei logischen Bildern die strukturelle Relation, also etwa die Skala einer Achse, deren Abstände für die Zunahme von Zeit-, Gewichts-, Temperatureinheiten oder Häufigkeiten steht, die Höhe einer Säule, die gemessen an der Achse das Bruttoinlandsprodukt anzeigt etc. Anders gesagt: Bilder enthalten in sich selbst Struktureigenschaften, die mit dem, was solche Bilder inhaltlich repräsentieren, übereinstimmen. Diese depiktionale Repräsentation mit ikonischen Zeichen

folgt damit anderen Konventionen als jene deskriptionale Repräsentation über Symbolzeichen.

Diese kursorisch dargelegten Unterschiede von verschiedenen Zeichensystemen und den daraus resultierenden (nicht-)konstitutiven Strukturrelationen zum Gegenstand von Text oder Bild betreffen die externe Manifestation, also die Art und Weise, wie sich ein außerhalb der lesenden Person befindlicher Text bzw. ein Bild seine Inhalte kodiert. Da es beim Lesen und Verstehen darum geht, eine interne Repräsentation aufzubauen, stellt sich die Folgefrage, ob sich diese Unterschiede der externen auch in der internen Repräsentation niederschlagen. Schnotz und Dutke (2004) modellieren das Verstehen von deskriptionalen und depiktionalen Repräsentationen als Mehrebenen-Konstrukt, bei dem sich einige Gemeinsamkeiten festhalten lassen, z. B. das mentale Modell und die hohe Bedeutung des Vorwissens für glückendes Verstehen. Allerdings gibt es auch Unterschiede. Das Äquivalent zur Textoberflächenrepräsentation bei Texten bildet bei Bildern die perzeptuelle mentale Repräsentation, bei der eine Person ein Bild aufgrund ihrer sensorischen Wahrnehmung repräsentiert, ohne notwendigerweise dessen Inhalt zu verstehen. Dies wiederum ist eine Gemeinsamkeit zur Textoberflächenrepräsentation.

Aufgrund der in Bildern explizit enthaltenen, für diese Form der Darstellung konstitutiven Relationen der Struktur, die direkt aus dem Bild ablesbar ist, und der ikonischen Zeichen wird beim Verstehen von Bildern – anders als bei Texten – angenommen, dass die nächste internale Repräsentationsform bereits das mentale Modell ist. Beim Textverstehen wird die propositionale Repräsentation als mittlere Ebene postuliert, welche bereits Integrationsprozesse nötig macht. Bei bildlichen Repräsentationen gehen hingegen eine Repräsentation als mentales Modell und eine semantische Analyse, deren Ergebnis der propositionalen Repräsentation beim Textverstehen ähnelt, Hand in Hand und erfolgen parallel. Textuelles und bildliches Verstehen haben also vergleichbare Ebenen, allerdings ist die Positionierung der propositionalen Repräsentation unterschiedlich gelagert.

Verarbeitungsprozesse des Verstehens und deren Ergebnisse in Form von internen mentalen Repräsentationen haben somit Gemeinsamkeiten und Unterschiede, die sich an der bislang leitenden separaten Betrachtung von externen deskriptionalen und depiktionalen Repräsentationen konturieren lassen. Aus der Sicht des erweiterten Verständnisses von Lesekompetenz kommt es jedoch darauf an, wie sich sowohl aus sprachlichen als auch aus bildlichen Texten eine *gemeinsame mentale Repräsentation* ergibt. Hierzu gibt es bereits verschiedene theoretische Vorschläge, die sich darin unterscheiden, welche Art von Gedächtnisleistungen sie beim Verstehen von Kombinationen von Texten und Bildern fokussieren, sich aber stark darin gleichen, dass sie von verschiedenen sensorischen Verarbeitungskanälen ausgehen und das Verfügbarhalten von visuell-depiktionalen und sprachlich-deskriptionalen Inhalten im Arbeitsgedächtnis mit dessen

Mehr-Komponenten-Aufbau vereinen (Reed, 2006). Zwei dieser Modelle werden an dieser Stelle konsultiert und präsentiert. Dabei handelt es sich um die *kognitive Theorie des Multimedia-Lernens* von Mayer (2022a) als ein allgemeines Modell und das *integrative Modell des Text- und Bildverstehens* von Schnotz (2022).

Die kognitive Theorie des Multimedia-Lernens

Eine der einflussreichsten Theorien im multimedialen Lernen stammt von Richard Mayer (2002a), der diese Theorie seit den 1980er Jahren entwickelt und verfeinert hat (Mayer, 2021). Er definiert *multimediales Lernen* als den *Aufbau einer internalen mentalen Repräsentation auf der Basis einer Kombination von externen sprachlichen und bildlichen Repräsentationen*. Es wird angenommen, dass Personen durch diese Kombination von Bild und Text bessere Lern- und Verstehensleistungen erbringen, als wenn sie Inhalte rein sprachlich präsentiert erhalten bekommen (Mayer, 2022b).

Das Elementare dieser Definition besteht darin, dass Personen eine einzelne Repräsentation aufbauen, die auf zwei verschiedenen Zeichensystemen basiert, die explizit auch mündliche und bewegte Bilder inkludiert. Die obige Definition des multimedialen Lernens greift die kognitive Theorie des Multimedia-Lernens konsequent auf, indem sie aus diversen anderen Theorien ein deskriptives Modell entwickelt, welches sich durch drei konstitutive Prinzipien auszeichnet, die sich im Modell deutlich niedergeschlagen haben. Diese Annahmen sind mit Mayer (2002a):

- *Die Annahme zweikanaliger Informationsverarbeitung*: Diese Annahme inkludiert, dass es zwei Modalitäten gibt, die sensorisch wahrgenommen und kognitiv bearbeitet werden, d. h. sprachliche Äußerungen werden (mündlich) über den Gehörsinn aufgenommen, bildliche (auch: Bewegtbilder) über die Augen. Sensorisch aufgenommene Informationen zudem in separaten Komponenten des Arbeitsgedächtnisses verarbeitet und verfügbar gehalten.
- *Die Annahme der begrenzten Informationsverarbeitungskapazität des Arbeitsgedächtnisses*: Gegenstand dieser Prämisse ist, dass Lernen ohne Nutzung des Arbeitsgedächtnisses nicht denkbar ist, dass aber die Besonderheiten dieses Bestandteils des menschlichen kognitiven Systems zu berücksichtigen sind (s. dazu Teilkap. [6.1.3.2](#)). Zwei Merkmale sind besonders wichtig: Die Kapazität des Arbeitsgedächtnisses ist erstens begrenzt, zweitens ist es aber auch jener Ort, an dem neue Informationen aus der Umwelt mit jenen aus dem Langzeitgedächtnis verbunden und aufeinander bezogen werden. Speziell die Anzahl von und Dauer des Zugriffs auf Informationen stellen Limitationen insofern dar, als es darum geht, diese Kapazitäten möglichst günstig zu nutzen, worauf sich ein Großteil der Anwendung der Theorie im Sinne einer Gestaltung möglichst lernförderlicher bzw. aus der Perspektive

kognitiver Lernprozesse optimierter Text-Bild-Kombinationen bezieht (s. dazu Teilkap. 4.4.2.3).

- *Die Annahme aktiver Informationsverarbeitung beim Lernen:* Internale Repräsentationen und Lernen, zumal wenn es sich um bedeutungsvolles Lernen bzw. tiefes Leseverstehen handelt, erfordern seitens der lernenden Person eine Transformation der Informationen aus der Umwelt. Die Theorie stützt sich auf drei grundsätzlich nötige Prozesse: a) die Auswahl von Informationen, b) deren strukturierende Organisation und die bereits im Rahmen dieses Teilkapitels ausführlicher behandelte c) Integration von Informationen. Dabei werden modalitätsspezifische Prozesse des Auswählens und Organisierens angenommen, während für das Integrieren von depiktionalen und deskriptionalen Inhalten in ein vereintes modalitätsunspezifisches mentales Modell die Integrationsprozesse nötig sind.

Entlang der drei Grundannahmen lässt sich der Aufbau der Theorie und vor allem die Logik der Informationsverarbeitungsprozesse unter Nutzung von Langzeit- und Arbeitsgedächtnis beschreiben. Das Modell ist in Abbildung 11 dargestellt, und in diesem Modell sind oben die externalen multimedialen, also verbal-sprachlichen und non-verbal-bildlichen Präsentationen und vor allem die drei internalen Gedächtnissystemen gefettet hervorgehoben. Darunter in den Kästen sind depiktionale und deskriptionale Repräsentationen in ihrer fortlaufenden kognitiven Verarbeitung dargestellt, wobei die Kästen in ihrer horizontalen Reihung die Zwei-Kanal-Informationsverarbeitung reflektiert. Mit kursiven Ausdrücken entlang der Pfeile sind im Modell die Prozesse der aktiven Informationsverarbeitung bezeichnet, welche sich einerseits aus der Verarbeitung sensorisch aufgenommener Informationen und der Langzeitgedächtnisinhalte (Vorwissen) im Arbeitsgedächtnis vollziehen. Diese Prozesse münden in verschiedene interne Repräsentationen mit variierendem Transformationsgrad.

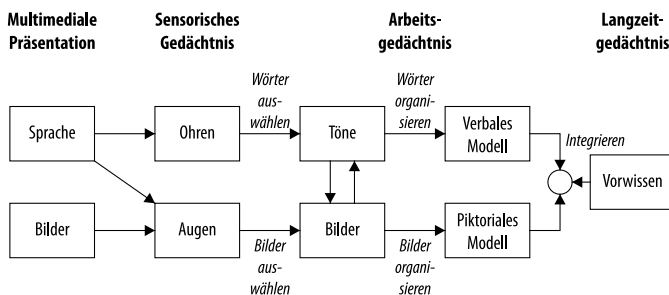


Abbildung 11: Die kognitive Theorie des multimedialen Lernens
(Quelle: Übersetzung von Mayer, 2022a, S. 62)

In der Abbildung 11 von links nach rechts gehend folgt das Modell also einer Informationstransformation von externalen Repräsentationen hin zu verschiedenen internalen Repräsentationen. Die folgende knappe Darstellung in Form einer Reihung wirkt aus darstellerischen Gründen als rein lineare Sequenz, Rückkopplungen und eine größere Dynamik sind jedoch möglich:

- In einem ersten Schritt nehmen Personen die deskriptionalen und depiktionalen Präsentationen aus der Umwelt über ihren Gehör- und Sehsinn wahr, wobei sprachliche Repräsentationen in der mündlichen Modalität akustisch und in ihrer schriftlichen Modalität über die Augen visuell wahrgenommen werden und im sensorischen Gedächtnis kurze Zeit verfügbar gehalten werden.
- Aus diesen flüchtigen Inhalten erreichen über Prozesse des *Auswählens* regelbasiert selektierte Ausschnitte das Arbeitsgedächtnis, nämlich die phonologische Schleife bei Tönen bzw. den visuellen Notizblock bei Bildern. Die Doppelpfeile von den Bildern und den Tönen im Arbeitsgedächtnis stellen zudem Wechsel dar, indem visuellen phonische Informationen zugeordnet bzw. diese gedächtnisbasiert koaktiviert werden.
- Diese Informationen im Arbeitsgedächtnis werden dann über verschiedene, das Vorwissen nutzende Prozesse in eine Struktur überführt, indem zentrale Verbindungen zwischen einzelnen deskriptionalen und depiktionalen Inhalten im Arbeitsgedächtnis erstellt werden, wobei Personen auf Konventionen von Bildern und Textsorten zurückgreifen. Das Ergebnis dieser *Organisations*prozesse ist ein piktoriales bzw. ein verbales Modell. Das verbale Modell ist mindestens mit der Textbasis, also einer propositionalen Repräsentation gemäß Construction-Integration-Modell gleichzusetzen. Das piktoriale Modell als nicht-sprachliche Repräsentation spezifiziert Mayer (2022a) nicht weiter, spricht aber von einer kohärenten Repräsentation aufgrund piktorialer Inhalte, welche mit Vorwissensbeständen angereichert sind.
- Der letzte Prozess ist das *Integrieren*, wobei hier explizit Vorwissen und die beiden separaten mentalen Modelle, das verbale und das piktoriale Modell, aufeinander bezogen werden. Dafür sind inferenzbasierte Verknüpfungen von zueinander gehörigen Elementen aus beiden Modellen nötig, damit aus zwei separaten ein gemeinsames kognitives Modell entsteht.

Die kognitive Theorie des Multimedia-Lernens ist damit ein sehr allgemeines Modell, gleichwohl ist es bereits aufgrund der Vielzahl von in- und externalen Repräsentationen bzw. Präsentationen und der drei kognitiven Prozesse komplex. Die Prämissen der kognitiven Theorie des Multimedia-Lernens kompilieren Vorannahmen und Bestandteile anderer Theorien, sodass der kognitiven Theorie des Multimedia-Lernens der Status einer Metatheorie zugewiesen werden kann, in der einzelne kleinteiligere Theorien verortbar sind. Das gilt aus der Sicht des

Lesens vor allem aus der hier besonders interessierenden Prozessperspektive für die Transformationsprozesse, bei denen Wittrocks (1989) generatives Lernen Pate stand für die Beschreibung der zentralen Prozesse des Auswählens, Organisierens und Integrierens. Das ist deshalb so zu betonen, weil das Integrieren und das Organisieren in ihrem Zusammenspiel als bedeutsam für das tiefe Verstehen von Sachverhalten gelten (Mayer, 1996).

Ihren Wert hat die kognitive Theorie des Multimedia-Lernens darin, entlang verschiedener Prämissen eine Logik der Verarbeitung von bildlichen und sprachlichen Inhalten beginnend von der sensorischen Wahrnehmung und modalitäts(un)spezifischen Verarbeitungen auf der Basis von kognitionspsychologischen Erkenntnissen kartiert zu haben und dabei einen breiten Geltungsbereich aufzuweisen. Gleichwohl ist die kognitive Theorie des Multimedia-Lernens eher grobkörnig, weil sie z. B. nicht spezifiziert, welche Art von mentalem Modell entsteht. Hier braucht es also mikrotheoretische Vertiefungen, und ein Modell, das sich hierfür anbietet, ist das im Folgenden dargestellte integrative Modell des Text- und Bildverstehens.

Das integrative Modell des Text- und Bildverstehens

Schnotz' (2014, 2022) integratives Modell des Text- und Bildverstehens teilt viele der Grundannahmen von Mayers im Abschnitt zuvor vorgestelltem Modell, darunter insbesondere die mehrkanalige transformierende Informationsverarbeitung bei der Mehrebenen-Repräsentation von Inhalten und die hohe Bedeutung des Arbeitsgedächtnisses für Verstehensprozesse. Besonders deutlich wird das in der Darstellung des Modells in Abbildung 12 daran, dass das Modell neben der Trennung von kognitiver Verarbeitung von Text (links) bzw. Bild (rechts) auch verschiedene Repräsentationsformen expliziert. Dieses Modell ist ursprünglich das genuine Modell des integrativen Bild- und Textverstehens gewesen, Schnotz (2022) hat es in seiner aktuellsten Version lediglich als Rahmen bezeichnet und ein neues Modell vorgelegt, das sich deutlich gegenüber auditiven Elementen öffnet und sprachliches Hörverstehen ebenso berücksichtigt wie das Verstehen von nicht-sprachlichen Inhalten. Dieses neue Modell ist erheblich detaillierter hinsichtlich der Komponenten und wird nachfolgend aus Platzgründen nicht entfaltet.

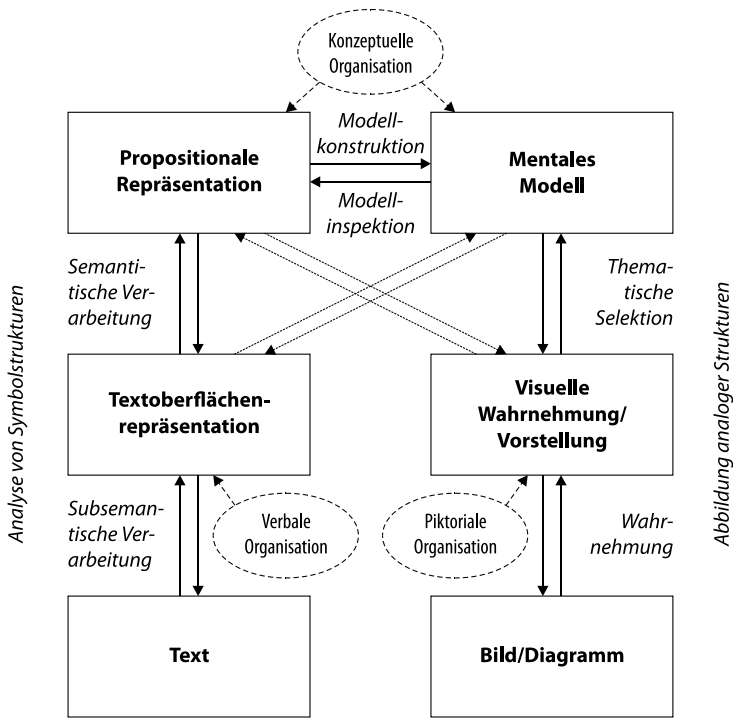


Abbildung 12: Das integrative Modell des Text- und Bildverstehens (Quelle: Übersetzung von Schnotz, 2022, S. 86)

Das integrative Modell des Text- und Bildverstehens nimmt erkennbare Anleihen an dem Construction-Integration-Modell, was sich im linken Teil der Abbildung 12 niederschlägt, nämlich dem, was in der Darstellung „Analyse von Symbolstrukturen“ genannt wird und für das Verstehen deskriptionaler Inhalte reserviert ist. Das Modell geht von der externalen Repräsentation, also einem Text, aus. Die darüber abgebildeten Textoberflächen- und propositionale Repräsentationen nebst mentalem Modell (rechts oben) entsprechen den drei Ebenen des Leseverstehens, wie sie im Teilkapitel [4.1.1.1](#) vorgestellt wurden und über das Zusammenspiel automatisch ablaufender Koaktivierung von Gedächtnisinhalten aufgrund des gelesenen Inhaltes und Konstruktions- und Integrationsprozessen interaktiv und dynamisch zustande kommen. Im Modell werden diese Prozesse der Informationstransformation als subsemantische und semantische Verarbeitung sowie als Modellkonstruktion bezeichnet. Ebenfalls im Modell ist vermerkt, dass verbale Strukturen bei der Textoberflächenrepräsentation einflussreich sind, während propositionale Repräsentation und mentales Modell auf konzeptuellen, also inhaltlich-thematischen Strukturen basieren.

Rechts im Modell ist das Bildverstehen, ausgehend von den depiktionalen Repräsentationen, verortet. Der Ausdruck „Abbildung analoger Strukturen“

verweist nicht auf analoge (vs. digitale) Bilder, sondern auf das konstitutive Merkmal von Bildern, dass Bild und Abgebildetes über eine Analogie in der Struktur oder im Inhalt miteinander verbunden sind, die für den Verstehensprozess bedeutsam ist. Wie sehr das der Fall ist, ist daran ersichtlich, dass die piktoriale Organisation, also die bildliche Struktur, eine visuelle Wahrnehmung mitsteuert, welche erste interne Vorstellung jemand aufgrund erkannter und voneinander unterschiedener Elemente aufbaut. Diese Vorstellung ist noch recht nah an der externalen Repräsentation. Ein mentales Modell stellt sie noch nicht dar, dazu bedarf es sogenannter Mapping-Prozesse, die bildliche und räumliche Elemente aus der visuellen Wahrnehmung mit semantischen Bezügen aus dem Vorwissen versehen. Damit geht – wie auch beim Verstehen von deskriptionalen Repräsentationen – eine auswahlbedingte Verknappung der Inhalte aus den Bildern einher, darum ist im Modell von einer thematischen Selektion die Rede. Diese thematische Selektion muss nicht nur darauf basieren, dass jemand ein Bild verstehen will. Vielmehr sind anforderungsbezogene Kriterien, also Relevanz für Zielerreichungen (s. Teilkap. 2.1.3), eine weitere Bemessungsgrundlage dafür, was eine lesende Person auswählt. Dies wurde insbesondere in den frühen Publikationen zum Modell dezidiert betont (Schnotz & Bannert, 1999). Das mentale Modell ist, selbst wenn es sich ursprünglich auf depikcional wahrgenommene Inhalte bezieht, jedoch kein reines Abbild externaler Inhalte, sondern in seiner Modalität unspezifisch, abstrahiert und mit Vorwissensbeständen angereichert.

Im Modell des Text- und Bildverstehens ist nur ein internes mentales Modell als Ergebnis der Verstehensprozesse enthalten, welches sich aus transformierten deskriptionalen und depikcionalen Inhalten und dem Vorwissen der lesenden Person speist und im Arbeitsgedächtnis präsent gehalten wird. Dieses mentale Modell enthält damit Informationen aus drei Quellen: dem eigenen Vorwissen und den aus den sowohl bildlichen als auch textuellen Bestandteilen stammenden Informationen in ihrem Zusammenhang. Die Bildung des mentalen Modells nimmt auch dieser theoretische Zugang als dynamisch an, was sich im oberen Teil der grafischen Darstellung in dem Doppelpfeil von Modellkonstruktion und -inspektion niederschlägt. Die Modellinspektion ist die neue Ergänzung, die besonders aus der Theorie zum Bildverstehen stammt. Schnotz und Bannert (1999) führen an, dass mentale Modelle, die auf der Basis depikcionaler (und auch deskriptionaler) Inhalte gebildet wurden, eine Grundlage dafür darstellen, weitere Propositionen in der entsprechenden Repräsentationsform in rekursiven Verarbeitungsschleifen zu ergänzen, wenn es nötig ist. Diese Propositionen können aus dem mentalen Modell selbst stammen, aber auch aus der erneuten lesenden Verarbeitung von Text und Bild. An dieser Stelle kulminieren also verarbeitete Informationen im mentalen Modell und lassen ihre Ausgangsmodalität hinter sich.

Damit impliziert das integrative Modell des Text- und Bildverstehens zugleich ein wichtiges Zusammenspiel (Schnotz, 2022): Es nimmt an, dass sensorisch als Bild bzw. Text aufgenommene Informationen zunächst modalitätsspezifisch im

Arbeitsgedächtnis verarbeitet werden, dass aber für eine modalitätsunspezifische Bildung eines mentalen Modells von Text und Bild ein Transfer von sprachlichen in nicht-sprachliche und von nicht-sprachlichen in sprachliche Repräsentationen erfolgt. Diese Übersetzung postuliert das Modell als Notwendigkeit für ein tiefes Verstehen des in deskriptionalen und depiktionalen externen Repräsentationen enthaltenen Inhalts.

4.1.2.3 Multiple Dokumente integrieren

Wenn Personen sich über Themen im Internet informieren, stoßen sie auf verschiedene Informationen aus verschiedenen Dokumenten. Dies bringt es mit sich, dass aus einer Kontingenz der Dokumente und ihrer Informationen eine Kohärenz in der mentalen internalen Repräsentation erwachsen soll. Die Fähigkeit dazu wird als „Multiple Documents Literacy“ bezeichnet, worauf dieses Teilkapitel in einem ersten Schritt der Darstellung eingeht. Die aus den Prozessen des Verstehens resultierenden internalen Repräsentationen variieren darin, welche Art von Information (eigentliche Inhalte und Informationen über Informationen) wie stark verbunden werden, was in einem zweiten Schritt entfaltet wird. Häufig müssen Personen nachvollziehen und interpretierend rekonstruieren, welche Perspektiven Dokumente einnehmen. Dies ist Gegenstand des dritten Schrittes. Und schließlich führt das erst im Prozess des Lesens zu klärende Verhältnis der Dokumente untereinander dazu, dass Leserinnen und Leser mit verstehenserschwerender Inkonsistenz und Inkohärenz umgehen können müssen, was über das Inferieren geklärt werden kann. Diese anspruchsvolle Fähigkeit wird im vierten Schritt der Darstellung fokussiert.

1) *Multiple Documents Literacy*

Dass Personen nicht nur Texte, sondern auch Kombinationen von Bildern und Texten lesen, ist durch die technischen Aspekte des multimedialen Lernens nicht nur möglich, sondern naheliegend (Mayer, 2002a). Eine weitere Besonderheit des Lesens vor allem im Internet besteht darin, dass Personen nicht mehr nur einen singulären, möglichst kohärenten Text bzw. eine Kombination von deskriptionalen und depiktionalen Bestandteilen lesen, sondern dass sie vielmehr auf diverse Dokumente stoßen, die Themen unterschiedlich behandeln (z. B. Folgen des Impfs, über welches sich Laien informieren wollen), deren Verhältnis untereinander nicht sofort evident ist und bei denen Leserinnen und Leser aktiv diverse Inferenzen über die Grenzen der einzelnen Dokumente hinweg bilden müssen, um zu einem Kohärenzaufbau zu gelangen. Die Besonderheit dieses Lesens wird als *Lesen multipler Dokumente* bezeichnet und sowohl prozess- als auch produktbezogen theoretisch modelliert und empirisch erforscht (Primor & Katzir, 2018).

Der Ausdruck „*Dokumente*“ ist im Zusammenhang mit dem Lesen mehrerer Texte nicht zufällig gewählt, aber explikationsbedürftig. Mit „Dokument“ wird eine Kombination von zweierlei Informationsarten gemeint. Das ist zum einen der Inhalt, der in deskriptionaler oder depiktionaler Modalität kodiert wird (Rouet & Britt, 2022), gemeinhin das, was als schriftlicher oder bildlicher Text gemeint ist. Zum anderen, und das ist das Besondere, kommen weitere potenziell bedeutungsmodifizierende Informationen hinzu, die man als „Metadaten“ bezeichnet, wobei es auch andere Bezeichnungen wie „Quellenmerkmale“ gibt (Brante & Strømsø, 2018). Unabhängig von der jeweils verwendeten, in der Forschung keineswegs einheitlichen Terminologie (Goldman & Scardamalia, 2013) ist für die Metadaten bedeutend, dass es sich um unterschiedlich gut zugängliche, in jedem Fall jedoch der Interpretation bedürftige Informationen über Informationen handelt, welche für gewöhnlich über den Prozess des Sourcing beigesteuert werden (Bråten et al., 2017, s. Teilkap. [3.1.2.1](#)). Solche Metadaten beziehen sich beispielsweise auf Merkmale der Person bzw. Institution, welche inhaltliche Aussagen getätigt haben, den Kontext der Informationen wie Erscheinungszusammenhang, aber auch zum Beispiel den Stil eines Dokuments (Goldman & Scardamalia, 2013).

Für das Lesen und Verwenden von multiplen Dokumenten und ihren beiden bedeutungstragenden Komponenten, Inhalt und Metadaten, hat sich international der Begriff „*Multiple Documents Literacy*“ etabliert. Anmarkrud et al. (2014, S. 64) reservieren ihn für die Fähigkeit, *verschiedene Informationsquellen zu finden, sie zu bewerten und so zu nutzen, dass man eine integrierte, bedeutungsvolle Repräsentation eines bestimmten Problems, Themas oder einer Situation konstruieren und kommunizieren kann*. Aus dieser Definition geht hervor, dass die Herstellung eines kohärenten mentalen Modells den Kern der Multiple Documents Literacy bildet, wofür ein Prozessverbund vor allem im intra- und intertextuellen Integrieren benötigt wird (Barzilai et al., 2018; Saux, Britt et al., 2021).

Bisherige Modelle zur Multiple Documents Literacy gewichten das Integrieren unterschiedlich, weisen allerdings die Gemeinsamkeit auf, dass inferenzbasierte Integrationsleistungen für den Aufbau intra- (einzelner) und intertextueller (dokumentenübergreifender) mentaler Modelle zum Kernbestand zählen (List & Alexander, 2017a; Philipp, 2018b). Dies wird an einer Aufzählung von vier zentralen Fähigkeiten der Multiple Documents Literacy deutlich, welche Britt und Rouet (2020) stellvertretend für andere Synopsen der Leseforschung (Cho, Afflerbach & Han, 2018; Philipp, 2020a; Rouet & Britt, 2022) vornehmen:

- 1) die Fähigkeit, Inferenzen innerhalb einzelner Texte (intratextuelles Integrieren) und zwischen Dokumenten (intertextuelles Integrieren) zu bilden;
- 2) die Fähigkeit, Informationen aufgaben- bzw. fragebezogen auszuwählen und zu strukturieren;

- 3) die Fähigkeit, relevante bzw. wichtige Informationen selektiv mithilfe externer Hinweise bzw. des sich aufbauenden mentalen Modells dafür zu verwenden, ein angemessen kohärentes Gesamtverstehen aufzubauen,
- 4) und die Fähigkeit, separate mentale Modelle einzelner Dokumente zu repräsentieren, wenn sich keine Gesamtkohärenz über alle gelesenen Dokumente herstellen lässt.

Aus der Liste oben lässt sich ablesen, dass das Lesen multipler Dokumente als Mehrebenen-Prozess und als Mehrebenen-Produkt beschreiben lässt. Auch wenn dieses gesamte Teilkapitel die Prozesse besonders fokussiert, so ist es dennoch nötig, die Besonderheiten des Lesens multipler Dokumente anhand der Produktperspektive zu umreißen.

2) *Das Dokumentenmodell als Zielvorstellung glückender intertextueller Integrationsleistungen*

Ein Initialzünder für die Theoriearbeit im Bereich der Multiple Documents Literacy war ein im Jahr 1999 vorgelegtes Modell, welches in zwei Veröffentlichungen vorgestellt wurde (Britt et al., 1999; Perfetti et al., 1999): das „Dokumentenmodell“. Das Dokumentenmodell ist eine direkt auf den Arbeiten des Construction-Integration-Modells aufbauende und damit kognitionspsychologische Beschreibung jenes Verstehensprodukts, wenn Personen mehrere Dokumente lesen. Seinen Ursprung hat das Modell in der Domäne Geschichte, wobei es die disziplinären Besonderheiten des Lesens in dieser Domäne explizit aufgreift. Erstens werden Sachverhalte durchaus sehr unterschiedlich und widersprüchlich in verschiedenen Dokumenten geschildert und eingeschätzt. Zweitens sind komplexe Kausalitätsabfolgen von Ereignissen typisch, deren Verständnis entscheidend ist, um den Sachverhalt angemessen kognitiv zu repräsentieren. Es handelt sich also um ein Lesen mit hohem Schwierigkeitsgrad, was sich darin zeigt, dass Expertinnen und Experten in solchen Lesesituationen auf diverse lesestrategische Vorgehensweisen zurückgreifen, um lesebezogene Ziele im Sinne des tiefen Leseverstehens (s. Teilkap. [4.1.1.2](#)) zu erreichen, darunter die mentale Modellbildung (Wineburg, 1991; Wineburg & McGrew, 2019).

Das Dokumentenmodell beinhaltet die Grundannahme, dass Dokumente nicht kontextlos zu verstehen sind, sondern in Diskurse eingebettet sind, worunter fällt, dass jemand sie mit einer Absicht zu einem Zeitpunkt an einem Ort hergestellt hat. Darum sprechen sich Britt et al. (2013) dafür aus, Dokumente nicht nur als Einheiten aus Text und mehr oder minder leicht per Sourcing mit den Inhalten rekonstruierbare Metadaten zu begreifen. Vielmehr plädieren sie sogar dafür, Dokumente als *soziale Artefakte* bzw. als *Teilnehmer in Diskursen* zu konzeptualisieren. Damit weitet sich der Blick erheblich: Er richtet sich nämlich nicht mehr nur auf ein akontextuelles Verstehen von einem Text und einer

lesenden Person und ihrem Vorwissen, wie es das Construction-Integration-Modell als Stellvertreter für kognitionspsychologische Modellierungen des Lesens gemeinhin tut (McNamara & Magliano, 2009). Stattdessen ist es demzufolge nötig, mehr und andere Informationen, nämlich zum Kontext von Aussagen und zur kommunikativen Absicht von Autorinnen und Autoren (Schnotz & Dutke, 2004), systematischer zu berücksichtigen, was als Desiderat in bisherigen Leseverstehensmodellen gilt (McNamara & Magliano, 2009). Dies greift das Dokumentenmodell insofern auf, als es eine eigene Repräsentationsebene postuliert, in der solche auf Evaluationen und Schlussfolgerungen basierenden leseenseitigen Einschätzungen mindestens implizit enthalten sind, nämlich die Ebene des Intertextmodells, welches Metadaten enthält (Perfetti et al., 1999).

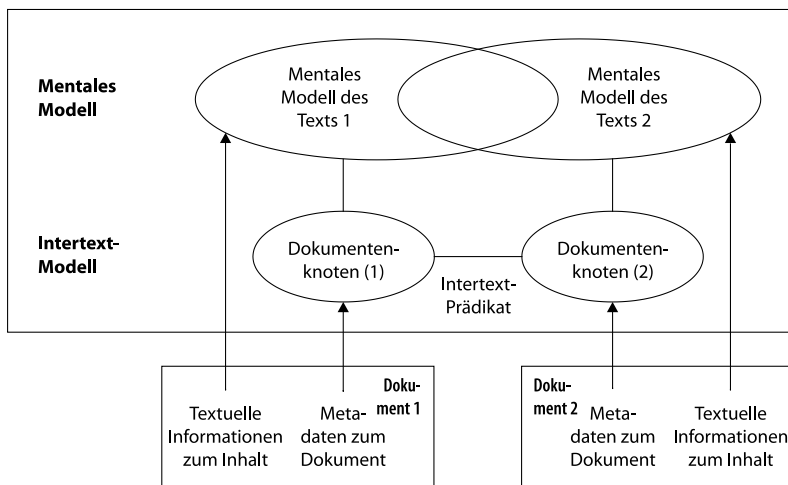


Abbildung 13: Das Dokumentenmodell und seine zwei Repräsentationsebenen, die sich aus verschiedenen Arten von Informationen aus Dokumenten speisen (Quelle: leicht modifizierte Darstellung gemäß Britt & Rouet, 2012, S. 285; Hinweis: Aus darstellerischen Gründen sind Oberflächen- und propositionale Repräsentation der Inhalte aus den Inhalten der textuellen Bestandteile in den Dokumenten 1 und 2 nicht dargestellt)

Was ist das Dokumentenmodell? Britt et al. (2018, S. 201) definieren es als *eine von der lesenden Person generierte interne Repräsentation sowohl des Inhalts einzelner Dokumente und integrierter, textübergreifender Informationen, die um die lesebezogene Anforderung herum organisiert sind (im integrierten mentalen Modell), als auch aller Metadaten über die verwendete(n) und abgelehnte(n) Dokumente*. Das Dokumentenmodell wird als Zwei-Ebenen-Repräsentation verstanden (s. Abbildung 13):

- Diese Repräsentation umfasst erstens auf der Ebene des mentalen Modells *das integrierte mentale Modell*, welches die konstruierte Bedeutung einzelner

Dokumente (genauer: deren Inhalte aus den Texten) enthält, und zwar sowohl der einzelnen Texte als auch der multiplen Texte. Je nachdem, wie stark diese Textinhalte überlappen, ist das integrierte mentale Modell eines, das eine große Schnittmenge aufweist und um die thematischen intertextuellen Makrostrukturen organisiert ist bzw. bei nur geringfügiger Überlappung eher lose verbundene Informationen aus allen multiplen Texten umfasst. Die mentalen Modelle der einzelnen Texte – unabhängig davon, ob und wie stark sie untereinander verbunden sind, – sind ebenfalls auf dieser Ebene zu verorten (Britt et al., 2018; Perfetti et al., 1999). Ein fiktives Beispiel hierfür können zwei Internetdokumente sein, in denen die Gefahren und Chancen von mRNA-Impfstoffen gegen Covid-19 Gegenstand sind. Solche Einschätzungen können Nebenwirkungen betreffen, die relative Wirksamkeit einzelner Impfstoffe, Herstellungsort und genaue Wirkmechanismen. All dies wären intertextuelle Makropropositionen im Sinne von Teilthemen der in beiden Internetdokumenten behandelten Thematik Covid-19-Impfstoff.

- Die zweite Ebene ist das *Intertextmodell*, welches sich auf die Dokumente und deren repräsentierter struktureller Relation untereinander bezieht. Das Intertextmodell enthält pro Dokument diverse Metadaten, welche eine Person zusätzlich zum Inhalt repräsentiert. Diese Metadaten enthält der sogenannte „*Dokumentenknoten*“. Zusätzlich zu den Metadaten beinhaltet der Dokumentenknoten noch die Hauptaussage aus dem Text des jeweiligen Dokuments. Diese Hauptaussage ist vor allem dazu nötig, das intertextuelle Verhältnis zwischen den Texten zu klären, was im Dokumentenmodell als „*Intertext-Prädikat*“ bezeichnet wird. Die Dokumentenknoten sind, wo sich die semantischen Verhältnisse ganzer Dokumente bestimmen lassen, also miteinander vernetzt sind, horizontal miteinander verquickt. Auch vertikal bestehen Beziehungen, die in der Abbildung 13 dargestellt sind und die beiden Ebenen des mentalen Modells und des Intertextmodells miteinander verbinden („*Quelle-Inhalt-Verbindungen*“). Im fiktiven Beispiel des Impfstoffs können zum Beispiel inhaltlich stark divergierende Aussagen zu den Nebenwirkungen auf zwei Websites getätigt worden sein, die beispielsweise mRNA-Impfstoffe als harmlos fürs menschliche Genom darstellen oder das Gegenteil behaupten. Wenn Personen diesen Widerspruch erkennen, dann hilft ihnen das bei der Explikation des Intertext-Prädikats, welches im Beispiel ein mindestens partieller Widerspruch zwischen den Dokumenten wäre; wenn sich die Dokumente noch stärker widersprechen, wäre sogar ein überwiegender oder gänzlicher Widerspruch die Bezeichnung für das Intertext-Prädikat. Daran deutet sich bereits an, dass Intertext-Prädikate Ergebnisse hochinferenten Leseverstehens darstellen, was sich auch in ihrer allgemeinen Funktion, der reduzierten kognitiven Repräsentation der Struktur der Dokumente als Netzwerk untereinander, ergibt. Hinzu kommen noch die Dokumentenknoten, die mit den einzelnen mentalen Modellen der textuell basierten Informationen vertikal

verknüpft sind. Diese Dokumentenknoten umfassen Metadaten, die dabei helfen, die Perspektiven einer Person, Personengruppe oder einer Institution auf den im Dokument beschriebenen Sachverhalt zu rekonstruieren. Wenn z. B. die eine stark impfstoffkritische Website von einer Bürgergruppe stammt, die aus Privatpersonen ohne erkennbare medizinische Ausbildung besteht, die impfstofffreundlichere Website hingegen vom Paul-Ehrlich-Institut als wissenschaftliche Institution von unabhängigen Forscherinnen und Forschern betrieben wird, dann modifiziert dies – natürlich in Verbindung mit Individualmerkmalen der lesenden Person, darunter thematischen Überzeugungen (s. Teilkap. 4.3.2) – nicht nur Plausibilität und Glaubwürdigkeit des Inhalts und dessen Ursprung. In solchen keineswegs untypischen Situationen bilden die Metadaten gleichfalls eine heuristische Grundlage dafür, Gewichtungen bei der Konstruktion des integrierten mentalen Modells auf der darüberliegenden Ebene im Dokumentenmodell vorzunehmen, weil jemand auf der Basis der Metadaten die Dokumente unterschiedlich priorisiert und sogar ganze Dokumente begründet ausschließen kann.

Die grundsätzliche Architektur des Dokumentenmodells ist auch heute noch in diversen Modellen des Verstehens und des Integrierens von Informationen aus multiplen Dokumenten implizit oder explizit enthalten (Philipp, 2018b). Dabei haben sich die Perspektiven darauf verändert, wie das Dokumentenmodell zustande kommt. Insbesondere ist viel deutlicher betont worden, dass Kontext- und Aufgabeneffekte erheblich mitsteuern, welche Informationen lesende Personen selektieren und in die Konstruktion komplexer mentaler Modelle aufnehmen (Britt & Rouet, 2012; Rouet et al., 2017; s. Kap. 2). Zudem wurde von Anfang an von der Forschungsgruppe hinter dem Dokumentenmodell darauf hingewiesen, dass es sich beim Dokumentenmodell lediglich um *eine* Variante handelt, wie Personen multiple Dokumente repräsentieren (Britt et al., 1999). Dabei ist das Dokumentenmodell zwar eine anspruchsvolle, aber noch nicht *die* anspruchsvollste Variante einer internalen mentalen Repräsentation (s. dazu Teilkap. 4.4.1). Die kognitiv aufwändigste und Leserinnen und Lesern mit hoher Expertise vorbehaltene Variante, das sogenannte „Modell mit allen Dokumentenetiketten“, geht über eine reine Verbindung von Dokumentenknoten und jeweiligem einzeltextbasierten mentalen Modell hinaus. Im Modell mit allen Dokumentenetiketten ist jeder inhaltliche Bestandteil auf der Ebene des integrierten mentalen Modells mit seinem Ursprung verbunden, also vollständig etikettiert. Die Menge an Informationen und ihrer Vernetztheit, die dabei eine lesende Person im Arbeitsgedächtnis präsent zu halten hat, ist daher hoch und mutmaßlich über eine vorwissensbasierte effektive Nutzung von Wissensbeständen im Kurz- und Langzeitkurzgedächtnis erklärbar (Ericsson, 2018).

Mit einem Spektrum verschiedener Varianten von mentalen Repräsentationen multipler gelesener Dokumente besteht eine Anschlussfähigkeit an die

Kohärenzstandards (s. Teilkap. [2.1.2](#)). Dokumentenmodelle erfordern vonseiten der lesenden Person sorgfältige und dynamische Leseprozesse beim Kohärenzaufbau, sodass dieser Kohärenzaufbau mittlerweile daran bemessen wird, ob er sich in Hinblick auf die angestrebten Leseziele lohnt. Darum nehmen neuere Modelle die Wahrnehmung der Aufgabenerfordernisse (durchaus im Sinne der Selbstregulation beim Lernen aus Texten, s. Teilkap. [6.1.1.3](#)) auf und stellen Aufgabenmodelle sowie Relevanzentscheidungen an den Anfang von deskriptiven Prozessmodellen (Rouet & Britt, 2011; Rouet et al., 2017). Dadurch entwickeln zuvorderst kompetente Leserinnen und Leser eine hinreichende Zielvorstellung des Verstehensprodukts, um die dafür nötigen lese- und lernstrategischen Prozesse planerisch auszuwählen und gleichwohl adaptiv zu handhaben (Bromme, Pieschl & Stahl, 2010; Winne & Marzouk, 2019). Das ist deshalb so betonenswert, um das Dokumentenmodell nicht als alleinig sinnvolle Zielvorstellung zu postulieren, sondern als lediglich *eine* kontextabhängige Variante zu begreifen, die keineswegs voraussetzungsarm wäre, insbesondere bei den Inferenzen, welche intra- und intertextuelle Kohärenz herzustellen helfen.

Das Verstehen von multiplen Dokumenten im Sinne einer kognitiven Repräsentation von Inhalten und Dokumenten auf trennbaren und gleichzeitig verbundenen Ebenen ist ein komplexes Konstrukt (Nagele et al., 2021; Schoor et al., 2020). Ferner ist es ein Konstrukt, das verschiedene Varianten aufweist, die sich taxonomisch klassifizieren bzw. typologisch beschreiben lassen. Ursprünglich legten Britt et al. (1999) eine Vierer-Typologie vor, welche jüngst allerdings von Barzilai et al. (2023) erheblich erweitert wurde, indem das Forschungsteam eine teils theoriebasierte und teils empirisch unterfütterte Typologie vorgelegt hat, die neun Ausprägungen umfasst (s. Tabelle 17).

Tabelle 17: Produktbasierte Typologie der Repräsentation von Inhalten und Dokumenten beim Lesen multipler Dokumente

Dimension Repräsentation der Dokumente	Dimension Repräsentation des Inhalts			
	Ausprägung	Kein Inhalt	Separater Inhalt	Integrierter Inhalt
	Keine Dokumente	–	–	Mischmodell: * Repräsentation des Verhältnisses der Inhalte aus Dokumenten ohne jegliche Informationen zur Herkunft der Inhalte
	Separate Dokumente	–	Separate Repräsentation: * Repräsentation der Inhalte jedes einzelnen Dokuments ohne jegliche intertextuelle Verknüpfung – weder bei den Dokumenten noch bei den Inhalten	Integrierter Inhalt mit Dokumentetiketten: vernetzte intertextuelle Repräsentation der Inhalte, Inhalte mit Informationen zur Herkunft verbunden, aber keine vernetzte Repräsentation der Dokumente untereinander
	Integrierte Dokumente	Quellen-Modell: Repräsentation des Verhältnisses von Dokumenten, aber ohne jeglichen Inhalt	Integrierte Dokumente mit Inhaltsetiketten: vernetzte intertextuelle Repräsentation der Dokumente, Dokumente mit inhaltlichen Informationen verbunden, aber keine vernetzte Repräsentation der Inhalte untereinander	Dokumentenmodell: * integrierte Inhalte (integriertes mentales Modell) mit Verbindungen von untereinander integrierten Dokumenten (Intertextmodell)

(Legende: * = Typus bereits in der Typologie von Britt et al., 1999, enthalten; Quelle der eigenen Darstellung: Barzilai et al., 2023, S. 825)

Barzilai et al. (2023) haben für die Typologie in Tabelle 17 zwei Dimensionen zugrunde gelegt. Die erste Dimension betrifft die Separation bzw. Integration von Dokumenten, welche sie empirisch als Bestandteil von grafischen Visualisierungen von multiplen Dokumenten erfasst haben. Die zweite Dimension erstreckt sich auf die Separation bzw. Integration von Inhalten. Diese beiden Dimensionen unterteilten sie in jeweils drei Ausprägungen, welche miteinander kombinierbar sind: keine Repräsentationen, separate Repräsentation und integrierte Repräsentation. In sechs Typen fanden sie solche Kombinationen tatsächlich empirisch vor, wobei auffällig ist, dass jeweils nur separierte Dokumente bzw. Inhalte ohne Kombination mit Inhalten bzw. Dokumenten empirisch nicht vorkamen. Anders gesagt: Die Typologie startet streng genommen erst dort, wo sowohl Inhalte als auch Dokumente mindestens separat repräsentiert worden sind. Dies erklärt zudem die Häufung der drei unbezeichneten Varianten im oberen linken Quadranten der Tabelle 17. Diese Typologie ist als Indikator des Integrierens wertbar, denn mit List, Du, Wang und Lee (2019) lässt sich sowohl theoriebasiert als auch auf der Basis einer empirischen Diskriminanzanalyse die separate Repräsentation

des Inhalts als Ausdruck geringer intertextueller Integration werten, während die separate Repräsentation von Dokumenten ihnen zufolge eine geringe strukturelle Repräsentation des Dokumentensets darstellt.

Was Barzilai et al. (2023) in ihrer Typologie, die sich auf Daten von Jugendlichen der Sekundarstufe stützt, nicht berücksichtigten, aber von Britt et al. (1999) beschrieben wurde, ist das „Modell mit allen Dokumentenetiketten“. Wie oben erwähnt wurde, ist diese Variante Expertinnen und Experten vorbehalten, sodass es angesichts der Testpersonen erwartbar gewesen ist, diese Variante empirisch nicht anzutreffen. Der Vollständigkeit halber ist allerdings erst eine Zehner-Typologie nach gegenwärtigem Kenntnisstand aus rein konzeptueller Warte komplett. Es ist wahrscheinlich, dass es je nach Komplexität des Dokumentensets und seiner inhaltlichen Bezüge untereinander mutmaßlich weitere Subtypen wie partiell integrierte Inhalte bzw. Dokumente gibt. Wenn man allerdings den Grad der Integration von Inhalten und von Dokumenten als separate, orthogonal kombinierbare Dimensionen begreift, wie es ursprünglich aus theoretischer Sicht angelegt worden ist (Britt et al., 1999), dann ergibt sich daraus eine breite Anwendbarkeit. Denn die Integrationsleistungen beim Lesen multipler Dokumente lassen sich dergestalt genauer beschreiben und verorten.

3) *Multiperspektivität angemessen und genau repräsentieren*

Wenn Leserinnen und Leser multiple Dokumente integrieren wollen, ist die Konstruktion des mentalen Modells des Inhalts schwerlich davon zu trennen, welche Perspektive die Autorinnen und Autoren auf den Gegenstand einnehmen, etwa auf Nutzen und Nebenwirkungen von Impfstoffen. *Perspektiven* werden verstanden als *Formen des Denkens und Wissens, welche die Interpretation und Beurteilung von Sachverhalten beeinflussen* (Barzilai & Weinstock, 2020). Solche Perspektiven haben verschiedene Bezugspunkte variierender Granularität, so kann sich eine Person aus persönlicher Sicht impfkritisch im Sinne einer Disposition erweisen (individuelle Perspektive), sie kann in einer Gruppe von gleichgesinnten Personen Mitglied sein, die Impfen ablehnt und dies als Teil der Gruppenidentität in sich trägt (soziale Perspektive), sie kann gleichwohl in einer medizinaffinen und aufgeklärten größeren Gemeinschaft leben (kulturelle Perspektive), die auf die Herstellung von medizinischem Wissen (disziplinäre Perspektive) mit geeigneten, auch ethisch begründeten Verfahren wie mehrstufigen Versuchsreihen (methodische Perspektive) vertraut. In diesem kleinen Beispiel deutet sich an, dass Für und Wider von Impfstoffen als rein inhaltliche Thematik je nach Perspektive unterschiedlich wahrgenommen, in schriftlichen oder bildlichen Texten unterschiedlich und nicht einmal notwendigerweise völlig explizit gerahmt und dargestellt wird sowie von der lesenden Person ihrerseits aus ihrer Perspektive unterschiedlich verstanden wird. Aus diesem Grund plädieren Barzilai und Weinstock (2020) dafür, dass angemessenes Verstehen sowohl eine

Repräsentation von Inhalten und Aussagen beinhaltet, wie es lesetheoriekompatibel bislang beschrieben wurde, als auch eine Repräsentation der Perspektiven auf den Inhalt inkludieren sollte. Dabei sind multiple Dokumente vermutlich besser geeignet, die „*Quellen-Perspektiven*“, wie Barzilai und Weinstock (2020, S. 125) die Perspektiven von Autorinnen und Autoren bzw. Organisationen als Urheber von Aussagen bezeichnen, durch intertextuelle Konflikte salienter werden zu lassen, weil Widersprüche zwischen Aussagen bestehen können, die sich aus divergierenden Perspektiven wie der positiven oder negativen Einschätzung von Impfnebenwirkungen erklären lassen. Damit ist auch deutlich, dass insbesondere implizite Quellen-Perspektiven, die bewusst oder unbewusst kaschiert wurden, durch elaborative Inferenzen von der lesenden Person zu rekonstruieren sind. Dafür muss die lesende Person einschätzen, ob die Aussage von einer medizinisch ausgebildeten Person kommt und ob sie Absichten hegen könnte, die positiv oder negativ sind. Daraus ergibt sich, dass eine Repräsentation, welche die Perspektiven systematisch berücksichtigt, über die bisherigen theoretischen Annahmen hinausgeht, vor allem über jene des Dokumentenmodells bzw. weiterer alleinig auf separate Dimensionen der Dokumenten- und inhaltlichen Repräsentationen abgestützter Repräsentationsformen.

Barzilai und Weinstock (2020) greifen darum auf das Konstrukt „*Meta-repräsentation*“ von Perner (1991) zurück, welches postuliert, dass Personen eine Repräsentation einer Repräsentation entwickeln, um verschiedenartige Perspektiven verschiedener Personen einnehmen zu können. Diese Annahme erweitert das Dokumentenmodell in seiner Komplexität erheblich, weil es nun nicht mehr nur darum geht, internal eine vernetzte Repräsentation zu Inhalten und deren Ursprung aufzubauen, sondern auch die Perspektiven und deren Einfluss zu rekonstruieren. Das wird deutlich an der Darstellung in Abbildung 14, welche sich der Einfachheit halber auf einen alleinigen Text (in diesem Fall: einen journalistischen Artikel) bezieht, wobei dieses Beispiel recht leicht um eingebettete Quellen (s. Teilkap. [3.1.2.2](#)), z. B. Zitatgeber oder paraphrasierte Inhalte anderer Personen und Institutionen, erweiterbar ist oder sich auf mehrere Dokumente applizieren ließe.

Das Beispiel aus Abbildung 14 trennt drei Repräsentationen voneinander, die miteinander verbunden sind, und diese Verbindungen sind Teil der Metarepräsentation. Im Exempel sind dies das eigentliche Phänomen (a) der Inhalt eines Gesetzes und dessen etwaige Konsequenzen), b) die Repräsentation des Inhalts des Textes über das Gesetz und d) die Repräsentation der Quelle, also jener Person, die den Text b) verfasst hat, sprich: der Journalistin, welche einen Artikel geschrieben hat. Diese Elemente sind miteinander verbunden, etwa dergestalt, dass der Inhalt des Dokuments sich auf das Phänomen bezieht und beispielsweise selektiv negative Aspekt hervorhebt (c), dass dies der politisch-weltanschaulichen Perspektive der Journalistin geschuldet ist (f), was mutmaßlich darauf zurückzuführen ist, wie sie das Gesetz und dessen Inhalt wahrgenommen hat (e).

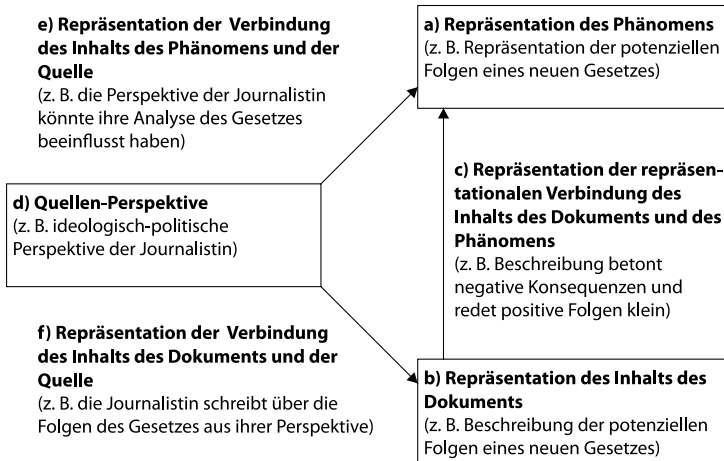


Abbildung 14: Metarepräsentation von Verbindung von Metadaten und Inhalt
(Quelle: Übersetzung von Barzilai & Weinstock, 2020, S. 126)

Es ist damit unschwer zu erkennen, dass eine solche Repräsentation nur eines einzelnen Dokuments erheblich mehr interpretative Leistungen der lesenden Person und diverse Kausalinferenzen zum Zustandekommen von Darstellungen in der inhaltlichen Repräsentation des Dokumenteninhalts (s. b) in der Abbildung 14) erfordert, als das Construction-Integration-Modell und das ihm entstammende Dokumentenmodell ursprünglich postulierten. Das Dokumentenmodell in der Darstellung aus der Abbildung 14 ist in den Komponenten (b) und partiell in (d) vorfindbar. Die Erweiterung dessen ursprünglicher Grundidee, Dokumenten- und inhaltliche Repräsentation abzubilden, um die Repräsentation von Multiperspektivität und den sich daraus ergebenden systematischen Implikationen für die Lesarten von Dokumenten helfen dabei, einige der ungelösten Probleme des Leseverstehens multipler Dokumente anzugehen. Solche Probleme manifestieren sich aktuell zum Beispiel darin, wie Personen aus widersprüchlichen, lückenhaften, inkohärenten und teils faktisch unzutreffenden externalen Repräsentationen in ihrem Gesamt dennoch zu kohärenten internalen Repräsentationen gelangen (Alexander & Disciplined Reading and Learning Research Laboratory, 2020; Cartiff & Greene, 2020; Kuhn, 2020; Saux, Britt et al., 2021).

4) Herausforderung Inkohärenz beim intertextuellen Integrieren

Die Frage intertextueller Kohärenzbildung ist zentral für das angemessene, tiefe Verstehen multipler Dokumente. Dabei ist das Verstehen multipler Dokumente keine additive Aneinanderreihung von mentalen Modellen einzelner gelesener Dokumente und ihren Dokumentenknoten. Entscheidend ist vielmehr, dass aufeinander beziehbare Informationen integriert werden, und zwar in einem

das Leseziel adressierenden Aufwand. In der bisherigen Forschung sind diverse Modelle vorgelegt worden, die zum einen mehr oder minder prominent die Integrationsprozesse und zum anderen den Verbund von Integrationsprozessen mit anderen Prozessen fokussieren, welche dabei helfen sollen, Kohärenz (wieder) herzustellen (Philipp, 2018b). In diese letztgenannte Gruppe fallen Modelle wie das „Discrepancy-Induced Source Comprehension Model“ (Braasch & Bräten, 2017; Braasch & Kessler, 2021), das „Two Step Model of Validation“ (Isberner & Richter, 2014; Richter & Maier, 2017), das „Content-Source Integration Model“ (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014), „Reading as Problem Solving“ (Britt et al., 2018) und „Knowledge Revision Components Framework – Multiple Documents“ (Butterfuss & Kendeou, 2021). Sie alle eint, dass sie sich insbesondere der Frage nach der Herstellung von Kohärenz bei widersprüchlichen Informationen widmen. In diesen Modellen wird die Frage behandelt, wie Personen damit verfahren, dass auf der einen Seite interne Informationen aus dem Vorwissen bzw. Überzeugungen oder bereits entwickelten Repräsentationen und externe Informationen auf der anderen Seite untereinander zu einem Widerspruch führen, der einen Kohärenzaufbau mindestens erschwert. Das „Discrepancy-Induced Source Comprehension Model“ geht davon aus, dass Personen in diesem Fall das Sourcing dafür nutzen, Widersprüche über Metadaten zu attribuieren und zu erklären (Braasch & Bräten, 2017; Braasch & Kessler, 2021). Das „Two Step Model of Validation“ sieht diese Möglichkeit optional vor, indem es eine zweischrittige Sequenz entwirft, in deren zweitem Schritt ein Widerspruch durch diverse lesestrategische Optionen aufgelöst werden kann (Isberner & Richter, 2014; Richter & Maier, 2017). Im „Content-Source Integration Model“ wird ebenfalls ein mehrschrittiges Vorgehen postuliert, welches in Abhängigkeit von Auflösungsnotwendigkeit bzw. dem überhaupt erkannten Kohärenzproblem wahlweise die Nutzung eigener Wissensbestände oder die Rekurrenz auf Metadaten und deren Evaluation zu einer interpretatorisch abgesicherten Kohärenzwiederherstellung nach sich zieht (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014). Demgegenüber gibt es Modelle, in denen das Auftreten von Widersprüchen und Inkohärenzen dazu genutzt wird, entweder im Rahmen sogenannter Nicht-Routine-Entscheidungen Leseziele zu adjustieren und kohärenzstiftende Maßnahmen durch Verhaltensmodifikationen zu ergreifen (Britt et al., 2018), oder aber das eigene Wissen zu revidieren, indem widersprüchliche Informationen zunächst im Arbeitsgedächtnis koaktiviert werden, um danach die neue und korrekte Information zu verstärken und im Langzeitgedächtnis zu verankern (Butterfuss & Kendeou, 2021; Kendeou & O’Brien, 2014).

Neben die fünf nur knapp skizzierten Einzelmodelle treten synthetisierende Modelle bzw. Rahmenmodellierungen, darunter das „Integrated Framework of Multiple Text Use“ (Alexander & Disciplined Reading and Learning Research Laboratory, 2020; List & Alexander, 2019), die einen weiter gefassten Geltungsbereich beanspruchen. Im Falle des „Integrated Framework of Multiple Text Use“

ist eine Synthese verschiedener Modelle vorgenommen worden, um umfassendere Verwendungszusammenhänge abbildbar zu machen. Das Integrieren ist hierbei ein Prozess des Verstehens, der als lesestrategisches Vorgehen beschrieben wird und ein Prozessverbund ist (Alexander & Disciplined Reading and Learning Research Laboratory, 2020; List & Alexander, 2019), der in neueren Publikationen dezidiert in den Dienst des „Relational Reasoning“ (relationales Schlussfolgern) gestellt wird. *Relationales Schlussfolgern* ist das Erkennen von bedeutungsvollen Verbindungen (Mustern) zwischen verschiedenen Informationseinheiten, die von einer Person aufgrund Inferenzen in eine kohärente Repräsentation überführt werden, wofür analytische und synthetisierende Fähigkeiten und Prozesse benötigt werden (Alexander & Disciplined Reading and Learning Research Laboratory, 2020). Im Falle des Lesens multipler Dokumente sind damit begründbare Zusammenhänge zwischen verschiedenen Informationseinheiten gemeint – seien es textuell oder grafisch kodierte Informationen, Metadaten und die Kombination von Inhalten mit Metadaten sowohl innerhalb einzelner Dokumente als auch bezogen auf ein ganzes Set an Dokumenten. Mit dem weiter gefassten Blick auf das Integrieren im Sinne des relationalen Schlussfolgerns ist auch einsichtig, dass Kohärenzstiftung bedeuten kann, das Verhältnis von widersprüchlichen Informationen zu erkennen und anzuerkennen, ohne dass dies notwendigerweise zwangsläufig aufgelöst werden muss, weil zum Beispiel Vorwissensbestände fehlen.

Das „Integrated Framework of Multiple Text Use“ sieht für die Lösung von Kohärenzproblemen explizit eine Sequenz von Schritten mit steigender kognitiver Anforderung für die lesende Person vor (List & Alexander, 2019). Die Reihenfolge erfordert damit immer mehr und vor allem immer abstraktere Inferenzen:

- 1) Der erste Schritt – „*Identifikation*“ – ist eine erste Anbahnung, da der lesenden Person aufgrund thematischer Überlappungen von Texten/Dokumenten deutlich wird, dass es potenzielle inhaltliche Bezüge zwischen ihnen gibt. In einem fiktiven Beispiel mit Impfstoffen wäre dies beispielsweise die aufkommende Erkenntnis, dass zwei Websites zum Thema etwas mit den Empfehlungen zum Nutzen eines Impfstoffs zu tun haben.
- 2) Im zweiten Schritt – „*separate Repräsentation*“ – werden solche potenziellen Bezugspunkte getrennt voneinander im Arbeitsgedächtnis koaktiviert. Das entspräche dem unverbundenen Verständnis des einen Dokuments, welches von einem Impfstoff abrät, und dem eines zweiten Dokuments, welches denselben Impfstoff dezidiert empfiehlt. Es kommt aber noch nicht zu Inferenzbildungen der Verknüpfung dieser Bestandteile.
- 3) Dies erfolgt erst im dritten Schritt – der „*simultanen Verbindung*“ –, denn hier hat die lesende Person die Bestandteile verknüpft. Jemand könnte also zur Aussage gelangen: „Während die eine Website den Impfstoff empfiehlt, tut die andere das nicht.“

- 4) Im vierten Schritt, der „*Bezeichnung der Verbindung*“, bildet die Leserin bzw. der Leser eine inferenzbasierte, abstrahierte Aussage über das Verhältnis der Aussagen bzw. der Dokumente: „Die beiden Websites widersprechen sich in ihren Einschätzungen zu den Nebenwirkungen des Impfstoffs.“ Dies entspricht den Leistungen beim Konstruieren eines Intertext-Prädikats.

Somit lässt sich diese Sequenz als ein Weg hin zur Klärung und vor allem Benennung des Verhältnisses von Aussagen aus Dokumenten begreifen, wobei hier nicht nur Widersprüche wie in anderen Modellen eine Rolle spielen, sondern prinzipiell jedwedes intertextuelle Verhältnis. Darüber hinaus beinhaltet das Modell zwei weitere Schritte, die als „Nachdenken über die intertextuellen Beziehungen“ bezeichnet werden:

- 5) Hierunter fällt das *Synthetisieren*, verstanden als das Bilden eines umfassenden integrierten mentalen Modells durch Kombination der Informationen in eine maximal kohärente Gesamtrepräsentation, darunter auch mittels komplexer Inferenzen. Ein Beispiel hierfür wäre: „Das Impfen ist ein umstrittenes, uneindeutiges Thema.“
- 6) Der letzte Schritt, der sich wiederum als Spezialfall von widersprüchlichen Aussagen begreifen lässt, wird als „*abgleichendes Auflösen der intertextuellen Inkonsistenzen*“ bezeichnet. Hierbei handelt es sich explizit um Anleihen aus dem „Content-Source Integration Model“ (Stadtler & Bromme, 2014), da dessen drei Vorgehensweisen hierfür angeführt werden, die vom Ignorieren von Inkohärenz, dem Erklären der Inkohärenz über eigenes Vorwissen über die vorwissensbasierte Einschätzung der Korrektheit von Positionen bis hin zum Zurückführen der Diskrepanz auf Metadaten der Dokumente reichen, bei denen man dem einen Dokument zur Impfung mehr Glaubwürdigkeit attestiert als dem anderen.

Zusammengefasst: Integrationsprozesse beziehen sich auf verschiedene Ebenen der Produkte, seien es einzelne Informationen, seien es Beziehungen zwischen Dokumenten, seien es die Vollständigkeits- und Vernetzungsgrade von mentalen Modellbildungen (List & Alexander, 2019; List, Du, Wang & Lee, 2019; Primor & Katzir, 2018; Segev-Miller, 2007). Gemeinhin gilt: Je stärker vernetzt und vollständiger die Integrationsprozesse und deren Produkte sind, als desto umfassender und tiefer gilt das Verstehen des in den Dokumenten geschilderten Sachverhalts. Hierin besteht eine Parallele zum Lesen singulärer Texte (s. Teilkap. 4.1.2.1). Allerdings machen multiple Dokumente es wahrscheinlicher, dass Personen auf wenig konsistente bis sogar widersprüchliche Informationen und Aussagen stoßen, die es ihnen abverlangen, auf der Basis von Evaluations- und

Inferenzprozessen Kohärenz zu stiften (Magliano et al., 2017; Rouet & Britt, 2022; Saux, Britt et al., 2021). Darum wird dies in der aktuellen Theoriebildung verstärkt adressiert, sodass das präzise Erkennen von informatorischen Bezügen nach gegenwärtigem Kenntnisstand und abgestützt auf diverse einzelne Theorien als zentraler Prozess bzw. – korrekter formuliert – als Prozessverbund gilt.

4.2 Prozesse und Produkte

Integrationsprozesse zum Zweck der kohärenten Repräsentation werden in diesem Teilkapitel in zwei Teilkapiteln entfaltet. Dabei wird sich aus Gründen der Komplexität des Gegenstands und Umfangsgründen die Darstellung rein auf das Integrieren von multiplen Dokumenten erstrecken, also nicht auf multimedial kodierte Inhalte. Das erste Teilkapitel (4.2.1) befasst sich mit dem Integrieren von verschiedenen Informationen im Allgemeinen, das zweite (4.2.2) geht auf die Integration vor allem widersprüchlicher Aussagen aus mehreren Dokumenten im Besonderen ein.

4.2.1 Intertextuelle Integration und Aufbau von Dokumentenmodellen

Intertextuelle Integrationsleistungen sind ein Forschungsgebiet, welches von einer großen Heterogenität gekennzeichnet ist (Primor & Katzir, 2018). In diesem Teilkapitel stehen Variablen bzw. Gruppen von Variablen im Vordergrund, die sich an das Dokumentenmodell anschließen lassen (s. Teilkap. 4.1.2.3). Mehrheitlich stammen diese aus Studien, in denen rezeptive Fähigkeiten gemessen wurden. Testverfahren, welche auf längeren eigenen Texten basieren, die Personen über gelesene Dokumente verfasst haben, werden an anderer Stelle (Teilkap. 5.2.1.2) behandelt. Weil das Dokumentenmodell zudem nicht nur auf Integrations-, sondern auch auf Sourcingprozessen (s. Teilkap. 3.1.2.3) basiert, wird zwecks Redundanzvermeidung auf eine erneute Darstellung der sourcingbezogenen Inhalte verzichtet. Der Vollständigkeit halber enthält die Tabelle 18 die Variablen dennoch; die Verweise in der mittleren Spalte geben an, in welchen anderen Teilen des Buchs sie vertieft behandelt werden.

Tabelle 18: Überblick über Bestandteile des Dokumentenmodells in seiner Operationalisierung

Bestandteil des Dokumentenmodells	Variable und Operationalisierung (Teilkapitel)	Beschreibung
Integriertes mentales Modell	Intertextuelle Inferenzen (als Verifikationsaufgabe) (s. „Intertextuelle Inferenzen“, 4.2.1.1) – Summenwerte	Vorgegebene Sätze mit Verknüpfungen aus zwei oder mehr Texten als gültig oder ungültig beurteilen
	Intertextuelle Konflikte (als Verifikationsaufgabe) (s. „Intertextuelle Konflikterkennung“, 4.2.1.1) – Summenwerte bzw. relative Maße	Vorgegebene Sätze mit Verknüpfungen zweier konfligierender Aussagen aus zwei oder mehr Texten als gültig oder ungültig beurteilen
	Intertextuelle Inferenzen in offenen Aufgaben (s. „lokale Kombinationen“, 5.2.1.2 bei Diskurssynthesen-Maßen) – Summenwerte	Zählung und teils Klassifikation von Verknüpfungen von Informationen und Informationseinheiten in Texten ohne vorab definierte Korrektheit von Lösungen
	Ausgewogenheit von inhaltlichen Positionen bei widersprüchlichen Dokumenten (s. „Textstrukturmaße“, 5.2.1.2 bei Diskurssynthesen-Maßen) – Ordinalskalen	Inhaltlich mehrperspektivische und unverzerrte Berücksichtigung mehrerer Perspektiven aus verschiedenen Dokumenten
	Intertextuelle Inferenzen zu Texten in Sets mit multiplen Dokumenten, die überzeugungskonsistent bzw. -inkonsistent sind (4.3.2 , 4.3.3.1) – Summenwerte	Indikator für unverzerrte bzw. verzerrte Repräsentation von konfligierenden Aussagen zu Themen, in denen die Leserinnen und Leser ontische Überzeugungen haben, welche eine balancierte Gesamtrepräsentation erschweren
Intertextmodell	Quelle-Inhalt-Verbindungen (s. Teilkap. 4.2.1.2) – Summenwerte	Verknüpfung der Quelle von Aussagen (als Dokumentenquelle oder eingebettete Quelle) mit der Aussage (Proposition)
	Referenzierungen (s. Teilkap. 5.2.1.2 für Diskurssynthesen-Maße) – Summenwerte	Anzahl von Nennungen von Metadaten (in Verbindung mit Inhalten)
Vollständige Dokumentenmodelle	Typologische Klassifikationen, meist ausgehend von Merkmalen des integrierten mentalen Modells (s. Teilkap. 4.4.1.1) – Ordinalskalen	Integrationsgrade von inhaltlichen Aussagen und Metadaten in einem eigenen grafischen oder Textprodukt

4.2.1.1 Integration von Informationen aus mehreren Dokumenten

In diesem Teilkapitel werden zwei wichtige Teilleistungen beim Verstehen multipler Dokumente behandelt, die sich auf die Kombination von Propositionen aus mindestens zwei Dokumenten stützen. Solche intertextuellen Inferenzen werden als generische Fähigkeit in einem ersten Schritt betrachtet. In einem zweiten Schritt erfolgt eine Vertiefung, indem ein Ausschnitt fokussiert wird, nämlich die Fähigkeit, intertextuelle Widersprüche als solche zu erkennen.

1) Intertextuelle Inferenzen

Produkte

Intertextuelle Inferenzen als Teil der Integrationsleistungen sind ein stark untersuchter Gegenstand der Multiple-Documents-Literacy-Forschung (Primor & Katzir, 2018). In diesem Abschnitt werden zwei Arten der Inferenzleistungen betrachtet. Erstens handelt es sich um geschlossene Aufgaben in rezeptiven Tests, bei denen vorgegebene Inferenzen als korrekt oder inkorrekt erkannt werden müssen. Zweitens werden Inferenzen betrachtet, die in expressiven Aufgabenformaten von den Testpersonen ex post ausgewertet werden.

a) Rezeptives Beurteilen der Zulässigkeit von intertextuellen Inferenzen bei rezeptiven Integrationsleistungen

Die Beurteilung der Korrektheit von Inferenzen ist ein Verfahren, das als „Informationsverifikationstechnik“ bezeichnet wird (Royer et al., 1996). Dabei werden Probandinnen und Probanden Aussagen vorgelegt, die sie aufgrund der zuvor gelesenen Textinhalte aus einem oder mehreren Texten als korrekt oder inkorrekt einschätzen sollen. Aus den einzelnen Inferenz-Antworten (Items) werden Summenwerte und Skalen gebildet, welche die Verstehensleistungen abbilden. Das Instrument ist deshalb von Interesse, weil es in der Forschung dazu eingesetzt wird, verschiedene Arten von Inferenzen zu messen:

- 1) *Inferenzen auf der Basis von textuellen Informationen allein oder aufgrund von textuellen Informationen und dem eigenen Vorwissen:* Royer et al. (1996) unterscheiden „nahe Inferenzen“ von „fernen Inferenzen“. Damit ist im ersten Fall gemeint, dass Personen zwei Informationen eines Texts kombinieren, um die Gültigkeit der Items einzuschätzen. Im zweiten Fall müssen Personen die Textinhalte mit ihrem Vorwissen verbinden, um die Items richtig zu lösen. Diese Grundlogik ist dafür verwendet worden, die unterschiedlichen intra- und intertextuellen Inferenzen als Grundlage für das Textverstehen zeitökonomisch zu erfassen.
- 2) *Intra- und intertextuelle Inferenzen:* Studien aus der Multiple-Documents-Literacy differenzieren etwas andere Leistungen, indem sie intra- und intertextuelle Inferenzen erfassen, die sie entweder als separate Verstehensprodukte oder als eindimensionales Konstrukt (z. B. Davis et al., 2017; Follmer & Tise, 2022; Incognito & Tarchi, 2024; Mateos et al., 2016) ausweisen. Der Unterschied liegt darin, dass intratextuelle Verifikationsaufgaben Verstehensleistungen auf der Ebene eines einzelnen Textes messen, was eher in den Bereich der separaten mentalen Modellbildung fällt, soweit sich dies mit singulären Inferenzen operationalisieren lässt. Intertextuelle Verifikationsaufgaben fokussieren demge-

genüber das Verstehen von mindestens zwei Texten, denn bei ihnen muss die gültige Inferenz von Aussagen aus zwei Texten als solche erkannt werden. Beide Leistungen korrelieren, wenn sie als separate Skalen behandelt werden, relativ stark miteinander (Philipp, 2020a). Diese Zusammenhänge erklären sich unter anderem darüber, dass aus theoretischer Sicht anzunehmen ist, dass integrierte mentale Modelle mit multiplen Texten als Grundlage darauf aufbauen, dass das Verstehen singulärer Texte eine Bedingung für das Verstehen multipler Texte ist (Britt et al., 1999; Philipp, 2018a, 2020a).

Die Verwendung von Batterien der intertextuellen Verifikationsaufgaben ist bis heute ein in der Forschung vorkommender Weg, Verstehensleistungen zu operationalisieren. Diverse Studien bedienen sich dieses Zugangs (Blaum et al., 2017; Braasch, Bråten, Strømsø & Anmarkrud, 2014; Bråten & Strømsø, 2006b, 2010a, 2010b, 2011; Bråten et al., 2008, 2009; Cheong et al., 2019; Follmer & Tise, 2022; Gil et al., 2010a, 2010b; Griffin et al., 2012, in Druck; Hagen et al., 2014; Karimi, 2015, 2018; Karimi & Atai, 2014; Karimi & Shabani, 2013; Lee & List, 2023a; Lehmann et al., 2019; List, 2021; List, 2023b, Studie 3; List & Lin, 2023; Salmerón et al., 2010; Strømsø & Bråten, 2009; Strømsø et al., 2008, 2010; Wei et al., 2024).

- 3) *Verknüpfung expliziter bzw. impliziter Informationen*: Schließlich lässt sich noch differenzieren, ob die Inferenzleistungen, sei es bei einem nahen oder fernen Inferieren, sei es beim intra- oder intertextuellen Inferieren, aufgrund explizit gegebener Textinhalte oder aufgrund implizit zu erschließender Inhalte zu leisten ist (Yukhymenko-Lescroart et al., 2022). Diese Differenzierung ist hilfreich, weil sich anhand der empirischen Daten nachweisen ließ, dass implizite intertextuelle Inferenzen einen besonders hohen Schwierigkeitsgrad aufweisen.

Inferenzen in der Multiple-Documents-Literacy-Forschung stellen einen gut erforschten Gegenstand dar. Dabei überwiegt in Verifikationsaufgaben ein Zugang, Kolonnen von Sätzen vorzulegen, die potenzielle textübergreifende inhaltliche Verbindungen repräsentieren. Je mehr valide Inferenzen als korrekt erkannt und nicht-valide Aussagen als inkorrekt detektiert werden, desto höher und vollständiger wird das Leseverstehen interpretiert. Dieses Vorgehen bildet jedoch nur einen Ausschnitt, da die zu bildenden Inferenzen vorab ausgewählt werden. Es gibt weitere Zugänge, in denen die Inferenzleistungen ex post rekonstruiert werden.

b) *Satzweise bzw. Subeinheiten fokussierende Auswertungen bei expressiven Integrationsleistungen*

In Studien, die mit offenen Aufgabenprodukten arbeiten, werden Inferenzleistungen nachträglich kodiert. Dabei bestehen diverse Freiheitsgrade in der Fokussierung auf die interessierenden Integrationsleistungen (Primor & Katzir, 2018). Dies mündet in diverse Operationalisierungen und Konstrukte, die hier nicht taxonomisch systematisiert, sondern in ihrem grundsätzlichen Zugang unterschieden werden. Der erste Zugang hilft dabei, im Gesamt besser abzuschätzen, dass und in welchem Ausmaß die intertextuellen Inferenzen bei den offenen Aufgaben aktiv im Produkt gebildet wurden. Der zweite Zugang, der eine Engführung auf einen Ausschnitt darstellt, klassifiziert die Antworten und verdeutlicht, welchen Explizitheitsgrad und welche Bezugspunkte die intertextuellen Inferenzen aufweisen.

Zugang 1: Intertextuelle Inferenzen als Bestandteil von satzweise stattfindenden oder anderen kleinteiligeren Analysen (sog. „Themeneinheiten“) der kognitiven Transformationsleistungen

In diesem Zugang zu Auswertungen werden die schriftlichen Produkte segmentiert und die Segmente danach kodiert, was die Quelle der Aussagen ist (Inhalte aus den Dokumenten, von der schreibenden Person sowie Hybridformen) und ob es sich um Inferenzen handelt (Greene, 1993; Rouet et al., 1996). Diese Inferenzen werden zudem noch danach untergliedert, ob sie intratextuell oder intertextuell sind. Mit diesem Vorgehen lassen sich schriftliche Antworten feinteilig danach analysieren, welche Reproduktions-, Elaborations- und Inferenzleistungen Personen erbracht haben. In den bisherigen Studien mit diesem Vorgehen werden teils die einzelnen Indikatoren separat ausgewertet (De La Paz & Wissinger, 2015; Greene, 1993; Kiili, Bråten et al., 2020; Kullberg et al., 2023; Le Bigot & Rouet, 2007; Linderholm et al., 2014, Studie 1; Naumann et al., 2009; Voss & Wiley, 1997; Wiley & Voss, 1996) oder aber mittels verschiedener Gewichtungen zu Gesamtscores (Gesamttransformationsindizes) aggregiert (De La Paz & Wissinger, 2015; Gil et al., 2010a, 2010b). Insbesondere der erste Zugang ist instruktiv, weil er als erstes wichtiges Ergebnis zutage bringt, dass die schriftlichen Antworten eine Kombination verschiedener Leistungen darstellen, die in individuellen Mischverhältnissen realisiert werden. Die Ergebnisse in den einzelnen Studien divergieren freilich: Es existieren Hinweise sowohl auf mehr übernommene als transformierte Inhalte (Kiili, Bråten et al., 2020; Kullberg et al., 2023; Linderholm et al., 2014, Studie 1; Wiley & Voss, 1996) als auch das Gegenteil davon (Naumann et al., 2009).

Ein zweites wichtiges Resultat offenbart sich mit Blick auf die Aufgabenstellungen. Denn offenbar beeinflusst diese, wie viele intertextuelle Inferenzen und Transformationen Testpersonen erbringen (s. a. „Illustrierendes Beispiel

7“ in Teilkap. [2.2.2.4](#)). Insbesondere das schriftliche Argumentieren führt (mit Ausnahme der Ergebnisse von Gil et al., 2010a, 2010b) zu mehr intertextuellen Inferenzen und damit zu einer höheren Integrationsleistung (De La Paz & Wisinger, 2015; Le Bigot & Rouet, 2007; Lehmann et al., 2019; Voss & Wiley, 1997; Wiley & Voss, 1999, Experimente 1 und 2).

Dieser erste Zugang mit Kodierungen, woher Inhalte stammen und welche Art von intertextueller Inferenzleistungen die Personen erbracht haben, hat seinen Wert darin, die Verschiedenartigkeit der Leistungen zu verdeutlichen. Allerdings ist gerade bei den Inferenzen unklar, worauf sie sich beziehen, welche Dokumente involviert sind und welche Wichtigkeit sie in den zu verarbeitenden Dokumenten bzw. welche Relevanz sie für das Aufgabenprodukt hatten. Außerdem divergieren in den Studien die Bezugseinheiten stark in ihrer Granularität und teils kommt es im Lauf der Zeit zu Veränderungen der kodierten Leistungen, sodass sich auf dieser heterogenen empirischen Grundlage kaum studienübergreifende Muster extrahieren lassen. Ein weiterer Kritikpunkt betrifft die Verwendung der Gesamtindizes, die als gezielte Aggregation kaum noch qualitative Unterschiede in den kognitiven Leistungen zu kommunizieren in der Lage sind. Andere methodische Zugänge greifen diese Schwächen gezielt auf.

Zugang 2: Gezielte mehrdimensionale Klassifikation der Integrationsleistungen

Die Taxierung der Inferenzen in einer mehrdimensionalen Zugangsweise ist in einer Reihe von Studien vorgelegt worden. Dieser Zugang fokussiert nur auf Inferenzen, die auf mehrere Arten kodiert wurden.

- 1) *Ex- und Implizitheit der Inferenzen*: Explizite Integrationen benennen deutlich, dass Informationen aus mindestens zwei Dokumenten stammen und nutzen dabei teils Konnektoren, welche die Relation von kombinierten Informationen indizieren. Implizite Integrationen sind bündelnde Aussagen, denen dann meist Aussagen aus einzelnen Dokumenten folgen, anhand derer rekonstruierbar ist, worauf die bündelnde Aussage basiert (Lee et al., in Druck; List, Du, Wang & Lee, 2019, Studie 1; Rickert & Philipp, 2022). Eine zusätzliche Kategorie in Erweiterung der Im- und Explizitheitsdichotomie beschreiben Hartmann et al. (2021): die Pseudo-Integration. Bei ihr werden Zusammenhänge zwischen Aussagen suggeriert, allerdings ist die Art und Korrektheit der Integration zweifelhaft. Die Häufigkeitsverteilung von im- und expliziten Inferenzen werteten List, Du, Wang und Lee (2019, Studie 1) aus. Sie kamen zu dem Ergebnis, dass sich die Integrationen, von denen es im Mittel etwas mehr als eine pro schriftliche Antwort gab, ungleichmäßig verteilten: Rund ein Viertel war explizit, drei

Viertel waren implizit. Bei Rickert und Philipp (2022) lag das Verhältnis bei Bestandteilen des integrierten mentalen Modells bei 1 zu 2 zugunsten der impliziten Bestandteile.

- 2) *Korrektheit der Inferenzen*: Inferenzen lassen sich danach auswerten, ob die gezogenen Schlussfolgerungen zulässig oder unzulässig sind. Diese Prämisse, auf der die Funktionsweise der Verifikationsaufgaben in rezeptiven Tests basiert, ist in expressiven Aufgabenformaten bislang kaum untersucht worden (Rickert & Philipp, 2022). In der Längsschnittstudie von Rickert und Philipp (2022) wurde dies anhand einer Musterlösung kodiert; das wichtigste Resultat lautete, dass explizite intertextuelle Inferenzen nicht nur selten vorkamen, sondern auch überwiegend inkorrekt waren. Ein weiteres Ergebnis war, dass explizites Vorhandensein und Korrektheit stark positiv zusammenhängen, während dies bei implizitem Vorhandensein nicht der Fall war. Demnach bestehen zwischen zwei Merkmalen der Inferenzen straffe positive Kopplungen: Wer mehr Inferenzen explizit berichtet, tut es inhaltlich angemessener.
- 3) *Spezifität der Integrationen*: Hierbei werden drei Formen unterschieden (List, 2019; List, Du, Wang & Lee, 2019, Studie 1): Erstens können sich Inferenzen auf *thematische Integrationen* beziehen, also auf dokumentenübergreifende Hauptideen oder zentrale Aussagen. Solche Inferenzen betreffen also inhaltliche Überlappungen bzw. Verknüpfungsmöglichkeiten, etwa zwischen zentralen Konzepten aus verschiedenen Dokumenten. Zweitens sind *vergleichende Inferenzen* anzuführen, welche die dokumentenübergreifende Nutzung von Informationen aus einzelnen Dokumenten betreffen, indem sie miteinander abgeglichen werden. Drittens sind *kontextuelle Inferenzen* kodierbar gewesen, in denen nicht die Inhalte, sondern die Metadaten miteinander kombiniert werden, um Perspektiven oder die Glaubwürdigkeit von Quellen zu vergleichen. Alle drei Formen der produktbezogenen Maße verwendeten List, Du und Lee (2021a) in einer offenen Aufgabe, um dadurch gezielt Verknüpfungen zwischen zwei Dokumenten zu erfragen. Zudem wurden diese Kategorien ebenfalls für die Analyse von Prozessdaten genutzt (List, Du & Lee, 2021b).

Die Ergebnisse der Studien haben Gemeinsamkeiten und Unterschiede. Eine Gemeinsamkeit ist, dass Inferenzen mit Metadaten am seltensten vorkamen. Gemessen an den Verteilungen pro Studie überschreiten sie mit einer Ausnahme (List, 2019) kaum die Zehn-Prozent-Marke (List, Du & Lee, 2021a, 2021b; List, Du, Wang & Lee, 2019, Studie 1). Dieser Befund spiegelt sich in den Prozessdaten wider (List, Du & Lee, 2021b). Den Großteil der intertextuellen Inferenzen, und das ist ein zweiter Befund, machen also thematische Integrationen und vergleichende Integrationen aus. Die Verteilungen in den Studien divergieren, sie liegen bei komplexen schriftlichen Aufgaben mit vielen Dokumenten ungefähr paritätisch (List, 2019; List, Du & Lee, 2021b; List, Du, Wang & Lee, 2019), bei wenigen Texten

kommen teils die vergleichenden Integrationen häufiger vor (List, Du & Lee, 2021b). Verlangt die Aufgabe hingegen ein grafisches Produkt, überwiegen die thematischen Integrationen deutlich (List, 2019).

Trotz der dünnen Datenlage lässt sich vorläufig festhalten: Die differenzierte Analyse der intertextuellen Inferenzen illustriert, dass implizite und sich auf die thematischen Bezüge fokussierende Inferenzleistungen überwiegen. Kaum in den Blick geraten sind zudem Zusammenhänge zwischen produktbezogenen Indikatoren von Integrationsleistungen und weiteren Maßen. Eine Ausnahme hier für bildet die Studie 1 von List, Du, Wang und Lee (2019). Sie korrelierten drei Konstrukte und kamen zum Resultat, dass je mehr integrative Aussagen – verstanden als Summe aller Integrationsformen – Personen in ihren schriftlichen Antworten aufwiesen, desto mehr Diskurskonnektoren zum einen und desto mehr Referenzierungen zum anderen in diesen Texten enthalten waren.

- 4) *Strukturebenen der intertextuellen Makrostruktur*: Ein aus der Forschung zu Makroregeln des Textverstehens und dem schriftlichen Zusammenfassen stammender Zugang besteht darin, die hierarchische Struktur von Propositionen zu fokussieren und dabei die übergeordneten Inhalte von Texten zu erkennen bzw. per Makroregeln zu bilden (van Dijk, 1980). Das Bilden von über- und untergeordneten Propositionen in integrierten mentalen Modellen („Superpropositionen“ gemäß Segev-Miller, 2007) auf der Basis von Inferenzen kann analog der Logik der Mehrebenen-Repräsentationen aus Modellen des Leseverstehens singulärer Texte (McNamara & Magliano, 2009) erfolgen. Empirisch untersucht wurde es bislang allerdings kaum, selbst wenn mittels Musterlösungen Abstraktionsleistungen wie das Benennen von Teilthemen auswertbar ist (Rickert & Philipp, 2022).
- 5) *Deklaration der intertextuellen Relationen*: Diese Kategorisierung nimmt auf, ob Divergenz, Konvergenz oder eine andere Form von intertextuellen Verhältnissen erkannt wurden. Dies wurde bisher selten erforscht – und zwar sowohl produktbezogen (Rouet et al., 1997) als auch mit Blick auf die Prozesse (List, Du & Lee, 2021b). Diese Fokussierung ermöglicht es, Intertextprädikate zu bilden, also ein theoriegeleitetes Konstrukt vor allem für das Intertextmodell zu fokussieren (Perfetti et al., 1999). Diese Variable ist nur wenig untersucht worden und zeichnet sich dadurch aus, dass sich in ihr nicht nur teils eine präzisere Einschätzung durch Personen mit höherer Expertise manifestiert (Rouet et al., 1997), sondern auch die Eigenheiten der Dokumentensets, indem beispielsweise mehr Dokumente in einem vorgegebenen Set zu mehr solcher Inferenzen führen, als wenn nur wenige Dokumente in kleineren Dokumentensets zu verarbeiten sind (List, Du & Lee, 2021b).

Inferenzen lassen sich in expressiven Aufgaben multipel klassifizieren und beschreiben. Sie können ge- oder misslingen (korrekt oder inkorrekt sein), deutlich markiert (explizit sein) oder implizit ausfallen. Sie können zudem verschiedene Bezugspunkte haben und diese verbinden sowie schließlich dazu dienen, das Verhältnis zwischen den informatorischen Einheiten zu deklarieren. Studien, die mehrere Merkmale von Inferenzen gemeinsam betrachten, sind jedoch selten.

c) Zusammenfassung

Die hier behandelten Inferenzleistungen aus den rezeptiven (geschlossenen) oder expressiven (offenen) Aufgaben stellen nur einen Ausschnitt aller Integrationsleistungen dar, gleichwohl zählen sie zu den häufigsten empirischen Zugängen, Integrationsleistungen zu operationalisieren (Barzilai et al., 2018; Primor & Katzir, 2018). Beide Zugänge erweisen sich komplementär und fokussieren gleichwohl sehr unterschiedliche Ausschnitte. So fällt auf, dass der Kardinalweg bei rezeptiven Tests, die Korrektheit der Inferenzen, bei den expressiven Tests kaum vorkommt. Dafür ist die Kleinteiligkeit der Kodierungen der expressiven Tests so bislang nicht im rezeptiven Pendant erkennbar. Dies spiegelt ein Muster wider, welches über diverse weitere hier nicht behandelte Variablen bei expressiven Tests besteht: Sie erlauben ein facettenreicheres, mehrdimensionales Bild der Verstehensleistungen (Primor & Katzir, 2018).

Prozesse

Intertextuelle Integrationsprozesse und – als bewusst gesteuerte Variante – Strategien des Integrierens haben in Theorie und Forschung zunehmend Aufmerksamkeit erhalten. Dies hat zu einer konzeptuellen und empirischen Heterogenität geführt, der dieser Abschnitt mit einer zweiteiligen Darstellungsstrategie begegnet. Zunächst werden taxonomische Zugänge gewählt, um die vielfältigen intertextuellen Prozesse (aus Lautdenk-Studien) zu ordnen. In einem zweiten Schritt stehen verschiedene Prozesse und Strategien im Zentrum, welche in einer dreiteiligen Sequenz entfaltet und mit Leistungsdaten in Verbindung gebracht werden.

Klassifikationen von intertextuellen Integrationsprozessen

Integrations- und Inferenzprozesse lassen sich analog zu den produktbezogenen Indikatoren verschiedenartig klassifizieren. Dadurch wird die in Studien berichtete Gesamtanzahl von intertextuellen Inferenzen als Variable präzisiert, was der empirisch beobachtbaren Heterogenität der Inferenzen (Afflerbach & Cho, 2009) und vor allem der Eminenz von inferenzbasierten Kohärenzleistungen

beim Lesen multipler (digitaler) Dokumente (Cho & Afflerbach, 2017) gerecht wird. In den Lautdenk-Studien werden hierzu verschiedene klassifikatorische Vorschläge genutzt:

- 1) *Spezifität der Integrationen*: Die Spezifität der Integrationen wurde in Studien teils auf ähnliche Weise kodiert wie bei den Produkten (s. Nr. 3 in der Liste im Abschnitt „b) Satzweises bzw. Subeinheiten fokussierende Auswertungen bei expressiven Integrationsleistungen“, dort: Zugang Nr. 2). Eine solche Studie stammt von List, Du und Lee (2021b) und ist im Kasten „Illustrierendes Beispiel 17“ Gegenstand. Die Ergebnisse dieser Studie stehen im Einklang mit denen anderer, produktbezogener Studien, welche vor allem thematische Integrationen als häufigste Form zum Ergebnis hatten.
- 2) *Art der intertextuellen Verhältnisse von Informationen*: Inferenzen geben nicht nur Auskunft darüber, welche Informationen miteinander kombiniert werden, es lässt sich auch das intertextuelle Verhältnis kodieren. Dies haben List, Du und Lee (2021b) fokussiert und ermittelt, welche Arten von intertextuellen Verhältnissen in den Inferenzen rekonstruierbar waren (s. Kasten „Illustrierendes Beispiel 17“). Sie unterschieden drei Arten: Komplementarität/Übereinstimmung, Divergenz und weitere Formen wie die Erklärung der einen Informationen über eine andere aus einem anderen Dokument. Diese Art von Kodierung von Inferenzen ist damit eine andere als die nachfolgende, welche sich auf die Bezugseinheit der Dokumente bezieht.
- 3) *Deklaration der intertextuellen Relationen*: Die Repräsentation, in welchem Verhältnis Aussagen bzw. Inhalte aus mehreren Dokumenten zueinander stehen, ist bereits im Produkt-Teilkapitel angesprochen worden (s. Nr. 5 in der Liste im Abschnitt „b) Satzweises bzw. Subeinheiten fokussierende Auswertungen bei expressiven Integrationsleistungen“, dort: Zugang Nr. 2). Eine präzise Repräsentation als Ergebnis der Analyse von Aussagen bedarf der prozessualen Verarbeitung der Inhalte mit einer inferenzbasierten relationalen Bezeichnung der Inferenz. Diese intertextuellen Relationen zu finden, beschreibt das Modell von List und Alexander (2019) als eine Abfolge von vier Schritten, bei der mit zunehmender Progression nicht nur die Inhalte kombiniert, sondern auch die Art der Beziehungen von Aussagen und ganzen Dokumenten geklärt wird (s. dazu Abschnitt „Herausforderung Inkohärenz beim intertextuellen Integrieren“ im Teilkap. [4.1.2.3](#)). Diese vier Schritte bzw. Arten des Inferierens sind: a) Identifikation (Erkennen potenzieller inhaltlicher Bezüge), b) separate Repräsentation als zwar koaktivierte, aber noch unverbundene Repräsentation der Inhalte aus mehreren Dokumenten, c) simultane Verbindung als thematische Verknüpfung und d) die Deklaration des intertextuellen Verhältnisses von Informationen. Diese theoretisch unterscheidbaren Leistungen ließen sich in der Studie von List, Du und Lee (2021b) nachweisen; dieser wichtigen Studie ist darum der Kasten „Illustrie-

rendes Beispiel 17“ gewidmet. Vorläufer dieser Art der Auswertung haben zudem List, Du und Wang (2019, Studie 2) vorgenommen. Sie differenzierten in Lautdenk-Protokollen, ob intertextuelle Integrationen nicht nur Rückschlüsse auf die gegenwärtige Koaktivierung von zwei Texten beinhalteten (Identifikation bis maximal separate Repräsentation in der zuvor genannten Taxonomie), sondern auch ob sie eine klare intertextuelle Inferenz beinhalteten. Das war in ungefähr zwei Fünfteln der Äußerungen der Fall.

- 4) *Richtung der intertextuellen Inferenzen*: Die intertextuellen Inferenzen werden bislang nur selten danach untersucht, welcher Text der Ausgangspunkt und welcher Text der Bezugspunkt ist. Es geht also nicht nur darum, dass eine Verknüpfung zwischen zwei Texten stattfindet, sondern auch auf welcher Grundlage, nämlich hinsichtlich der Reihenfolge. Dies kodierten McCrudden et al. (2021) in einer Grundlagenstudie und replizierten das gefundene Muster später in einer experimentellen Studie (McCrudden et al., 2022). Dazu wurde in einem Text ein allgemeines Prinzip der natürlichen Auslese beschrieben und im anderen Text anhand konkreter Beispiele das Prinzip illustriert, ohne dass dies explizit benannt wurde. Die Leserinnen und Leser mussten also das Beispiel mit dem Prinzip aus den beiden komplementären Texten kombinieren. Dies gelang ihnen besser, wenn der Prinzipien-Text als erster Text gelesen wurde (McCrudden et al., 2022), und diese Inferenzen vom generischen Thema zum konkreten Beispieltext führten zu einem besseren Verstehen der Thematik in offenen Tests (McCrudden et al., 2021, 2022). Dieses sich erst allmählich abzeichnende Muster macht darauf aufmerksam, dass es vertiefender Analysen wert ist, die Dynamik der Integrationsprozesse ausgehend von der Dokumentensequenz in ihrer Direktionalität und den sich daraus ergebenden Effekten genauer zu analysieren.

Illustrierendes Beispiel 17

Intertextuelle Integrationsprozesse – mehrdimensional und aufgabenspezifisch betrachtet

A) Was war das Erkenntnisinteresse?

Welche Art von Integrationsleistungen nehmen Personen vor, die zwei unterschiedlich komplexe Dokumentensets lesen und dabei explizit aufgefordert werden, im Prozess des Lesens Verbindungen zu benennen? Und welche Leistungen erbringen sie in geschlossenen bzw. offenen Aufgaben, die das intertextuelle Integrieren erfordern? Diesen beiden Fragen verschrieb sich eine Studie mit Studentinnen und Studenten (List, Du & Lee, 2021b).

B) Wer wurde untersucht?

Das Sample umfasste 33 Studierende einer Universität aus dem amerikanischen Nordwesten, die im Mittel 19 Jahre alt waren. Drei Viertel der Studierenden waren Frauen, ebenfalls drei Viertel waren europäische Amerikanerinnen und Amerikaner. Sie wurden aus entwicklungspsychologischen Kursen rekrutiert. Die Studierenden bearbeiteten zwei Aufgaben, sodass hier ein Intrapersonenvergleich vorgenommen wurde.

C) Welche Instrumente kamen zum Einsatz?

Die Studierenden hatten in einem Within-Subjects-Design zwei Aufgaben zu lösen. In der ersten, einfacheren Aufgabe lasen sie drei ausgedruckte Dokumente zum Thema Stonehenge, und diese Dokumente enthielten konfligierende Aussagen, es galt also Widersprüche zu erkennen. Die zweite, komplexere Aufgabe, basierte auf einem Dokumentenset von acht ausgedruckten Dokumenten, welche verschiedene Positionen auf die Legalisierung von Sexarbeit einnahmen. Die eigentlichen Texte waren zwischen 240 und 400 Wörtern lang.

Die Aufgaben waren ähnlich: Die Personen sollten die Dokumente lesen, um eine Einschätzung zum kontrovers dargestellten Sachverhalt zu erlangen, vor allem aber sollten sie während des Lesens immer dann eine Markierung anfertigen, wenn ihnen intertextuelle Bezüge zwischen den Texten auffielen, und die Art der Verbindung mit lautem Denken erklären. Diese prozessbezogenen Aussagen zum Integrieren wurden auf viererlei Art pro Aussage kodiert. Erstens kodierten die Forscherinnen, *auf welche Dokumente* sich die Personen bezogen, was später ausgezählt wurde. Zweitens wurden die Bezugspunkte bzw. die *Spezifität von Aussagen* kodiert: a) vergleichende Verknüpfungen zwischen mehreren Texten, b) thematisch übergeordnete Hauptideen sowie c) kontextuelle Aussagen, etwa zur Glaubwürdigkeit einzelner Autorinnen und Autoren. Drittens wurde die *Art der intertextuellen Verhältnisse von Informationen* kodiert, z. B. Widersprüche, Komplementarität oder Erklärungen der Inhalte aus dem einen Dokument aufgrund anderer Dokumente bzw. deren Inhalte. Viertens wurde die Integrationsleistung nach *Klärung des Verhältnisses der jeweiligen betroffenen Dokumente* kodiert: a) relationale Beziehung als Erkennen potenzieller thematischer Bezüge (Identifikation), b) separate Repräsentation als Erkennen von einzelnen Themen über die Dokumentengrenzen hinweg, c) simultane Verbindung, also das Erkennen von thematischen Bezügen zwischen den Dokumenten bei Benennung des intertextuellen Verhältnisses, und d) Bezeichnung der Verbindung als Erklären, warum die Aussagen aus verschiedenen Dokumenten auf eine bestimmte Weise zueinander stehen. Auch Leistungsdaten im Sinne von Produkten wurden erhoben. Im Rahmen der einfacheren Aufgabe sollten die Studierenden ohne Zugriff auf die drei Dokumente sechs Karten mit Aussagen danach sortieren, ob diese Aussagen in einem der Dokumente vorgekommen waren oder in mehreren. Hier wurden die richtigen Aussagen als Punktwert in der geschlossenen Aufgabe gezählt. Bei der schwierigeren Aufgabe sollten sie ein Diagramm zeichnen, welches über die Verhältnisse der Aussagen aus den gelesenen Dokumenten informieren sollte. Dabei konnten die Testpersonen die acht Dokumente zur Aufgabenbewältigung nutzen. Gezählt wurde hier die Anzahl der Verbindungen im Diagramm als

offenes Aufgabenformat für Integrationsleistungen. Außerdem wurde eine theoriebasierte Klassifikation der Diagramme in Typen von mentalen Modellen vorgenommen, auf die hier aber nicht gesondert eingegangen wird.

D) Was waren die Hauptbefunde?

Die Prozessdaten standen in dieser Studie im Vordergrund und lieferten reichhaltige Befunde. Bei der *Anzahl der Integrationen* ergaben sich Differenzen in der durchschnittlichen Anzahl von Verbindungen bei hoher Streuung: Bei der einfachen Aufgabe waren es knapp vier (3,9) Verbindungen, davon bezogen sich 87 Prozent auf Verbindungen zwischen zwei Dokumenten; bei der komplexeren Aufgabe waren es mehr als doppelt so viele (8,6), 57 Prozent inkludierten nur zwei Dokumente, die restlichen mindestens drei Dokumente. Die komplexere Aufgabe führte also zu mehr integrativen Aussagen, die zudem mehr gelesene Dokumente umfassten.

Bei der *Spezifität bzw. den Bezugspunkten* überwogen in beiden Aufgaben Aussagen, die entweder thematische Aussagen kombinierten oder Aussagen untereinander abglich. Kontextuelle Integrationsleistungen kamen demgegenüber seltener vor.

Auch die *Art der intertextuellen Verhältnisse von Informationen* war in beiden Aufgaben ähnlich: Gut drei Viertel der identifizierten Verbindungen wurden als Übereinstimmungen eingeordnet, im Falle der schwierigeren Aufgabe wurde ein größerer Anteil von Verbindungen als Widerspruch erkannt (14 Prozent vs. 5 Prozent bei der einfacheren Aufgabe). Die *Klärung der intertextuellen Verhältnisse* unterschied die beiden Aufgaben hingegen deutlich. Relationale Bezüge als niedrigste zu kodierende Stufe stellten mehr als die Hälfte (52 Prozent) der Inferenzen in der einfacheren Aufgabe dar, in der komplexeren Aufgabe waren es 37 Prozent. Bei den separaten Repräsentationen waren die Differenzen ebenfalls vorhanden (27 vs. 16 Prozent). Besonders markant wurden die Differenzen in den beiden höchsten Inferenzkategorien: Bei der einfachen Aufgabe entfielen 21 Prozent der Integrationen entweder auf erkannte simultane Relationen oder relationale Elaborationen, im Falle der schwierigeren Aufgabe waren dies jedoch 48 Prozent. Demnach führten komplexere Aufgaben zu mehr und zu abstrakteren Integrationsleistungen.

Die *Produktdaten* ergaben eine relativ hohe Korrelation zwischen beiden Tests ($r = .51$), das heißt, geschlossene und offene Aufgabe messen mit dem Fokus auf das intertextuelle Integrieren etwas Ähnliches. Bemerkenswert ist jedoch, dass die Anzahl von verbal geäußerten intertextuellen Integrationen unkorreliert war mit der jeweiligen Testleistung.

E) Diskussion

Diese Studie zeigt, dass integrative Aussagen innerhalb der Bearbeitung von Aufgaben zum Zweck des intertextuellen Integrierens auf verschiedene Art differenziert kodierbar sind und sich in den Auswertungen Aufgabenmerkmale bemerkbar machen, da sich sämtliche Probandinnen und Probanden beiden Aufgaben widmeten. Die einfachere der beiden Aufgaben hat im Vergleich zu weniger Inferenzen geführt, die sich zudem mehrheitlich auf die Verbindungen von zwei Dokumenten stützten. Auch fiel diese Aufgabe dadurch auf, dass fast vier Fünftel der Aussagen eher oberflächliche Verstehensleistungen demonstrierten,

weil sie überwiegend Inhalte separat und additiv beinhalteten oder nur vage thematisch verknüpften. Bei der schwierigeren Aufgabe mit mehr als doppelt so viel zu lesenden Dokumenten ergab sich ein gegenläufiges Muster. Nicht nur erwähnten die Studierenden mehr Dokumente in einem größeren Anteil der integrativen Aussagen, sie konnten die Inhalte aus Dokumenten zusätzlich thematisch stärker bündeln und erklären. Ähnlich war in beiden Aufgaben, dass es den Testpersonen leichter fiel, Gemeinsamkeiten zwischen den Dokumenten zu identifizieren, die sich zudem mehr auf den Zusammenhang von einzelnen Aussagen bezogen denn auf übergeordnete Themen.

Solche Befunde sind wichtig, weil sie die Aufgabenmerkmale noch stärker in den Fokus rücken und die Differenzierbarkeit und Differenzierungsnotwendigkeit bei Integrationsprozessen betonen. Selbiges gilt für die nach wie vor teils widersprüchliche Befundlage zu Integrationsprozessen und -produkten beim Lesen multipler Dokumente.

Die prozessbezogene Auswertung von (intertextuellen) Inferenzen folgt teils ähnlichen Zugängen wie die produktbezogenen Auswertungen und Taxonomien. Im Vergleich zu den Produktmaßen fallen die empirischen Ergebnisse weniger zahlreich aus, doch die Klassifikationen erlauben eine Differenziertheit, welche ertragreich wirkt, das Zusammenspiel von verschiedenen kognitiven Prozessen beim intertextuellen Kohärenzaufbau (unter Berücksichtigung des intratextuellen Verstehens) auszuleuchten und die Lesetheorie anzureichern (List, 2020). Allgemein variieren die prozessualen Inferenzen in ihrer Granularität und ihren Bezugspunkten. Das wiederum ist abstrahiert eine wichtige Überschneidung zu Produktmaßen.

Zusammenhänge von Prozessen und Produkten

Obwohl Prozesse und Produkte der intertextuellen Inferenzbildung jeweils für sich eine stabile Basis an Forschungsliteratur aufweisen, ist auffällig, dass Untersuchungen, welche sich parallel Prozessen und Produkten widmen, bislang in überschaubarem Ausmaß vorgelegt wurden. Die Vorstellung dieser empirischen Ergebnisse erfolgt dreigeteilt: Erstens werden Erträge des Kreierens externer Repräsentationen als prozessnahe Zwischenprodukte beim Lesen betrachtet. Zweitens lassen sich auch nachträglich aufgabennah die Prozesse erfassen und zwei Zugänge benennen. Drittens wurden anhand mehrerer Strategien Vorgehensweisen extrahiert und teils kontrastiert.

a) Notizen und Hervorhebungen

Notizen und Hervorhebungen sind zwei Strategien, die sich dadurch auszeichnen, dass sie von außen beobachtbare Spuren hinterlassen. Sie werden darum als *Strategien für externale Repräsentationen* bezeichnet (Kobayashi, 2007) und erfüllen verschiedene Funktionen, die sich dem generativen Lernen und seinen

drei Hauptprozessen der Informationsauswahl, kohärenten Organisation und Anreicherung und Verbindung mit dem Vorwissen (Philipp, 2021a) bzw. dem Verstehen von (Sach-)Texten subsumieren lassen (Mayer, 1996). Das Hervorheben von Textbestandteilen wie Unterstreichen oder farbiges Markieren dient beispielsweise dazu, Informationen auszuwählen und ihre Salienz zu erhöhen, was zu weiteren tiefenorientierten Verarbeitungsprozessen führen kann, aber nicht muss (Ponce et al., 2022). Das Anfertigen von Notizen kann ebenfalls die Funktionen des Hervorhebens erfüllen, einzelne Informationen mit höherer Salienz zu versehen und das Arbeitsgedächtnis zu entlasten, indem notierte Inhalte eine externe Repräsentation darstellen (Peverly & Wolf, 2019). Anders als bei den Hervorhebungen können jedoch auch potenziell mehr Restrukturierungen bzw. Formen der makrostrukturellen Verstehensleistungen beim Textverstehen über reduktive Verarbeitungen erfolgen (Philipp, 2021a). Hinzu kommen Optionen für die Leserinnen und Leser, verschiedene Funktionen in den Notizen wie kritisches Kommentieren zu realisieren. Die eigenen schriftlichen Annotationen erlauben damit dezidiert Inferenzen, die sich in der eigenen schriftlichen Repräsentation manifestieren und dort für spätere Verarbeitungszyklen zur Verfügung stehen. Die Befunde sind zusammengefasst:

- *Die Effekte des Anfertigens von Notizen sind uneindeutig und bedürfen einer genaueren Analyse, welches Notizenmaß und welche Art von kognitiver Transformation fokussiert wurde.* Lee (2022) ermittelte zum Beispiel, dass Personen dann ein besseres intertextuelles Verstehen erzielten, wenn sie ihre Notizen gebündelt so anfertigen konnten, dass sie tabellarisch strukturiert und gebündelt an einem Ort angefertigt wurden. List und Lin (2023) werteten die informatorischen Bezugspunkte bzw. Spuren kognitiver Transformationsprozesse in den Notizen aus und ließen diese insgesamt sechs Variablen als Prädiktoren für intertextuelle Verifikationsaufgaben einfließen. In dieser Analyse ergab sich nur ein positiver Effekt: Hatten Personen mehr kategorisierende, die Textinhalte abstrahierend bezeichnende Notizen hergestellt, erreichten sie bessere Testwerte bei der Verifikationsaufgabe. Zusätzlich war in einer weiteren Regressionsanalyse die Anzahl aller Notizen ein positiver Prädiktor korrekter intertextueller Inferenzbildung. Anders gesagt: Positiv wirkte sich in dieser Studie aus, dass die Personen viele bzw. abstrahierende Notizen angefertigt hatten. Dieser Befund ist deshalb von Belang, weil in der Studie nur ein Drittel aller Testpersonen etwas notiert hatte und sich trotzdem positive Effekte einstellten. In einer anderen Studie (Hagen et al., 2014) mit mehrdimensionaler Auswertung der Notizen mit insgesamt fünf Variablen gab es einen anschlussfähigen Befund in Korrelationsanalysen: Für das intertextuelle Verstehen erwies es sich als günstig, wenn Personen in den Notizen dokumentenübergreifende Inferenzen gebildet hatten. Alle anderen Variablen waren nicht-signifikant korreliert. Wenn auch nicht statistisch

signifikant, aber inhaltlich bemerkenswert ist, dass Personen, die viele inhaltliche Paraphrasen in ihren Notizen hatten, tendenziell schlechtere Inferenzleistungen erbrachten. Es scheint also auf die Art der Inhalte in den Notizen anzukommen. Relativierend muss auch hier angemerkt werden, dass ebenfalls lediglich ein Drittel der Personen Notizen produziert hatte.

Zu anders gelagerten Befunden kommen andere Untersuchungen. Keine korrelativen Effekte für das Leseverstehen mit einem Inferenzen-Gesamtscore gab es bei zwei Notiz-Maßen in der Studie von Castells et al. (2023): Es war weder von Vorteil noch von Nachteil, überhaupt Notizen angefertigt zu haben oder im Falle des Anfertigens zu verbatimem Übernehmen geneigt zu haben. Diese recht grobe und vor allem oberflächenorientierte Auswertung ergab keine Vorteile des Notierens.

Damit scheint es nicht auf das Notieren als solches anzukommen, vielmehr sind *spezifische Notizen für das Herstellen von intertextuellen Bezügen hilfreich*, nämlich solche, bei denen *Informationen über Dokumentengrenzen hinweg intertextuell kombiniert werden bzw. intratextuell Rückschlüsse auf Abstraktionsleistungen (Kategorisierungen) zulassen*. Wenn damit Indikatoren tieferer Verarbeitungsprozesse vorliegen, profitieren die Personen später in Testsituationen. Dieser Befund wird unterstützt aus dem Bereich der Diskurssynthesen-Forschung (s. Teilkap. [5.2.3.2](#) und dort den Abschnitt „Spezifische Schreibprozesse I: Planungsaktivitäten“).

- *Der wenig erforschte Nutzen des Hervorhebens von Informationen in multiplen Dokumenten fällt differenziert aus.* Die Möglichkeit, multiple (digitale) Texte mit Hervorhebungen zu bearbeiten, führt per se nicht zu besseren intertextuellen Inferenzleistungen in offenen Fragen (Li et al., 2016). Dafür konnten List und Lin (2023) mit ihren spezifischen Analysen der getätigten Hervorhebungen differenziertere Ergebnisse erzielen. Sie unterschieden fünf Arten von hervorgehobenen Informationen in den digitalen Dokumenten. Neben Metadaten handelte es sich um vier weitere: Hauptideen, Details, Schlüsselbegriff-Definitionen und statistische Informationen. Die am meisten hervorgehobenen Elemente waren die statistischen Informationen, Schlüsselbegriffe und die Hauptideen. Für die Inferenzleistungen erwiesen sich in zwei separaten Regressionsanalysen jeweils zwei Variablen als positive Prädiktoren. Zum einen waren die Inferenzleistungen besser, wenn die Personen mehr statistische Informationen hervorgehoben hatten, zum anderen war es die Gesamtzahl der Hervorhebungen, die sich als vorteilhaft erwies. Für diesen Summenwert war freilich die Anzahl der statistischen Informationen in den Hervorhebungen der größte Summand. Wie schon bei den Notizen verdeutlichen die feinteiligeren Analysen, dass differenzielle Effekte bestehen. Vor allem ist die Forschungssituation noch ausbaufähig, da kaum Studien dazu vorliegen, wie Hervorhebungen Inferenzbildungen begünstigen oder beeinflussen können (s. aber Ponce et al., 2022, für das Verstehen von

singulären Texten, sowie Mason et al., 2024, für positive Effekte von makrostruktursensiblen Hervorhebungen auf Transferverstehensleistungen beim Lesen digitaler Texte).

Empirisch ist der Nutzen jener Prozesse, mit denen Personen externe Repräsentationen gelesener Inhalte aus multiplen Texten kreieren, für das erfolgreiche Herstellen von intertextuellen Inferenzen (in geschlossenen, rezeptiven Tests) nur ansatzweise untersucht. Dieser Sachverhalt gestaltet sich anders für expressive Aufgaben, allen voran die Diskurssynthesen, und doch konvergieren die Ergebnisse: Notizen mit Indikatoren einer Verarbeitung jenseits der reinen Übernahme von Informationen erweisen sich als günstige Vorleistungen für später erfasste inferenzielle Verstehensleistungen.

b) Selbstauskünfte zum intertextuellen Integrieren: Informationsakkumulation sowie intertextuelle Elaboration

Ein aus Norwegen stammendes Instrument wurde von Bråten und Strømsø (2011) entwickelt. Es erfasst retrospektiv für Aufgaben mit multiplen Dokumenten zwei Vorgehensweisen, die unabhängig voneinander bestehen (Bråten & Strømsø, 2011; Hagen et al., 2014; List, Du, Wang & Lee, 2019, Studie 2; Tarchi, 2019). Eine Vorgehensweise wird über die Skala „intertextuelle Elaboration“ abgebildet und bezeichnet ein intertextuelles Integrieren, welches über Items wie dieses erfasst wurde: „Ich habe versucht, herauszufinden, wie der Inhalt der verschiedenen Texte zusammenhängt.“ Die zweite Skala („Informationsakkumulation“) operationalisiert das Vorgehen, Inhalte in ihrer Breite aufzunehmen, über Aussagen wie „Ich habe mich darauf konzentriert, mir so viele Fakten wie möglich aus allen Texten zu merken“. Diese Skala wurde im Originalwortlaut oder in weiterentwickelten Skalen sowohl in Studien mit rezeptiven Tests als auch in solchen mit expressiven Tests verwendet. Die Ergebnisse aus Studien mit expressiven Tests sind in Teilkapitel [5.2.2.2](#) im Abschnitt „Spezifische Leseprozesse 1: Integrationsprozesse“ Gegenstand. Die Resultate aus Studien mit rezeptiven Leseaufgaben folgen an dieser Stelle.

Wer in Selbstauskünften der Skala „*intertextuelle Elaboration*“ hohe Werte aufweist und damit angab, mehr um intertextuelle Inferenzen bemüht gewesen zu sein,

- erkannte mehr intertextuelle Inferenzen in Verifikationsaufgaben korrekt (Bråten & Strømsø, 2011; List, 2021),
- hatte höhere intra- und intertextuelle Verstehensleistungen erreicht, die in kombinierten Skalen zusammengefasst wurden (Follmer & Tise, 2022; List, 2021),

- konnte korrekter in Differenzierungsaufgaben Aussagen ihren Dokumenten zuordnen (List, 2021).

Wer hingegen auf der Skala „*Informationsakkumulation*“ höhere Werte aufwies und damit nachträglich berichtete, sich stärker der Sammlung von isolierten Informationen gewidmet zu haben, erzielte schlechtere Leistungen in der richtigen Beantwortung von intertextuell basierten Aussagen in Verifikationsaufgaben (Bråten & Strømsø, 2011).

Die Ergebnisse verweisen selbst bei einer recht dünnen Datenbasis und nur wenigen Studien auf ein Muster. Es deuten sich sowohl Vorteile beim Inferieren als auch beim Verknüpfen von Aussagen mit ihrer Quelle, also den Quelle-Inhalt-Verbindungen, als den kognitiven Produkten an. Damit sind sowohl die Ebene des integrierten mentalen Modells als auch die des Intertextmodells im Dokumentenmodell angesprochen. Das aufwändigere Vorgehen vor allem der intertextuellen Integration korrespondiert also mit besserem Verstehen in diesem Segment der Verstehensleistungen. Das gilt ebenfalls für expressive Tests (s. Teilkap. [5.2.2.2](#)).

c) *Verbünde von Strategien*

Eine dritte Gruppe von Studien, zahlenmäßig handelt es sich um die kleinste Gruppe, hat geprüft, ob der Einsatz mehrerer Strategien mit besseren intertextuellen Integrationsleistungen einhergeht:

- Die erste hier berichtete Studie ging der Frage nach, ob *kritisches Denken* und intertextuelles Leseverstehen miteinander zusammenhängen, wobei das kritische Denken nicht direkt mit Prozessdaten erhoben wurde (List, 2023b, Studie 3). Das Design sah es vor, dass die Testpersonen unmittelbar nach den zu lesenden Dokumenten aufgefordert waren, die argumentative Stärke des jeweiligen Dokuments einzuschätzen und ihre Urteile zu begründen. Ausgewertet wurde, wie viele kritisch-analytische Einschätzungen sie getroffen hatten, und es wurde ein Gesamtwert gebildet, der als Indikator für das kritische Denken beim Lesen galt. Je stärker ausgeprägt die Werte dieser Variable waren, desto besser waren die erwachsenen Testpersonen dazu in der Lage, Aufgaben des intertextuellen Verifikationstests erfolgreich zu bewältigen. Dieser Effekt bestand noch über andere Merkmale wie das Vorwissen hinaus, er ist aber kein über Prozessmaße abgesicherter Befund.
- Eine weitere Studie (Karimi & Shabani, 2013) traf eine Unterscheidung von besseren und schwächeren Leserinnen aufgrund diverser Tests, in denen die Lesefähigkeit erfasst und das Vorwissen kontrolliert wurde. Die Gruppen divergierten in ihrem Leseverstehen, aber nicht in ihrem Wissen zur Thematik der zu lesenden Texte. Während des Lesens wurden Lautdenk-Protokolle erhoben und die Gruppe der besseren Leserinnen mit jener der schwächeren

lesenden Studentinnen verglichen. Die Strategien wurden zu zwei Kategorien verdichtet. Die *oberflächenorientierten* („pragmatisch“ benannten) *Strategien* bestanden aus dem Notieren, Hervorheben, Unterstreichen und wiederholtem Lesen. Die *tiefenorientierten* („analytischen“) *Strategien* umfassten 16 Einzelstrategien wie Inferieren, Suchen, Überprüfen, Evaluieren etc. In beiden Gruppen war das Muster gleich: Bessere Leserinnen nutzten mehr Strategien. Da das intertextuelle Integrieren zwar erfasst, aber die Werte in der Studie nicht berichtet wurden, ist diese zweite Studie als indirekter Hinweis auf den Einsatz multipler Strategien beim erfolgreichen Lesen zu werten.

Der Einsatz von tiefenorientierten Strategien, der sich bereits bei den Selbstausskünften aus dem Abschnitt zuvor als günstig erwiesen hat, wird mittelbar von den Ergebnissen der hier berichteten Studien unterstützt. Jedoch fehlt es noch an Studien, die methodisch eindeutig Prozess- und Produktdaten bei Gruppen von Strategien und intertextuellen Integrationsleistungen aufeinander beziehen.

2) Intertextuelle Konflikterkennung

Intertextuelle Konflikte bilden einen Hauptgegenstand der gesamten Multiple-Documents-Literacy-Forschung und diverser Modellierungen (Philipp, 2018b; Saux, Britt et al., 2021). Denn intertextuelle Konflikte – gemeint sind damit diskrepante oder konfligierende Perspektiven oder Sichtweisen auf einen Sachverhalt in multiplen Dokumenten (Bråten & Braasch, 2018) – bilden für Leserinnen und Leser ein Hindernis beim Aufbau von Kohärenz. Entsprechend fokussieren Modelle darauf, wie Leserinnen und Leser damit verfahren, wenn Aussagen einander widersprechen. Eine Bedingung dafür ist, dass die Widersprüche als solche erkannt werden. Entsprechend werden in Studien, die diesen Spezialfall der intertextuellen Kohärenzherstellung unter erschwerten Bedingungen behandeln, Prozesse des intertextuellen Erkennens von Divergenz und Widerspruch als Ausgangspunkt für weitere Prozesse und Produktmaße erfasst und behandelt. Darum sind solche Studien mit Prozessdaten eher im Teilkapitel [4.2.2.2](#) Gegenstand, während sich dieser Abschnitt primär Produktdaten – seien es zunächst die rezeptiven, seien es im Anschluss die expressiven – widmet und danach die empirischen Befunde zu Zusammenhängen von Prozessen und Produkten behandelt.

a) Rezeptive Produktmaße (Verifikationsaufgaben)

Ein produktbezogenes und rezeptives Instrument, welches in einer Reihe von Studien zum Einsatz gelangt ist, ist die *Konfliktverifikationsaufgabe*. Die Operationalisierung funktioniert analog zur Inferenzverifikationsaufgabe (s. oben den Abschnitt 1) Intertextuelle Inferenzen und dort bei den Produkten den Unterabschnitt a) Rezeptives Beurteilen der Zulässigkeit von intertextuellen Inferenzen

bei rezeptiven Integrationsleistungen): Personen erhalten eine Batterie von Aussagen, die sie aufgrund der gelesenen Dokumente beurteilen sollen. Der Bezugspunkt der Beurteilung ist – zumindest in der breit verwendeten Variante, die von Stadtler, Scharrer, Brummernhenrich und Bromme (2013) entwickelt worden ist, – das Vorhandensein von intertextuellen Konflikten. Dazu werden Aussagen präsentiert, bei denen die Personen zunächst einschätzen müssen, ob sie in den gelesenen Texten vorgekommen sind. In einer Folgefrage beurteilen sie dann, ob diese Aussagen mit gelesenen Aussagen inhaltlich konfligierten. Jede Aussage muss also doppelt geprüft werden. Diese Aussagen unterscheiden sich darin, ob die intertextuellen Konflikte tatsächlich bestanden haben, denn das trifft nur für einen Teil der Items zu. Damit lassen sich echte Konflikte von fälschlicherweise als konfliktuös deklarierten Aussagen trennen und auswerten.

Die Studien divergieren, wie sie das Erkennen von intertextuellen Konflikten als Variable modellieren. Zwei grundsätzliche Vorgehensweisen lassen sich unterscheiden:

- In einigen Studien wurden die richtigen Antworten als Summenwert gebildet (Braasch et al., 2021; Kobayashi, 2015, Experiment 1; Schuster et al., 2023). In dieser Variante ist ein hoher Gesamtwert auf der Skala Indikator der Verstehensleistung.
- Andere und bislang auch mehr Studien nutzten ein anderes Vorgehen, indem die Anteile falscher Antworten in den Tests von Anteilen richtig beantworteter Items abgezogen wurden; es gab hier also eine Korrektur und damit ein relatives Maß (Delgado et al., 2020; Kammerer et al., 2019; Keck et al., 2015; Stadtler et al., 2020; Stadtler, Scharrer, Brummernhenrich & Bromme, 2013; Stadtler et al., 2014; Stang Lund et al., 2017, 2019). Dementsprechend werden die Anteile falsch positiv erkannter Aussagen von denen Anteilen der korrekt positiven Aussagen subtrahiert, sodass die Personen einen Wert zwischen -1 (keine richtigen Antworten gegeben, aber alle falsch positiven Items als richtig bezeichnet) und 1 (nur korrekte Antworten, keine falsch positiven Aussagen) aufweisen. Je höher der Wert innerhalb dieser Spanne ist, desto korrekter ist das Verständnis der intertextuellen Konflikte gewesen.

Die beiden Varianten unterscheiden also, wie sensibel Personen nicht nur für die intertextuellen Konflikte sind (erste Variante), sondern ob sie zusätzlich auch dazu in der Lage sind, zwischen echten Konflikten und nur vermeintlichen zu differenzieren (zweite Variante).

b) Expressive Produktmaße (zweiseitige Konfliktberichte in Texten)

Eine weitere Form, die intertextuellen Konflikte zu erfassen, besteht darin, sie zum Gegenstand einer expressiven Aufgabe zu machen. Studien nutzen bislang zwei Arten von Tests:

- So setzten Stadtler et al. (2020) auf der einen Seite einen Test ein, in dem ein Wiki-Eintrag zu der kontroversen Thematik korrigiert werden musste. Der Wiki-Text beinhaltete bereits die Themen, die Gegenstand von Divergenz waren, aber er berichtete nur einseitig davon. Wenn es den Testpersonen gelang, diese einseitige Sicht bei den vorab definierten Themen um die kontroversen Aspekte zu ergänzen und damit die Strittigkeit darzulegen, galt dies als Indiz dafür, dass sie die Konflikte erkannt und memoriert hatten.
- Weitere expressive Aufgaben auf der anderen Seite zählen aus, ob in Texten der Testpersonen die Konfliktthemen dezidiert als mehrperspektivisch dargestellt werden, sodass die Anzahl zweiseitig berichteter Konfliktthemen die Variable für intertextuelle Konflikterkennung bildete (Stadtler, Scharrer, Brummernhenrich & Bromme, 2013; Stadtler et al., 2014). Eine weitere Form der Auswertung besteht darin, dass binär erfasst wird, dass eine Kontroverse explizit benannt wurde (Schoor et al., 2023).

Die zweite Variante wirkt kognitiv anspruchsvoller, da die lesende Person selbst die Konfliktthemen erkennen und berichten muss, während bei der Korrektur eines Textes zumindest ein Teil der Leistung angelegt ist.

c) Zusammenhang von Prozess- und Produktmaßen

Die Studienlage bei der intertextuellen Konflikterkennung unterscheidet sich von jener zum intertextuellen Inferieren in der Datenbasis. Insbesondere Studien, die Prozess- und Produktindikatoren zusammen erfasst und aufeinander bezogen haben, sind bislang selten vorgelegt worden und teils Gegenstand der Forschung zu Diskursynthesen (s. dazu Teilkap. [5.2.1.2](#) und [5.2.2.2](#)). Die Ergebnisse aus drei Studien lauten:

- *Längere Lesezeiten und Wechsel zwischen Texten stehen in einem positiven Verhältnis dazu, ob jemand explizit Konflikte berichtet.* Dies ist das Ergebnis der Studie von Schoor et al. (2023). Demnach korreliert das binäre Produktmaß des Maßes für explizites Ausweisen des Konfliktthemas damit, dass Personen die Dokumente länger gelesen haben und häufiger zwischen ihnen hin und her gewechselt haben. Diese Logfile-Daten verdeutlichen damit, dass längeres und nicht-lineares Lesen auf tieferes Verarbeiten verweist. Gerade die Wechsel sind von Belang, weil sie als Indikator eines textvergleichenden

Lesens gelten, welches als „Corroboration“ bezeichnet wird. Die Indikatoren in der Studie enthalten allerdings anders als die sogleich vorzustellenden Studien keine Hinweise darauf auf, dass gezielt die inhaltlich konfligierenden Textteile konsultiert wurden. Andere Studien haben dies methodisch über andere Studiendesigns realisiert.

- *Nicht-reaktive Prozessdaten ergeben, dass eine vertiefte Verarbeitung von konfligierenden Informationen zu besseren Leistungen bei der korrekten Konfliktdetektion führt.* Eye-Tracking-Daten zeigen: Personen waren sowohl in expressiven als auch in rezeptiven Aufgaben besser darin, Konflikte korrekt zu erkennen bzw. zu berichten, wenn sie beim Auftreten des Konflikts längere Lesezeiten aufgewiesen haben. Dies war das übereinstimmende Resultat aus den Eye-Tracking-Daten in der Studie von Stadtler et al. (2020). Dabei war die Fixationsdauer bei den Erklärungen zum biologischen Mechanismus, der den Konflikt betraf, besonders entscheidend. Lasen die Personen also die Details, welche die konfligierenden Gründe zu einer vorher gelesenen Aussage erläuterten, waren ihre Gedächtnisleistungen besser. Methodisch anders, aber inhaltlich komplementär war das Ergebnis von Delgado et al. (2020). In der Studie wurden rezeptive Maße genutzt, und die Versuchsanlage sah es vor, dass die Personen am Ende von Absätzen Hyperlinks aufriefen, die zu den konfligierenden Aussagen in anderen Dokumenten führten. Sie hatten also die Option, ganz gezielt Aussagen aus konfligierenden Dokumenten des digitalen Dokumentensets anzuwählen. Je häufiger sie dies taten, desto besser waren ihre Ergebnisse in der rezeptiven Aufgabe, wobei dieser Effekt dadurch zustande kam, dass die Testpersonen insbesondere jene Items korrekter lösten, welche auf den intertextuellen Konflikten mit den Hyperlinks basierten. Gezieltes Konsultieren steigerte also die Leistungen (und mittelbar die Fähigkeit, sich korrekt an die Quellen von Aussagen zu erinnern).

Es ergibt sich damit ein Muster in den Studien mit unspezifischen Prozessmaßen: Wenden sich Personen länger und häufiger jenen Informationen zu, die das intertextuell konfligierende Verhältnis verursachen, dann erkennen bzw. memorieren sie diese Konfliktthemen besser. Diese Ergebnisse stellen eine Ergänzung zu den Zusammenhangmaßen von Prozessen und Produkten der intertextuellen Inferenzbildung bzw. Integrationsleistungen dar.

4.2.1.2 Integration von Metadaten und Textinhalten

Das Integrieren von propositionalen Inhalten und Metadaten ist aus der Perspektive des Dokumentenmodells entscheidend dafür, die beiden Ebenen des integrierten mentalen Modells und des Intertextmodells miteinander kognitiv zu verbinden. Darum sind die gleichsam vertikalen Verbindungen, die Quelle-Inhalt-Verbindungen als einzelne Elemente eine Variable von Interesse. Sie stehen

deshalb in der folgenden Darstellung an erster Stelle der zweigeteilten Darstellung. Ihnen folgt eine zweite und deutlich umfassendere Repräsentationsleistung, nämlich die Fähigkeit, Dokumentenmodelle mental zu repräsentieren.

1) Quelle-Inhalt-Verbindungen

Das umfassende Verstehen multipler Dokumente setzt nicht nur intertextuelles Integrieren von Inhalten voraus, welches im Teilkapitel [4.2.1.1](#) betrachtet wurde. Eine zentrale Prämisse des Verstehens von multiplen Dokumenten besteht vielmehr darin, Inhalte und deren Quelle miteinander kognitiv zu repräsentieren, um so spezifische Beiträge von einzelnen Dokumenten zu gesamthaften Repräsentationen gemäß dem Dokumentenmodell separieren zu können (Saux, Britt et al., 2021). Diese Fähigkeit wird von manchen Autorinnen und Autoren dem Sourcing zugerechnet (Rouet & Britt, 2022), allerdings lässt sich durchaus argumentieren, dass sich die Repräsentation von Quelle und Inhalt ebenfalls als Bemühen der Leserinnen und Leser um Kohärenz werten lässt (Strømsø & Bråten, 2010). Das gilt insbesondere bei konfligierenden Dokumenten (s. Teilkap. [4.2.2](#)).

Quelle-Inhalt-Verbindungen bezeichnen die Integrationsleistung, eine Aussage ihrem Ursprung zuzuordnen und diese Verbindung von Metadatum und Aussage mental zu repräsentieren (Saux, Britt et al., 2021). Diese Leistungen werden als *Produkt*maße in verschiedenen Studien auf zweierlei Arten erhoben (McCrudden et al., 2023; Saux et al., 2018):

- Die erste und in Studien häufiger zum Einsatz gekommene Möglichkeit besteht darin, *direkt* Metadaten und ihre Verbindung zu Inhalten zu erfragen, sei es als Zuordnungsaufgabe in Multiple-Choice-Tests (Strømsø et al., 2010), sei es als Aufforderung, die Quellen ohne Zugriff auf die Dokumente zu benennen und sich an so viele Inhalte wie möglich zu erinnern und diese den Dokumenten korrekt zuzuordnen (s. z. B. Steffens et al., 2014, Experimente 1 und 2). In solchen Aufgaben werden Quelle-Inhalt-Verbindungen demnach gezielt elizitiert und die Korrektheit der Leistungen bestimmt.
- Das ist anders bei den indirekten bzw. *unaufgeforderten Maßen*, denen es eigen ist, dass sie ex post ausgezählt werden und ohne direkte Aufforderung zustande kommen, z. B. in schriftlichen Aufgabenprodukten (Anmarkrud et al., 2022). Teils sind in solchen Studien jedoch die Quelle-Inhalt-Verbindungen nicht eine eigene Variable, sondern werden als Bestandteil übergeordneter Indikatoren des Leseverstehens behandelt (Hämäläinen et al., 2023; Kiili et al., 2021; Mason et al., 2020). Teils lassen sich nur die Häufigkeiten der Quelle-Inhalt-Verbindung auszählen, aber nicht anhand eines Maximalwerts beurteilen (Anmarkrud et al., 2014; Bråten, Ferguson et al., 2014). Hinzu kommt, dass bei diesen Auswertungen nicht die Korrektheit im Vordergrund steht,

sondern die Frequenz. Damit erfassen die beiden Vorgehensweisen nicht das gleiche Konstrukt.

Prozessbezogene Studien sind bisher kaum vorgelegt worden. Wenn die Quelle-Inhalt-Verbindungen ausgewertet wurden, handelte es sich um Indikatoren der unaufgeforderten Quelle-Inhalt-Verbindungen während der Aufgabenbearbeitung, die in Studien mit Lautdenk-Protokollen kodiert wurden (Barzilai et al., 2015). Damit ist aus Sicht der empirischen Forschung bereits ein Desiderat und ein Muster benannt: Es gibt ein Ungleichgewicht von Produkt- und Prozessmaßen zugunsten der Produktindikatoren (mit geschlossenen Antwortformaten und gezielter Aufforderung zum nachträglichen Verknüpfen von Inhalten und deren Quellen).

Produkte

Die Ergebnisse zu den Produktmaßen in den Quelle-Inhalt-Verbindungen sind bündelbar:

- *Die Leistungen divergieren erheblich, je nachdem ob die Quelle-Inhalt-Verbindungen explizit erfragt oder ohne Aufforderung von den Testteilnehmerinnen und -teilnehmern zu erbringen war.* Dieses Befundmuster ergibt sich deutlich aus den zusammengetragenen Ergebnissen aus der Tabelle 19. Auffällig ist natürlich, dass die Datenbasis divergiert, da offene Aufgaben erheblich weniger häufig gestellt wurden. Doch selbst wenn dies berücksichtigt wird, zeigen sich eklatante Differenzen. Diese spiegeln das Muster wider, welches Anmarkrud et al. (2022) in ihrem Forschungsüberblick über sourcingbezogene Studien berichten: Geschlossene Aufgabenformate gehen mit höheren Leistungen und einer höheren Streuung einher. Das lässt sich ebenfalls aus der Tabelle 19 in der Spalte mit den geschlossenen Aufgaben ermesen, da dort die Mittelwerte der geringsten und der höchsten Ausprägung der Mittelwerte um den Faktor 3 divergieren.

Tabelle 19: Studien aus der Multiple-Documents-Literacy-Forschung mit Operationalisierung der Quelle-Inhalt-Verbindungen und den berechneten Quoten der Leistungen, bezogen auf die Mittelwerte der jeweiligen Samples

Offene Aufgaben ohne Aufforderung		Geschlossene Aufgaben mit Aufforderung	
Studie	Prozentualer Anteil von Quelle-Inhalt-Verbindungen	Studie	Prozentualer Anteil korrekter Antworten
Braasch et al. (2016), Experiment 3 ***	3,8	Barzilai & Eshet-Alkalai (2015)	41,9
Florit, Cain & Mason (2020) ***	2,7	Delgado et al. (2020)	65,3
		Kammerer, Meier & Stahl (2016) *	21,3
		Lee et al. (2024)	47,8
		Leroy et al. (2023) **	57,1
		Leroy & Kammerer (2023) ***	59,0
		List (2021)	36,5
		Merkt et al. (2017)	58,8
		Merkt & Huff (2020, Experiment 1)	64,2
		Merkt & Huff (2020, Experiment 2)	66,2
		Paul et al. (2018, Studie 2)	56,5
		Paul et al. (2019)	59,5
		Salmerón, Sampietro & Delgado (2020) **	50,0
		Schuster et al. (2023) ***	46,5
		Stang Lund et al. (2017)	60,6
		Stang Lund et al. (2019)	65,3
		Stadtler et al. (2016) ****	51,9
		Strømsø et al. (2010)	37,6
Mittelwert	3,3	Mittelwert	52,6

(Legende: * = nur Werte der Kontrollgruppe; ** = gewichteter Mittelwert über alle Experimentalkonditionen aufgrund eigener Berechnungen; *** = Mittelwert über alle Experimentalkonditionen aufgrund eigener Berechnungen; **** = Mittelwert über alle Gruppen im Pretest)

- *Wer bessere Leistungen in der Repräsentation von Inhalten und ihrer Quelle erbringt, weist auch in anderen Verstehensleistungen bei expressiven, offenen Aufgaben bessere Werte auf.* Solche Leistungen betreffen verschiedene Integrationsleistungen wie schriftliche Antworten mit Gesamtindizes der *Integrationsleistungen*. Effekte wurden sowohl in Korrelations- (Bråten, Ferguson et al., 2014; Lee et al., 2024) als auch in Regressionsanalysen (Florit, Cain & Mason, 2020) ermittelt. Nicht in allen Studien war dies jedoch der Fall (Anmarkrud et al., 2014). Zusätzlich bestehen Hinweise aus Einzelstudien mit geschlossenen Testformaten, dass Quelle-Inhalt-Verbindungen und intertextuelles Integrieren aus Verifikationsaufgaben

zusammenhängen und erstgenannte letztgenannte regressionsanalytisch präzisieren (Strømsø et al., 2010). Ferner liegen erste Belege vor, dass bessere Werte in geschlossenen Testformaten und Indikatoren der Integrationsleistungen aus expressiven Aufgaben Hand in Hand gehen (Salmerón, Sampietro & Delgado, 2020): Wer sich besser an die Herkunft von Aussagen erinnert, ist besser dazu in der Lage, Informationen aus gelesenen Texten intra- und intertextuell miteinander über Inferenzen zu verknüpfen.

Evaluative Leistungen und das *Sourcing* hängen ebenfalls mit Produktmaßen der Quelle-Inhalt-Verbindungen zusammen. Wer Quellen und Inhalte gemeinsam korrekt repräsentiert, zitiert und referenziert mehr in eigenen Texten (Anmarkrud et al., 2014), kann Perspektiven von Dokumenten korrekter einschätzen (Barzilai & Eshet-Alkalai, 2015) bzw. ausgewogener berichten (Mason, Moè et al., 2022), die Glaubwürdigkeit von Dokumenten besser beurteilen (Barzilai & Eshet-Alkalai, 2015) und das eigene Glaubwürdigkeitsurteil besser begründen (Barzilai & Eshet-Alkalai, 2015).

Die Ergebnisse aus reiner Produktperspektive sind damit zusammengefasst: Es macht einen Unterschied in den Ergebnissen, ob die Leistungen als Skalen mit direkter Aufforderung der Zuordnung von Aussagen zu ihrem Ursprung oder als offene Aufträge formuliert werden, welche das Verbinden von Inhalten und Quellen implizit erfragen. Quelle-Inhalt-Verbindungen, die nicht explizit erfragt werden, bleiben in den Produktindikatoren bei den untersuchten Testpersonen randständig bis marginal. Jedoch ergaben sich Zusammenhänge vom explizit erfragten kognitiven Repräsentieren von Quelle-Inhalt-Verbindungen mit anderen Integrations-, Sourcing- und evaluativen Leistungen beim Verarbeiten von multiplen Dokumenten. Diese Zusammenhänge sind allesamt positiv und können so gewertet werden, dass mentales Kombinieren von Quelle und Inhalt innerhalb eines Ensembles mehrerer kognitiver Verstehensleistungen zu situieren ist.

Zusammenhänge von Prozessmaßen und Produkten

Die Prozessmaße sind ein Desiderat, darum werden hier zwei Perspektiven konturiert. Die erste geht von den marginal erforschten Prozessen des genuinen Kombinierens von Daten zur Quelle und Propositionen aus dem Inhalt aus, die mit Produktmaßen statistisch in Verbindung gebracht werden. Die zweite Perspektive ist anders gelagert: Sie versammelt Studienbefunde, in denen verschiedene Prozessindikatoren erhoben wurden, wobei diese Prozessdaten nicht genuin auf Quelle-Inhalt-Prozesse abzielten.

- *Prozessdaten zur Quelle-Inhalt-Verbindung während des Lesens korrespondieren mit Verstehensleistungen in offenen Aufgaben.* Es ergaben sich substantielle positive Zusammenhänge solcher Prozesse mit verschiedenen Maßen

aus eigenen argumentativen Texten in schriftlichen Antworten (Barzilai et al., 2015). Die Studie von Barzilai et al. (2015) bezieht sich auf Diskurssynthesemaße und wird darum in Teilkap. 5.2.2.2 behandelt.

- *Verschiedene Prozesse, die auf eine intensivere Verarbeitung der Inhalte hinweisen, hängen mit besseren Leistungen in der Quelle-Inhalt-Repräsentation zusammen.* Dies betrifft unspezifische Maße ebenso wie spezifische Maße (s. Teilkap. 5.2.2 für die Erläuterung der Unterschiede zwischen beiden Operationalisierungsweisen).

Zunächst zu den *unspezifischen Prozessindikatoren* und ihrem Zusammenhang mit Leistungen in geschlossenen Aufgabenformaten: Bereits eine längere Lesezeit ist ein Prädiktor für bessere Testergebnisse (List, 2021). Je häufiger Personen laut Logfile-Daten die zu lesenden Dokumente mehrfach konsultierten – ein Indikator für dokumentenvergleichendes Lesen –, desto höher waren ihre Testwerte in Quelle-Inhalt-Produktindikatoren (Leroy et al., 2023; Leroy & Kammerer, 2023). Auch das häufige eigenaktive Aufrufen von konfligierenden Aussagen korreliert positiv mit Testleistungen (Delgado et al., 2020). Die Resultate aus Studien mit *spezifischen Lesemaßen* in Lautdenk-Studien erbrachten zudem, dass Strategien der intertextuellen Verarbeitung und der Beurteilung von Inhalten mit besseren Integrationsleistungen in offenen Aufgaben korrespondierten (Anmarkrud et al., 2014). In geschlossenen Formaten erwies sich das intertextuelle Elaborieren als positiver Prädiktor der richtig gelösten Quelle-Inhalt-Verbindungsaufgaben (List, 2021).

Erstaunlicherweise sind die Prozesse des Verbindens von Quelle und Inhalt bei multiplen Dokumenten bislang nicht sonderlich intensiv erforscht worden, zumal hinsichtlich der Frage, welche produktbezogenen Verstehensleistungen sie nach sich ziehen. Wechselt man jedoch die Perspektive und überblickt, welche spezifischen und unspezifischen Leseprozesse mit Produkten der Quelle-Inhalt-Verbindungen zusammenhängen, zeigt sich ein Muster: Prozesse, die auf ein vergleichendes und teils kritisches Lesen schließen lassen, erweisen sich in korrelativen Studien musterhaft als günstig für die Produktmaße.

2) Dokumentenmodelle

Produktperspektive 1: reine Dokumentenmodelle

Dokumentenmodelle sind eine besonders umfassende kognitive Mehrebenen-Repräsentation multipler Dokumente (s. Teilkap. 4.1.2.3). Damit stellen sie lediglich eine unter verschiedenen Repräsentationen dar, welche Leserinnen und Leser aufbauen. Aufgrund der Vielzahl von potenziell zu integrierenden Informationen und Informationsarten aus den Dokumenten selbst, aber auch wegen des komplexen Zusammenspiels von Aufgabenstellung und individuellen Merkmalen

ergeben sich unterschiedliche konkrete Leistungen in den Artefakten der Studien. In Studien werden darum die meist schriftlichen, teils auch grafisch realisierten Lösungen bei Testaufgaben danach kodiert, inwieweit es den Testpersonen gelungen ist, inhaltliche Integrationsprozesse und die Nutzung und Verknüpfung von Metadaten mit Inhalten zu realisieren. Das übereinstimmende Ergebnis solcher Studien (List, 2019, 2023a; List, Du & Lee, 2021b; List, Du, Wang & Lee, 2019, Studien 1 und 2; List, Oaxaca et al., 2021) ist, dass Personen mehrheitlich nicht dazu in der Lage sind (s. dazu Teilkap. [4.4.1.1](#) und dort die Tabelle 27). Aus den über 600 ausgewerteten Aufgaben aus allen Studien konnten nur 32 Prozent als echte Dokumentenmodelle klassifiziert werden. Allerdings variierte dieser Mittelwert in den Studien teils eklatant.

Produktperspektive 2: Dokumentenmodelle und andere Produktmaße

In einigen der Studien wurden die Globalindikatoren der Diskurssynthese in Dokumentenmodell-Taxonomien bzw. andere komplexe Integrationsleistungen von der Ebene des integrierten mentalen Modells mit weiteren Produktmaßen in Beziehung gesetzt. Wer laut diesen Studien komplexere, stärker inferenzbasierte Repräsentationen aufgebaut und in seinem Produkt wiedergegeben hatte,

- hatte mehr integrierte Inhalte aus den verwendeten Texten genutzt (List, 2019; List, Du & Lee, 2021a; List, Du, Wang & Lee, 2019, Studie 1), was bedeutet, dass es mehr Inferenzleistungen in diesen Texten gab – Wiley et al. (2009) differenzieren dies aus, indem sie zeigten, dass solche umfassenderen integrierten mentalen Modelle deutlich mehr korrekte Inhalte enthielten und weniger falsche Konzepte inkludierten;
- zitierte mehr der zugrundeliegenden Dokumente aus dem zur Verfügung gestellten Set (List, Du, Wang & Lee, 2019, Studie 1);
- konnte Perspektiven aus verschiedenen Disziplinen der Dokumente klarer benennen und deren Methoden einschätzen (List, 2023a);
- hatte reflektiertere und kritischere inhaltliche Schlussfolgerungen formuliert und damit ein tieferes Leseverstehen demonstriert (Hartmann et al., 2021).

Komplexere kognitive Repräsentationen wie die Dokumentenmodelle gehen demnach mit anderen Leistungen einher, die auf eine breitere und korrektere Informationsverwendung zum einen verweisen und zum anderen ein inhaltlich tieferes Verstehen der Thematik nahelegen. Diese Zusammenhänge, welche studienübergreifend konsistent sind, verdeutlichen die hohen kognitiven Leistungen, welche Personen im Verbund zu erbringen imstande sind.

Experimente, die gleichzeitig Prozesse und Dokumentenmodelle fokussierten, sind rar. Doch gibt es Hinweise aus zwei Studien für die Bedeutung des intertextuellen Inferierens für umfassende kognitive Repräsentationen. Denn je mehr intertextuelle Beziehungen beim Lesen Personen bemerkten bzw. äußerten, desto mehr zeugten ihre später erbrachten produktbezogenen Integrationsleistungen auch tatsächlich von integrierten mentalen Modellen bzw. stellten Dokumentenmodelle dar (List, Du & Lee, 2021b; List, Du, Wang & Lee, 2019, Studie 2). In diesen Studien überwogen die thematischen Integrationen (s. dazu die Spezifität von Inferenzen im Teilkapitel [4.2.1.1](#)), also die Kombination von Propositionen aus multiplen Dokumenten zu einzelnen dokumentenübergreifenden Themen.

4.2.2 Integration angesichts inkohärenter, widersprüchlicher und fehlerhafter Informationen

Der Spezialfall inkohärenter bzw. konfligierender Inhalte beim (digitalen) Lesen hat neuere Theoriarbeiten nötig gemacht, die sich der Frage verschrieben haben, wie es trotz erschwelter Bedingungen glückt, aus inkohärenten (einander inhaltlich widersprechenden) Aussagen und Dokumenten mit eigenem kognitivem Aufwand eine kohärente Repräsentation aufzubauen. Neuere Theorien nehmen dies gezielt auf und gehen in der Regel davon aus, dass Inkonsistenzen von der lesenden Person entdeckt und durch verschiedene Kohärenzwiederherstellungsprozesse behoben werden. Dabei ist für eine gelingende Kohärenzleistung die Analyse der Konfliktursache(n) nötig, um dann über verschiedene optionale Prozesse mittels Elaborationen eine intertextuelle Kohärenzleistung zu erbringen.

4.2.2.1 Erste- und Zweite-Hand-Einschätzungen als funktionale Kategorien

Wer sich lesend im Internet informieren will, zumal zu kontroversen Themen oder solchen Gegenständen, auf die es divergierende Perspektiven gibt, muss damit rechnen, dass Aussagen sich nicht ohne weiteres kohärent aufeinander beziehen lassen. Die Gründe können vielgestaltig sein: Informationen und Dokumente können inhaltlich so fragmentarisch sein, dass Leserinnen und Leser separate, also unverbundene Repräsentationen aufbauen (Britt et al., 1999; Goldman, 2004). Dokumente können sich im- oder explizit widersprechen, also ein intertextuelles Kohärenzproblem aufweisen, welches Leserinnen und Leser erkennen und – je nach Leseziel und Wissensbeständen sowie epistemischen Überzeugungen – ignorieren oder auflösen (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014). Außerdem kann es sich um irreführende und falsche Aussagen handeln, die teils

mit manipulativer Intention verbreitet werden. Solche Desinformationen (Kendeou et al., 2019) als solche zu erkennen und sie gezielt nicht im Aufbau kognitiver Repräsentationen zu berücksichtigen, schützt davor, dass mentale Modelle und Gedächtnisinhalte mit schwer korrigierbaren Informationen verzerrt werden. Harris (2022, S. 5) nennt solche Aussagen eine „epistemische Bedrohung“, die einen Erwerb oder die Speicherung von Wissen unterminiert. In diesen Fällen sollten Leserinnen und Leser also Dokumente gezielt *nicht* verarbeiten und über die mehrdimensionalen Evaluationsprozesse (s. Teilkap. 3.1.3) filternd potenzielle, aber ungeeignet wirkende Dokumente und Aussagen ausschließen.

Der Umgang mit Widersprüchlichkeit und Divergenz von Perspektiven ist jedoch nicht zwangsläufig ein Umgang mit Falsch- oder Fehlinformationen. Mitunter werden Themen aufgrund ihrer Komplexität und Vielschichtigkeit bzw. einer Vielzahl heterogener Zugänge in multiplen Dokumenten behandelt (Goldman, 2015). Eine der erwartbaren Folgen ist, dass das intertextuelle Verhältnis dieser Dokumente konfligierend sein kann. Dies stellt aus kognitionspsychologischer Sicht ein Kohärenzproblem bzw. eine Erschwernis dar: Personen können nämlich nicht mit einer übergreifenden Kohärenzerwartung multiple Dokumente lesen, sondern müssen die Beziehung zwischen Aussagen, Quellen und Dokumenten im Leseprozess selbst über Inferenzleistungen bestimmen (Saux, Britt et al., 2021). Diese Kohärenzleistung bildet den Ansatzpunkt verschiedener theoretischer Modellierungen mit unterschiedlicher Art, die Kohärenzleistungen zu erklären und zu beschreiben (s. Abschnitt „Herausforderung Inkohärenz beim intertextuellen Integrieren“ im Teilkap. 4.1.2.3). Dreh- und Angelpunkt solcher Modelle ist die Frage, wie Personen trotz objektiv vorhandener Inkohärenz von Aussagen dennoch zu einer kohärenten Gesamtrepräsentation des konfligierend beschriebenen Sachverhalts gelangen können, welche Prozesse dazu führen und teils welche Bedingungen erfüllt sein müssen. Es wurden verschiedene solcher Modelle (s. Tabelle 20 und Philipp, 2023) vorgelegt; sie eint, dass eine erkannte und je nach Modell auflösungsbedürftige bzw. auflösbare Inkohärenz den Ausgangspunkt für Prozesse und Produkte des Umgangs mit dieser Inkohärenz bildet (s. dazu auch Teilkap. 3.1.2.3). Die drei Modelle aus Tabelle 20 sind vor allem aus der Perspektive von Belang, welche prozessualen Reaktionen die Modelle auf die durch Inkohärenz hervorgerufene kognitive Dissonanz vorsehen und welche Produkte erwartbar sind. Hierin variieren die Modelle in den Folgeschritten und deren Resultaten.

Tabelle 20: Vergleich dreier Modelle der Multiple-Documents-Literacy-Forschung mit Fokus auf Umgang mit inkohärenten Aussagen aus multiplen Dokumenten

Modell	Auslöser der Inkohärenz	Prozessuale Reaktion(en) der Lesenden auf Inkohärenz	Erwartbare Ergebnisse
Discrepancy-Induced Source Comprehension Model (Braasch & Bräten, 2017; Braasch & Kessler, 2021)	Erkannte Inkohärenz von Aussagen verschiedener Quellen	Sourcing	• <i>Prozesse</i>: vermehrte Sourcingprozesse • <i>Produkte</i>: stärkere kognitive Repräsentation von Metadaten (in Kombination mit Inhalten)
Two Step Model of Validation (Richter & Maier, 2017; Richter et al., 2020; in partieller Erweiterung: Richter & Tiffin-Richards, 2024)	Erkannte Text-Überzeugungs-Inkonsistenz	Epistemische Prozesse zur produktiven Überwindung einseitiger, überzeugungskonsistenter Inhalte	• <i>Prozesse</i>: Prozesse aus der Sammelkategorie der „epistemischen Elaboration“ (Integration überzeugungsinkonsistenter Inhalte, strategische Suche nach Zusatzinformationen, Sourcing, Corroboration) • <i>Produkte</i>: balanciertere Repräsentationen inhaltlich konfligierender Inhalte, insbesondere der überzeugungsinkonsistenten Aussagen und Perspektiven
Content-Source Integration Model (Braasch & Scharer, 2020; Stadtler & Bromme, 2014)	Erkannte und auflösungsbedürftige Inkohärenz von Aussagen verschiedener Quellen	Kohärenzwiederherstellung durch a) oberflächliches/falsches Verarbeiten (= keine Auflösung der Inkohärenz), b) Erste-Hand-Einschätzungen (Validieren), c) Zweite-Hand-Einschätzungen	• <i>Prozesse</i>: je nach Phase divergierende Prozesse der Informationsbeurteilungen und -verarbeitung wie Plausibilitätseinschätzungen und Sourcing • <i>Produkte</i>: je nach Phase mittels Elaborationen angereicherte Repräsentation der Inhalte mit ihren (mutmaßlichen) Dissensursachen oder per Evaluation vorgenommene Gewichtungen von Aussagen mit einer Auflösung der Konflikte

Angesichts der bereits aus den drei Modellen aus Tabelle 20 hervorgehenden Breite an Prozessen und Strategien ist eine integrative, modellübergreifende Analyse zielführend. Dazu ist eine Differenzierung von Prozessen hilfreich, die sich im „Content-Source Integration Model“ bereits deutlich niederschlägt: diejenige nach Einschätzungen, die aus erster Hand oder aus zweiter Hand erfolgen (Stadtler & Bromme, 2014). Diese Differenzierung hat ihren Ursprung in der Frage, wie Laienpersonen in zunehmend komplexen Gesellschaften mit ungleich verteilter Expertise in Wissensgebieten zu begründbaren und hinreichend sicheren Einschätzungen mit epistemisch unsicheren Informationen gelangen (Bromme

& Goldman, 2014; Bromme, Kienhues & Porsch, 2010; Keren, 2018). Dabei helfen zwei voneinander trennbare Herangehensweisen, welche der Beantwortung zweier Fragen dienen:

- *Was ist wahr?* Diese *Erste-Hand-Einschätzungen* dienen einer direkten Beurteilung von Aussagen, indem Personen die Kohärenz und Konsistenz von konfligierenden Aussagen prüfend beurteilen und sie mit ihren eigenen Wissensbeständen vergleichen. Es handelt sich um eine kritisch-analytische Vorgehensweise, die Bromme, Kienhues und Porsch (2010) als zentrales Element des kritischen Denkens bezeichnen. Demnach wären derartige Erste-Hand-Einschätzungen, die sich auf verschiedene Facetten von Aussagen beziehen, im Kontext größerer metatheoretischer Konstrukte wie kritischem Denken (Cartiff & Greene, 2020), kritisch-analytischem Denken (Alexander, 2023) und dem Denken höherer Ordnung zu situieren (List & Sun, 2023). Dies verdeutlicht zugleich ihren hohen Anspruch an die Fähigkeiten der lesenden Person, mehrdimensionale Gültigkeitszuweisungen vorzunehmen, namentlich hinsichtlich des nötigen Vorwissens und weiterer dispositionaler Fähigkeiten (Dwyer et al., 2014).
- *Wer ist vertrauenswürdig?* Sind die Erste-Hand-Einschätzungen aufgrund ihrer hohen Voraussetzungen keine realistische Option (z. B. weil Laien keine Aneignungschancen haben oder weil für das zu lösende Problem zu wenig Zeit besteht), dann wird aus der Frage nach der Wahrheit/Plausibilität und deren direkter Beurteilung seitens der lesenden Person eine andere Frage. Diese betrifft als *Zweite-Hand-Einschätzung* ein indirektes Gültigkeitsurteil, indem die Quelle von Aussagen beurteilt wird (Bromme, Kienhues & Porsch, 2010) und aufgrund dieser Grundlage die Plausibilität eingeschätzt wird (Lombardi et al., 2016; von der Mühlen et al., 2016). Diese Verlagerung des Urteilsgegenstands nimmt den Grundgedanken auf, dass Expertise und Wissen ungleich verteilt sind, sodass Quellen mit einer höheren epistemischen Autorität (Personen mit echter Expertise) eine höhere Glaubwürdigkeit zugeschrieben bekommen (Kienhues et al., 2020). Dies hilft wiederum dabei, nicht nur die Ursachen von Inkohärenz auf Quellen zurückzuführen, sondern diese Quellen und deren Inhalte zu gewichten. Entsprechend ist das Sourcing mit seinen vielfältigen Funktionen bei der Kohärenzwiederherstellung hier anzuführen (Philipp, 2023). Ergänzend kann Corroboration, also das dokumentenvergleichende Lesen, zum Zweck des verifikatorischen Abgleichs von Aussagen mit anderen vertrauenswürdigen Dokumenten dabei helfen, epistemisch unsichere Informationen und Kohärenzwidersprüche aufzulösen (Philipp, 2024a). In diesem Fall versuchen Personen, externe Quellen zu konsultieren, um sich ausreichend über intertextuelle Konvergenz zu vergewissern. Die Zuweisung von Plausibilität (strenggenommen also eine

Erste-Hand-Einschätzung) erfolgt dann auf der Basis vergleichender Zweite-Hand-Einschätzungen.

Die analytische Trennung von Erste- und Zweite-Hand-Einschätzungen ist eine kategoriale Trennung, die sich an der Frage scheidet, ob der Inhalt oder die Quelle beurteilt wird. Die Modelle aus Tabelle 20 – mit Ausnahme des „Discrepancy-Induced Source Comprehension Model“ – sehen dezidiert beide Formen der Einschätzungen vor. In dieser Modellarchitektur liegen die Differenzen zwischen den Modellen vor. Das „Discrepancy-Induced Source Comprehension Model“ postuliert, dass Widersprüche zwischen Dokumenten das Sourcing als Reaktion hervorrufen und dazu führen sollen, die kognitive Dissonanz zu reduzieren (Braasch & Kessler, 2021). Das „Two Step Model of Validation“ behandelt im Rahmen der epistemischen Elaboration das Sourcing als eine Möglichkeit unter mehreren, damit Personen balancierte Repräsentationen überzeugungsinkonsistenter Inhalte bzw. mehrperspektivischer Darstellungen erlangen und insbesondere jene Inhalte berücksichtigen, die ihren eigenen ontischen Überzeugungen zuwiderlaufen (Richter & Maier, 2017; Richter et al., 2020). Das Modell sieht für das Erreichen dieses Ziels verschiedene Prozesse als Mittel vor; ursprünglich waren Erste-Hand-Einschätzungen besonders deutliche Bestandteile. Das „Content-Source Integration Model“ modelliert anders als die anderen beiden Modelle zwei Phasen, bei denen die Plausibilisierung des Zustandekommens von kognitiver Dissonanz auf „soziale Gründe“ (Zweite-Hand-Einschätzung) und/oder „epistemische Gründe“ (Erste-Hand-Einschätzung) attribuiert wird, was in der Phase der Kohärenzwiederherstellung verortet wird (Braasch & Scharer, 2020). Dieser Phase, welche eine große Übereinstimmung mit der Phase der Dissonanzreduktion in der „Discrepancy-Induced Source Comprehension Model“ bzw. der epistemischen Elaboration im „Two Step Model of Validation“ aufweist, kann, muss aber nicht von einer weiteren Phase gefolgt werden. Diese Phase wird im „Content-Source Integration Model“ als „Evaluation“ (Braasch & Scharer, 2020) bzw. als „Konfliktauflösung“ (Stadtler & Bromme, 2014) bezeichnet und sieht eine entweder inhaltliche (also Erste-Hand-)Einschätzung oder eine die Herkunft betreffende (ergo Zweite-Hand-)Einschätzung vor. Diese Trennung von Erkennen und Verarbeiten von Ursachen von Kohärenzbrüchen auf der einen Seite und auf der anderen Seite dem Evaluieren der Gründe, um zu einer Entscheidung zur Glaub- bzw. Vertrauenswürdigkeit zu gelangen, spiegelt wider, dass es sich um Prozessverbünde handelt, die verschiedenen Zwecken dienen (s. dazu exemplarisch Deroover et al., 2022, und Philipp, 2023). Eine derartige Trennung ist analytisch sinnvoll, gleichzeitig werden diese Prozesse in verschiedenen Modellen als Prozessverbund bzw. als lesestrategische Optionen behandelt, darunter bereits im „Two Step Model of Validation“ und vor allem im „Content-Source Integration Model“. Erste- und Zweite-Hand-Einschätzungen

erfüllen damit verschiedene und komplementäre Funktionen, damit Leserinnen und Leser ausreichend kohärente mentale Repräsentationen entwickeln können (List, 2020).

Um diese komplexe Thematik eingangs zu verdeutlichen, enthält der Kasten „Illustrierendes Beispiel 18“ die Darstellung zweier Studien. Beide Studien eint, dass sie die Komplementarität beider Formen von Einschätzungen herausstellen. Zudem ist ein hervorhebenswertes Muster, dass Personen zu teils unterschiedlichen Ausprägungen neigen, mittels welcher Einschätzungen sie ihre Urteile begründen, wenn sie mit konfligierenden Aussagen konfrontiert werden.

Illustrierendes Beispiel 18

Wie und auf welcher Basis beurteilen Leserinnen und Leser die Plausibilität und die Glaubwürdigkeit, wenn sich Perspektiven unterscheiden und das Thema ihnen vertraut oder unvertraut ist?

A) Was war das Erkenntnisinteresse?

Die Fragestellungen der Studie aus Israel (Barzilai, Thomm & Shlomi-Elooz, 2020) waren vielfältig, zwei stechen heraus. Erstens interessierten sich die Forscherinnen dafür, welche Effekte die inhaltliche Vertrautheit darauf hat, für wie inhaltlich wahr Personen Aussagen halten und welche Glaubwürdigkeit sie den Quellen dieser Aussagen zuweisen. Zweitens ging es darum, die offenen Begründungen für die Urteile zu verstehen und Muster zu finden.

B) Wer wurde untersucht?

Die Publikation enthielt zwei separate Studien, die inhaltlich aufeinander aufbauten. An Studie 1 partizipierten 104 im Durchschnitt 25-Jährige (davon etwas mehr als zwei Drittel weiblichen Geschlechts). An Studie 2 nahmen weitere 113 Studierende (80 Prozent Frauen, im Mittel knapp 28 Jahre alt) derselben Universität teil. Gezielt wurden nur Studierende ausgewählt, die als Laienpersonen in der Domäne Geschichte galten; sie stammten aus diversen sozial- und geisteswissenschaftlichen Fächern, aber nicht aus der Geschichtswissenschaft.

C) Welche Instrumente kamen zum Einsatz?

Das Experiment sah vor, dass die Testpersonen entweder zu einer ihnen vertrauten Thematik (Jom-Kippur-Krieg) oder zu einem fingierten Thema (Fünfter Libanischer Krieg) einen Text lasen, der zwei sehr unterschiedliche Positionen enthielt, die von zwei eingebetteten Quellen postuliert wurden. Die beiden Dokumentensets waren auf sehr ähnliche Weise konstruiert, sodass lediglich die Vertrautheit bzw. Unvertrautheit das Unterscheidungsmerkmal bildete. Beide Experimente unterschieden sich in Details der Durchführung, wiesen aber die gleiche Grundstruktur auf.

Die Probandinnen und Probanden lasen die Dokumente und gaben dann für die Plausibilität jeder der diskrepanten Aussagen und für die Glaubwürdigkeit ein numerisches Urteil ab, zusätzlich beantworteten sie in einer offenen Frage mittels einer Begründung,

wieso sie zu dem Urteil gelangt waren. Bei der Beantwortung hatten sie Zugriff auf das jeweilige Dokumentenset. Die numerischen Plausibilitäts- bzw. Glaubwürdigkeitsurteile wurden varianzanalytisch ausgewertet, die offenen Antworten inhaltsanalytisch kodiert und über Chi-Quadrat-Tests verglichen. Insgesamt vier Kategorien wurden bei den offenen Antworten gebildet, die für die Plausibilitäts- und Glaubwürdigkeitsfragen aber die gleichen Bezugspunkte aufwiesen. Erstens wurden *vorwissensbasierte Begründungen* gezählt, in denen die Verfügbarkeit von Vorwissen bzw. die Übereinstimmung mit dem Vorwissen oder eigenen Überzeugungen die Grundlage bildeten. Zweitens wurden *diskursbasierte Begründungen* kodiert, in denen die Begründungen im Text anhand von Belegen, Details, Datenquellen, Fakten, Erklärungen, logischer Widerspruchsfreiheit sowie dem Schreibstil erfolgten. Als dritte Variante galten *vergleichende Begründungen (Corroboration)*, welche auf detektierten Gemeinsamkeiten und Unterschieden fußten. Zu guter Letzt wurden *sourcingbasierte Begründungen* ausgewertet, in denen Quellenmerkmale wie Fachwissen, Standpunkte und Voreingenommenheit beurteilt worden waren. Vorwissens-, diskurs- und vergleichsbasierte Begründungen wertete das Forschungsteam als Erste-Hand-Einschätzungen, nur das Sourcing wurde als Zweite-Hand-Einschätzung klassifiziert.

D) Was waren die Hauptbefunde?

Die Ergebnisse zeigen die Eminenz der thematischen Vertrautheit für die Urteile und deren Begründungen:

- *Vertrautheit führt zu stärker polarisierten Urteilen in puncto Plausibilität und Vertrauenswürdigkeit.* Bei unvertrauten Themen gab es in beiden Studien keine signifikanten Unterschiede in der Beurteilung der Plausibilität und der Glaubwürdigkeit. Bei den vertrauten Themen war dies deutlich anders gelagert: Beide Urteile fielen in beiden Studien deutlich positiver in Richtung der Aussagen und der Quellen aus, welche die vertrautere Position vertraten.
- *Bei der Plausibilitätsbeurteilung divergieren die Evaluationen teils je nach Kategorie in Abhängigkeit von der Vertrautheit.* War das Thema vertraut, so neigten Personen vor allem dazu, vorwissensbasierte Einschätzungen vorzunehmen. Bei unvertrauten Themen waren diskursbasierte Einschätzungen die häufigste Kategorie. Dieses Muster erwies sich in der zweiten Studie als noch prononcierter als in der ersten (s. Abbildung 15).
- *In der Einschätzung der Vertrauenswürdigkeit zeigten sich weniger Unterschiede im stark genutzten Sourcing als bei den Erste-Hand-Einschätzungen.* Analog zur Plausibilitätseinschätzung nutzten Personen bei vertrauten Themen mehr vorwissensbasierte Einschätzungen, während bei dem unvertrauten Thema die diskursbasierten und teils textvergleichenden Begründungen stärker ausgeprägt waren (s. Abbildung 15). Das Sourcing scheint damit unabhängig vom Vorwissen eine Option zu sein, um die Vertrauenswürdigkeit einzuschätzen.
- *Personen bemühen tatsächlich mehrere (kategorial ähnliche) Strategien und sie nutzen sie sowohl für Plausibilitäts- als auch für Vertrauenswürdigkeitsbegründungen.* Die (per Experimentaufbau explizit eingeforderten) Erste- und Zweite-Hand-Einschätzungen

sind demnach nicht einander ausschließende Vorgehensweisen. Vielmehr überschneiden sie sich und koexistieren. Außerdem erfüllen die gleichen Einschätzungen ihre Funktionen, um zu Urteilen über den berechtigten Geltungsanspruch zu gelangen. Dies ist bei aller Differenz in der Vertrautheit bzw. Unvertrautheit ein wichtiger Befund.

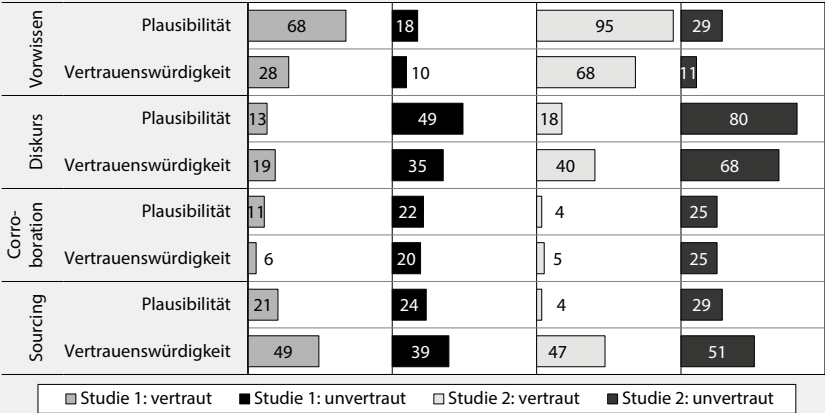


Abbildung 15: Anteile der Testpersonen, die in den beiden Studien Begründungen in vier Kategorien geäußert hatten (Mehrfachnennungen möglich, Angaben in Prozent, Quelle: eigene Darstellung auf der Basis von Barzilai, Thomm & Shlomi-Elouo, 2020, Tabellen C1 und C2; Legende: vertraut = vertrautes Thema, unvertraut = unvertrautes Thema)

E) Diskussion

Personen greifen je nach Bezugspunkt des Urteils und der inhaltlichen Vertrautheit auf verschiedene Muster bei Erste- und Zweite-Hand-Einschätzungen zurück. Bei inhaltlicher Vertrautheit fungieren innere Ressourcen dabei, beide Arten von Urteilen abzustützen, bei geringer Vertrautheit erfolgt ein Umweg über das prüfende Lesen (vor allem durch diskursbasierte Prüfungsgänge, teils über vergleichendes Lesen). Sourcing hingegen scheint weniger vorwissenssensitiv zu sein, zumindest bei den Vertrauenswürdigkeitsurteilen. Dieses wichtige Befundmuster aus zwei Studien unterstreicht, dass Erste- und Zweite-Hand-Einschätzungen dynamisch zu interagieren scheinen. Insbesondere die Erste-Hand-Einschätzungen wirken dort, wo sie an inhaltliches Vorwissen gebunden sind, als gangbarer Weg, während sie bei unvertrauteren Themen ein analytischeres Vorgehen einfordern.

4.2.2.2 Erste-Hand-Einschätzungen: Inhalte prüfen und gewichten (Produktperspektive)

Einschätzungen aus erster Hand zur Plausibilität bzw. zu anderen Facetten der inhaltlichen Beurteilungen (Lombardi et al., 2016), darunter auch der argumentativen Überzeugungskraft, betreffen Urteile über den Geltungsanspruch von Aussagen als vermutlich wahr oder unwahr. Diese Fähigkeit zum Validieren wird vor allem im Zusammenhang mit dem Umgang mit Fake News aktuell

diskutiert (Chinn et al., 2021; Sinatra & Lombardi, 2020). Die Beurteilung des potenziellen Wahrheitsgehaltes von Aussagen wird nicht nur in diesem Spezialfall des Umgangs mit epistemisch unsicheren und Faktizität nur vorgebenden Aussagen benötigt. Vielmehr gilt das Einschätzen von Plausibilität ebenfalls als zentral für das Verstehen allgemein (Singer, 2013). Die Annahme ist, dass Leserinnen und Leser automatisch beim Lesen neue Informationen mit anderen im Arbeitsgedächtnis verfügbaren Informationen abgleichen und zur Vermeidung von Inkohärenz plausible Informationen privilegieren (O'Brien & Cook, 2016a). Diesem Mechanismus wohnt eine Ambivalenz inne, da er im Optimalfall routinisierte Selektionen plausibler Informationen fördert. Allerdings kann im Negativfall bei einem ungeeigneten Bemessungsbenchmark kein ausreichendes Korrektiv vorhanden sein, objektiv zulässige, aber unplausible Informationen korrekt als plausibel einzuschätzen. In diesem Fall würden die Leserinnen und Leser nicht nur einseitige, sondern sogar unzulässige Repräsentationen entwickeln (Chinn et al., 2014).

Die Zuweisung von Plausibilität bei sich widersprechenden Aussagen wurde als ein Element der Multiple-Documents-Literacy verschiedentlich aufgenommen. Neben dem „Content-Source Integration Model“ und den Erste-Hand-Einschätzungen (Stadtler & Bromme, 2014) hat der Umgang mit intertextuellem Dissens besondere Betonung im „Two Step Model of Validation“ erfahren (Richter & Maier, 2017). Letztgenanntes Modell widmet sich der Frage, wie thematische Überzeugungen ausbalancierte Repräsentationen unterminieren, wenn der Validierungsmechanismus nicht gezielt überwunden wird (s. dazu das Teilkap. 4.3.2). In Rahmenmodellen wie jenem von List und Alexander (2019) wird die Fähigkeit zu inhaltlich begründeten Einschätzungen von intertextuellen Konflikten am oberen Ende des Fähigkeitspektrums verortet. Das strategische Validieren wirkt damit wichtig, allerdings zugleich kognitiv anspruchsvoll.

Gemessen an den Forschungsbemühungen zu den Glaubwürdigkeitseinschätzungen aus zweiter Hand für eine indirekte Plausibilitätsbeurteilung ist die Empirie zu den Erste-Hand-Einschätzungen im Umgang mit multiplen Dokumenten deutlich randständiger. Aktuell zeigt sich dies besonders prononciert in einem Forschungsüberblick von List und Sun (2023) mit 57 sekundäranalysierten Studien. Die Beurteilung des Inhalts wurde dort in 11 Studien bei den Prozessen und in 14 Studien bei den Produkten kodiert. Es deutet sich darin eine gewisse Marginalisierung des Forschungsgegenstands an, welche dadurch noch an Brisanz gewinnt, dass die Nutzung von direkten Plausibilitätsurteilen für Verstehensleistungen wiederum nur einen Bruchteil der Studien ausmacht. Die folgende Darstellung des limitierten Forschungsstands fällt deshalb knapp und produktorientiert aus. Strukturiert wird sie entlang zweier Fokusse, nämlich zum einen dem Zusammenhang von primär inhaltsbezogener Kritik und Verstehensleistungen (Fokus 1) und den inhaltlichen Erklärungen von Dissens (Fokus 2). Beide Fokus-Darstellungen bzw. die ihnen zugrundeliegende Forschungsstände

sind aktuell noch zu unsystematisch, als dass sie ein kohärentes Gesamtbild erlauben.

Fokus 1: Inhaltliche Kritik als spezifische Form der Plausibilitätseinschätzung

Das kritische Lesen lässt sich als eine Form des Validierens werten, da es eine lesestrategische, meist mehrdimensionale Einschätzung des Geltungsanspruchs von Aussagen betrifft und sich auf den Inhalt bezieht (List, in Druck; List & Oaxaca, 2024). Einige Studien haben sich mit dem kritischen Denken befasst (List & Sun, 2023), ihre Ergebnisse werden an dieser Stelle gebündelt.

Mehrere Studien mit Studierenden ergaben ein eindeutiges Bild, wenngleich diese empirische Basis aktuell noch schmal ist. Personen, die nach der Lektüre von Studien zu einem kontroversen Thema mehr berechnigte Kritikpunkte an den Texten zu formulieren imstande sind, können die Kontroverse ausgewogener in Argumentationen darstellen (List, in Druck, Studien 1 und 2; List & Oaxaca, 2024). Das bedeutet, dass ein kritisches Lesen zu einer vollständiger dargestellten Kontroverse führt und damit dem Inhalt gerechter wird.

Für berechnigte Kritik ist es aus empirischer Sicht besonders wichtig, dass die Personen in ihrem eigenen Verständnis des Konzepts „Kritik“ das Untersuchen von Stärken und Schwächen angeführt und bei retrospektiven Schilderungen des Vorgehens dies berichtet hatten (List, in Druck, Studie 2). Außerdem ist für das Evaluieren eine Passung von dem Wissen über die Belege in Aussagen sowie die dazu passenden Kriterien nötig, um die Evaluationen zu begründen, z. B. die Nutzung von Methoden für die Einschätzung von korrelativen oder kausalen Zusammenhängen (List et al., 2022). Entsprechend illustrieren Studien, dass Personen mit höherem Fachwissen sehr viel stärker zu mehr Evaluationen des Inhalts und seiner Plausibilität neigen: Leistungsstärkere Personen evaluieren beispielsweise den Inhalt von Websites häufiger hinsichtlich der Qualität der Informationen (Goldman et al., 2012) bzw. prüfen sie bei Texten stärker die Schlüssigkeit der Argumentation (von der Mühlen et al., 2015, 2016).

Zusammengefasst: Die Fähigkeit, Inhalte aus multiplen Dokumenten inhaltlich angemessen zu kritisieren, korrespondiert mit inhaltlich vollständigeren und stärker mit Inferenzen angereicherten Repräsentationen. Im Gegenzug hängen einseitigere Repräsentationen mit einem weniger kritischen Lesen zusammen. Diese Befunde sind zwar ermutigend, jedoch haben die Studien mit ihrer bislang sehr geringen Anzahl von zu verarbeitenden Dokumenten wenig Aussagekraft, wenn es darum geht, die Gewichtung von Inhalten zu beleuchten und zu plausibilisieren. Darum können die bisherigen Studienergebnisse als erste punktuelle Hinweise betrachtet werden, die zugleich markieren, dass dieser Bereich mehr Grundlagenforschung verdient.

Die Erklärung von Dissens wurde in einem eigens dafür kreierten Instrument fokussiert, dem „Explaining Conflicting Scientific Claims (ECSC) Questionnaire“ (Thomm et al., 2015). Dieser erfasst vier Facetten, wie Konflikte erklärt werden können. Zwei davon betreffen epistemische Gründe, nämlich zum einen die divergierenden Methoden in der Herstellung von Erkenntnissen und zum anderen die Komplexität des Gegenstands. Die zwei anderen beiden Facetten offerieren metadatenbezogene Evaluationen, indem auf der einen Seite die Aufrichtigkeit der Quelle im Sinne ihrer Redlichkeit bzw. Integrität ein Erklärungsmuster bildet oder auf der anderen Seite die variierende Kompetenz (Expertise) als Ursache genutzt werden kann. Die ersten beiden Facetten – Methoden und Komplexität – sind konzeptuell näher an den Plausibilitätsbeurteilungen als die anderen beiden und werden darum hier fokussiert.

Auffällig ist, dass inhaltlich divergierende Aussagen selten auf die Methoden bzw. die Komplexität attribuiert werden, da sich in Experimenten mit variierender Glaubwürdigkeit bzw. Expertise keine signifikanten Unterschiede nachweisen lassen (Gottschling & Kammerer, 2021; Gottschling et al., 2020; Thomm & Bromme, 2016; Thomm et al., 2015, Studie 2). Vereinzelt gab es Differenzen: So wurde in einer Studie (Gottschling et al., 2019) bei unterschiedlich vertrauenswürdigen Quellen mehr Zustimmung zur Komplexität als Grund für divergierende Aussagen geäußert, nicht aber für die Forschungsmethodik. Das in den zitierten Studien überwiegend vorgefundene Null-Ergebnis bei der Konflikterklärung spiegelt – gemessen an den komplementären Befunden, dass die Testpersonen die Konflikte vor allem auf jene metadatenbezogenen Merkmale zurückführten, welche eigens experimentell variiert worden waren (s. dazu Teilkap. 4.2.2.3 und dort im Abschnitt „Phase 2“ den Teil „Fokus 3: Konflikterklärung über die Interpretation von Metadaten“) – möglicherweise nur einen sehr limitierten Ausschnitt wider. Das bedeutet im Umkehrschluss, dass das Erklären von Dissens bezogen auf die Inhalte und deren Merkmale ohne Zusatzinformationen wenig wahrscheinlich ist bzw. empirisch mit den meist kurzen Texten nicht sinnvoll möglich wirkt, zumal in Wissensgebieten, in welchen die Laienpersonen wenig Wissen aufweisen. Hierzu passt der Befund, dass Personen die Plausibilität von widersprüchlichen Aussagen unterschiedlich beurteilen, wenn ihnen die Themen vertraut sind (Barzilai, Thomm & Shlomi-Elooz, 2020, Studien 1 und 2): Sie differenzieren dann stärker zwischen plausiblen und unplausiblen Aussagen, während bei unvertrauten Themen die Unterschiede in den Urteilen nur gering ausfallen.

Die Auflösung von intertextuellen Konflikten anhand der Inhalte und ihrer Beurteilung ist damit ein Forschungsgebiet, welches aktuell nur fragmentarische Erkenntnisse bezogen auf die Multiple-Documents-Literacy offeriert. Insbesondere mangelt es an Studien, die gezielt untersuchen und experimentell variieren, wie die Plausibilitätseinschätzungen kontextsensitiv erfolgen und sich in den

kognitiv anspruchsvollen Verstehensleistungen kohärenter mentaler Modelle niederschlagen. Insbesondere Prozessstudien bilden ein Desiderat.

4.2.2.3 Zweite-Hand-Einschätzungen: Metadaten nutzen und beurteilen

Die Nutzung der Metadaten ist eine Variante, intertextuelle Konflikte aufzulösen, wenn das eigene Vorwissen nicht ausreicht bzw. wenn die Unsicherheit zu groß wirkt und Personen darum über die Zuweisung von Vertrauenswürdigkeit indirekt zu Plausibilitätseinschätzungen gelangen. Wie dies erfolgt, steht aktuell im Zentrum der Modelle zum Verstehen multipler konfligierender Dokumente und wird besonders im „Discrepancy-Induced Source Comprehension Model“ (Braasch & Bråten, 2017; Braasch & Kessler, 2021) sowie im „Content-Source Integration Model“ (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014) behandelt. Beide Modelle eint die Prämisse, dass intertextuelle Konflikte zu einer kognitiven Dissonanz führen, die über das Sourcing auflösbar und auflösungsbedürftig ist. Der Geltungsbereich der Modelle jedoch variiert. Beide Modelle gehen davon aus, dass Sourcing eine Reaktion auf die kognitive Dissonanz wegen eines unterschrittenen Kohärenzschwellenwerts ist. Das „Discrepancy-Induced Source Comprehension Model“ postuliert, dass lesestrategisches Sourcing stattfindet. Das „Content-Source Integration Model“ spezifiziert, wozu das Sourcing genutzt wird, indem es mehrere Phasen modelliert. Darum lässt sich das zweitgenannte Modell als eine Präzisierung des erstgenannten beschreiben.

Diese Spezifikation des Sourcings im „Content-Source Integration Model“, auf einer von zwei Routen Kohärenz produktiv zu stiften, ist für die folgende Darstellung leitend, indem zwei Phasen – nicht in der Diktion und Logik der Phasenlogiken beider Modelle, da diese als erste Phase und damit als Bedingung die Konfliktdetektion an den Anfang stellen – dargestellt werden. Die erste Phase in der folgenden Darstellung widmet sich den Befunden zur Nutzung von Metadaten als kohärenzstiftendem Prozess, während die zweite Phase auf die evaluativen Leistungen beim Sourcing eingeht.

Phase 1: Metadaten erkennen und zum Zweck der Kohärenzstiftung mit konfligierenden Inhalten kombinieren

In diesem Abschnitt werden die Produkt- und Prozessperspektiven behandelt, wie Personen Metadaten konsultieren und zunächst noch ohne evaluative Bestandteile nutzen. Die Darstellung beginnt mit der Produktperspektive, es folgt die Zuwendung zu den Prozessen.

Die Produktmaße aus den bisher vorgelegten Studien behandeln unterschiedliche Aspekte des konfliktinduzierten Sourcings und der Repräsentation von Metadaten bzw. Dokumenten als Ursprung resp. Träger von Inhalten. Ein erster, besser erforschter Bereich behandelt, inwieweit konfligierende Dokumente zu mehr korrekten Quelle-Inhalt-Verbindungen bzw. zu mehr Quellennennungen in offenen Aufgaben führen. Der zweite Bereich behandelt den Zusammenhang von Konfliktverifikationen und Quelle-Inhalt-Verbindungen. Der dritte Bereich der Forschung geht den Erinnerungsleistungen an Metadaten nach, die bei konfligierenden Sets von Dokumenten untersucht wurden.

- *Kohärenz(wieder)herstellung unter Zuhilfenahme von Metadaten ist bei konfligierenden Dokumentensets häufiger feststellbar als bei konvergierenden.* Dieses Muster zeigte sich in zwei Arten von Tests in Experimenten, in denen Gruppen konvergierende oder konfligierende Dokumente lasen. Erstens konnten Personen der Experimentalkondition mit konfligierenden Dokumenten in *geschlossenen Aufgabenformaten* besser Quelle-Inhalt-Verbindungen korrekt an- bzw. wiedergeben (Barzilai & Eshet-Alkalai, 2015, in Interaktion mit epistemischen Überzeugungen; Braasch et al., 2012, Experimente 1 und 2; Braasch et al., 2016, Experimente 1–3; Rouet, Saux et al., 2021, Experiment 2; Saux et al., 2017, Experimente 1 und 2; Saux et al., 2018, Experimente 1 und 2; Saux, Vibert et al., 2021, Experimente 1 und 2) und machten zudem weniger Fehler in solchen Aufgaben (Braasch et al., 2016, Experimente 1–3; Saux et al., 2017, Experiment 1). Zweitens bestand das Muster ebenso in *offenen Aufgaben*: Konfligierende Dokumente führten zu mehr Referenzierungen, also angeführten Quelleninformationen (Braasch et al., 2012, Experimente 1 und 2; Kammerer & Gerjets, 2012b, 2014a; Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2; Rouet et al., 2016, Experimente 1–3; Schoor et al., 2023), – allerdings ist dieser Effekt nicht durchgängig belegt (Braasch et al., 2016, Experiment 3; List, Du & Lee, 2021a; Londra & Saux, 2023, Experiment 1). Relativierend muss zudem angeführt werden, dass Referenzierungen nicht mit einer kritischen Verwendung von Inhalten aus den Dokumenten gleichzusetzen sind (Bräten et al., 2015). Damit sind vor allem die Befunde von geschlossenen Tests auf der Basis meist sehr kurzer Texte (teils Satzfolgen) eindeutig: Dokumentensets mit Widersprüchen führen zu korrekteren Quelle-Inhalt-Verbindungen. Metadaten und Aussagen werden demnach kognitiv stärker koaktiviert und kombiniert repräsentiert.
- *Je besser Personen dazu in der Lage sind, Konflikte korrekt als solche zu erkennen, desto besser repräsentieren sie Quelle-Inhalt-Verbindungen.* In einigen Studien wurden die Fähigkeiten, Konflikte mittels Verifikationsaufgaben zu

bestimmen (s. Teilkap. [4.2.1.1](#) und dort Nr. 2a), und die Fähigkeit, die Inhalte von Dokumenten ihrem Ursprung korrekt zuzuordnen (Quelle-Inhalt-Verbindungen, s. Teilkap. [4.2.1.2](#) und dort Nr. 2), erfasst. Beide Erinnerungsleistungen hängen zusammen: Wer Konflikte besser erkannt hatte, konnte auch besser gesamthaft die Dokumente als Ursprung von Aussagen memorieren (Delgado et al., 2020; Stang Lund et al., 2017, 2019); regressionsanalytisch erwies sich die Konfliktverifikation überdies als Prädiktor der Quelle-Inhalt-Verbindungen über andere Merkmale hinaus (Stang Lund et al., 2017). Dieser Zusammenhang betrifft nicht die spezifisch die Kohärenzprobleme generierenden Konflikte und die Nutzung der spezifischen Metadaten, welche den konfligierenden Aussagen anheften. Anders gesagt: Der Zusammenhang von korrekter Konfliktdetektion und den spezifischen Quellen der konfligierenden Aussagen ist noch zu wenig erforscht. Insofern ist damit ein generischer Zusammenhang zwischen zwei Memorierungsleistungen angesprochen, dass intertextuelle Konflikte zu reichhaltigeren und differenzierten mentalen Repräsentationen mit mehr metadatenbasierten Bestandteilen führen (Bråten & Braasch, 2018).

- *Konfligierende Aussagen zwischen Dokumentensets ziehen selbst dann keine konsistent besseren Erinnerungsleistungen nach sich, wenn Metadaten auf kontrastive Ausprägungen der Glaubwürdigkeit schließen lassen.* Einige Experimente widmeten sich der Frage, ob verschiedene Ausprägungen der Metadaten (z. B. hohe vs. geringe Expertise oder hohe bzw. geringe Vertrauenswürdigkeit) dazu führen, dass Personen sich besser an die Metadaten erinnern. In einem Teil der Studien gab es weder bei rein offenen noch bei mit Hinweisen versehenen Aufgaben bessere Erinnerungsleistungen an Metadaten (Gottschling & Kammerer, 2021; Gottschling et al., 2019; Rouet, Saux et al., 2021, Experiment 2). Das gleiche Null-Resultat bestand bei Multiple-Choice-Aufgaben, in denen die Aussagen mit den Quellen zu verbinden waren (Gottschling et al., 2020). Diese Ergebnisse werden jedoch von anderen Studien mit genaueren Analysen spezifiziert. Beispielsweise erinnern sich Personen besser an jene Quellen, die eine höhere Expertise aufweisen (Thomm & Bromme, 2016) bzw. denen eine größere Aufrichtigkeit zugestanden werden kann (Scharrer et al., 2019, Studie 1; Thomm & Bromme, 2016; Thomm et al., 2015, Studie 2), wobei auch dies nicht durchgängig belegt wurde (Scharrer et al., 2019, Studie 2).

Zusammengefasst: (Erkannte) Konflikte zwischen Dokumenten bzw. Aussagen aus Dokumenten führen zu vermehrt repräsentierten Quelle-Inhalt-Verbindungen in rezeptiven Tests sowie mehr Quellenangaben bei expressiven Aufgaben. Zugleich erinnern sich Personen jedoch nicht durchgängig besser an die Metadaten, wenn innerhalb der konfligierenden Dokumentensets gezielt Merkmalsausprägungen experimentell manipuliert wurden. Die Produktmaße sprechen also

insgesamt dafür, dass intertextuelle Konflikte zu einem Mehr an sourcingbasierten Repräsentationen führen können, dass dies aber nicht für alle Arten der Sourcing-Maße gilt.

Prozesse

Wie schon bei den Produktdaten im Abschnitt zuvor gibt es über den Vergleich von Experimentalkonditionen mit konsistenten und konfligierenden Dokumentensets die Möglichkeit, Unterschiede zu eruieren. Diese Prozessdaten sind seltener als Produktdaten erhoben worden und zeigen vor allem bei den spezifischen Prozessen Auffälligkeiten:

- An den reinen *Lesezeiten* allein ließen sich keine Differenzen festmachen, denn Personen lasen kongruente Texte länger als konfligierende (Braasch et al., 2016, Experiment 3) bzw. lasen sie ähnlich lang (Saux et al., 2017, Experiment 1; Schoor et al., 2023).
- Allerdings zeigte sich, dass selbst bei vergleichbarer Netto-Lesezeit ein Indikator für andere Prozesse auffiel. Denn Personen aus der Experimentalgruppe mit konfligierenden Dokumenten wechselten beim Lesen häufiger zwischen diesen Dokumenten als die Personen mit konvergierenden Dokumenten. Das gilt als Indikator für *Corroboration*-Prozesse, also für ein textvergleichendes Lesen (Schoor et al., 2023).
- Bezogen auf das *Sourcing* waren relativ konsistente Ergebnisse studienübergreifend extrahierbar. Personen, die konfligierende Dokumente lasen, wählten erstens das Impressum als Ort für Quelleninformationen auf Websites häufiger an (Kammerer, Kalbfell & Gerjets, 2016, Experiment 2). Sie betrachteten zweitens (visuelle bzw. im Impressum anwählbare) Quelleninformationen häufiger und länger (Braasch et al., 2012, Experiment 1; Gottschling & Kammerer, 2021; Gottschling et al., 2019; Kammerer & Gerjets, 2014a; Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2; Saux, Vibert et al., 2021, Experiment 2) und neigten stärker zu evaluativen Äußerungen (Kammerer, Kalbfell & Gerjets, 2016, Experiment 1). Außerdem – und das ist ein eher deskriptiver Befund – absolvierten sie drittens in Dokumentensets mit konvergierenden und konfligierenden Dokumenten mehr Sourcingaktivitäten bei jenen Dokumenten, die sich inhaltlich am stärksten widersprachen (Strømsø et al., 2013).

Damit zeigt sich vor allem bei den Prozessindikatoren, dass Widersprüche von Aussagen in Dokumenten zu einem stärker textvergleichenden Lesen und zu einer stärkeren Zuwendung zu Metadaten führen, primär zur Herkunft von Aussagen. Dies kann als nötige, aber nicht hinreichende Bedingung für die Kohärenzwiederherstellung durch die Nutzung betrachtet werden, die im folgenden Abschnitt behandelt wird.

Phase 2: Metadaten evaluieren und auf dieser Grundlage Inhalte beurteilen, um Kohärenz zu stiften (Produktperspektive)

Das Erkennen und Repräsentieren von Metadaten ist nur eine Facette dessen, was Sourcing im Allgemeinen (Scharrer & Salmerón, 2016) bzw. beim Auflösen von intertextuellen Konflikten im Besonderen leistet (Braasch & Scharrer, 2020, s. Teilkap. [3.1.2.3](#) und [4.1.2.3](#)). Intertextuelle Konflikte auf die Quellen zurückzuführen („Die Diskrepanz kommt deshalb zustande, weil sich die beiden Dokumente widersprechen“), ist für Leserinnen und Leser aber lediglich eine Option. Verlangen die Leseaufträge eine Positionierung, etwa weil eine Argumentation zu schreiben ist oder weil das Lesen einer Entscheidung für oder gegen eine Option dient, bedarf es einer begründbaren Einschätzung („Die Dokumente widersprechen sich, aber das Dokument A ist von einer Expertin und darum insgesamt belastbarer“). Hier helfen neben den Erste-Hand-Einschätzungen zur Beurteilung der Plausibilität als Supplement metadatenbezogene Evaluationen.

Diese Zweite-Hand-Einschätzungen sind ein heterogenes, primär hinsichtlich der Produktmaße untersuchtes Forschungsfeld mit einer Vielzahl von Variablen und Zugängen. Grob lassen sich zwei Zugänge in den Studien unterscheiden, nämlich solche mit mittelbaren und unmittelbaren Maßen. Mittelbar sind die Variablen, die auf unterschiedliche Verwendungen des Sourcings in Produktmaßen abzielen, unmittelbare Operationalisierungen bestehen aus Indikatoren, welche spezifisch die theoretisch postulierte Nähe von intertextuellen Konflikten und spezifischer Nutzung der Metadaten erfassen. Mittelbar wären demnach allgemeine Sourcing-Maße, unmittelbar hingegen spezifische Indikatoren, welche direkt auf evaluative bzw. elaborative Leistungen hinweisen.

Der folgende Forschungsüberblick beschränkt sich notgedrungen auf Produktmaße, was bereits das Desiderat bei Prozessmaßen, insbesondere bei den Beurteilungsprozessen, markiert. Die Darstellung erfolgt in einem Dreischritt, indem jeweils ein spezifischer Fokus gesetzt wird. Den Auftakt bilden zunächst die diversen Beurteilungsleistungen, deren Bezugspunkte variieren. Es folgt die Zuwendung zur Verwendung von Metadaten für die Auflösung von Konflikten. Zum Abschluss werden die Muster bei der interpretatorischen Auflösung der Konfliktursachen fokussiert.

Fokus 1: Beurteilungsleistungen bei kontroversen Aussagen

Die Auflösung von inhaltlichen Konflikten über die evaluativen Anteile des Sourcings ist ein mehrerer Einzelprozesse umfassender Verbund, dessen Bezugspunkte variieren. Dies zeigt sich unter anderem darin, dass die Studien die Beurteilung von Quellen ebenso erfassen wie die Zustimmung zu Inhalten in Abhängigkeit der Glaubwürdigkeitszuweisung. Hinzu kommen Studien, die sich

aus dem Paradigma der Metakognition kommend der Frage widmeten, wie sicher sich Person in ihren Inhaltsevaluationen waren. Diese drei Facetten gestalten sich empirisch folgendermaßen:

- *Beurteilungen der Quellen von Aussagen:* Die Beurteilung der Quelle erfolgt empirisch überwiegend mittels unmittelbarer, also direkter Maße, bei denen die jeweilige Quelle anhand mehrerer oder einzelner Teildimensionen zu beurteilen ist. Wenn in Studien Expertise bzw. die Aufrichtigkeit über die Metadaten systematisch variiert wurde, beurteilen Personen bei Differenzen zwischen den Quellen deren Glaubwürdigkeit (hier: Aufrichtigkeit und Ehrlichkeit) und Expertise unterschiedlich. Personen mit größerer *Expertise* wird eine höhere Kompetenz (Gottschling et al., 2020; Hendriks et al., 2015, Studie 3; Lescarret, Magnier et al., 2023; Thon & Jucks, 2017) bzw. eine allgemein höhere Glaubwürdigkeit (Lescarret, Magnier et al., 2023; Wertgen & Richter, 2020, Experiment 1 und Normstudie bei Experiment 2) attestiert. Ähnliches gilt für die Beurteilungen der Glaubwürdigkeit bei *aufrichtiger wirkenden Quellen* (Gottschling & Kammerer, 2021; Gottschling et al., 2019, 2020; Hendriks et al., 2015, Studie 3; Kammerer & Gerjets, 2012b, 2014a; Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2; Londra & Saux, 2023, Experimente 1 und 2; Thomm & Bromme, 2016). Hinzu kommt, dass Leserinnen und Lesern den Quellen mit höherer Glaubwürdigkeit eine höhere Überzeugungskraft bescheinigen (Kammerer, Kalbfell & Gerjets, 2016, Experimente 1 und 2). Eine andere Form von Glaubwürdigkeitsoperationalisierung wurde in den beiden Experimenten von Wertgen und Richter (2023) gewählt: Hier wurde Glaubwürdigkeit als Passung von Quelle und Aussage operationalisiert. Unglaubwürdig war beispielsweise ein fingierter Post des „World Wide Fund for Nature“, der die Bedeutung des Klimaschutzes herunterspielte. In beiden Experimenten konnten Personen dergestalt unglaubwürdige von glaubwürdigen Quellen mittels unterschiedlicher Plausibilitätsurteile unterscheiden.
- *Inhaltliche Zustimmung zu Positionen:* Nicht nur die Quellen als solche, sondern auch ihre inhaltliche Perspektive wird in Studien hinsichtlich des Urteils der lesenden Personen erforscht. Das Ergebnis ist dabei eindeutig: Laienpersonen stimmen bei intertextuellen Konflikten jenen Aussagen stärker zu, deren Quellen sie eine höhere Vertrauenswürdigkeit attestieren. Diese Vertrauenswürdigkeit wurde über verschiedene Facetten operationalisiert. Das Muster aus Experimenten ist (mit Ausnahmen: Thomm & Bromme, 2016) einhellig: Personen stimmen Aussagen eher zu, die von einer aufrichtigeren Quelle (Gottschling & Kammerer, 2021; Gottschling et al., 2019, 2020; Scharrer et al., 2019, Studien 1 und 2) bzw. einer Quelle mit höherer Expertise stammen (Gottschling et al., 2020; Thon & Jucks, 2017). Dieses Ergebnis ist deshalb zu betonen, weil es zu einem großen Teil aus den Studien stammte, in denen auch die Beurteilung der Quellen erhoben wurde. Die Ergebnisse

korrespondieren miteinander: Quellen mit höherer Glaubwürdigkeit stimmen Personen stärker zu, wenn sie auf Konflikte stoßen, bei denen sie wenig Vorwissen haben.

- *Urteilssicherheit bei der inhaltlichen Einschätzung:* Zusätzlich wurde in einigen Studien neben der kognitiven Einschätzung auch das metakognitive Urteil erhoben, wie sicher sich die Personen in ihrer inhaltlichen Einschätzung waren. Bei Aussagen aus vertrauenswürdiger wirkenden Quellen ist die subjektive Urteilssicherheit größer (Scharrer et al., 2019, Studien 1 und 2; Thon & Jucks, 2017).

Bei experimentell bewusst salient gestalteten Metadaten, die im direkten Vergleich auf Unterschiede der Quellen von Aussagen schließen lassen, gelingt den (erwachsenen) Testpersonen also dreierlei:

- *sinnvolle Attributionen des Dissens' auf Quellenmerkmale:* Leserinnen und Leser nutzen Differenzen in den Metadaten als Anlass für angemessene Urteile in geschlossenen Antwortformaten;
- *Kongruenz von Glaubwürdigkeitsurteil und inhaltlicher Zustimmung:* Als glaubwürdiger geltende Quellen erhalten höhere Zustimmungsurteile in puncto Inhalt;
- *inhaltliche Einschätzungen und das eigene Zutrauen in das eigene Urteil:* Die inhaltliche Einschätzung und das Vertrauen in diese Einschätzungen sind bei vertrauenswürdigen Quellen höher.

Mit forcierten Entscheidungen der lesenden Personen in experimentellen Settings lässt sich ein konsistentes studienübergreifendes Muster abstrahieren: Leserinnen und Leser können glaubwürdigere Quellen differenzierend erkennen, denen sie inhaltlich stärker zustimmen.

Fokus 2: spezifische Nutzung der Metadaten im Zusammenhang mit den intertextuellen Konflikten

Der zweite Fokus der Forschung liefert angesichts der noch deutlich ausbaufähigen Basis an Studien weder ein kohärentes noch stark belastbares Fundament für generalisierende Aussagen. Die fragmentarisch zu nennenden Resultate lauten:

- *Die mittelbar operationalisierten Zusammenhänge von expressiven Produktmaßen im Zusammenhang mit intertextuellen Inferenzen zu Konflikten und Metadaten wirken unerforscht und unklar.* Ein indirektes Maß für das Herstellen von Kohärenz mittels Metadaten stammt aus einer ersten Studie, in der die Personen zu inhaltlich konfligierenden Themen Diskurssynthesen schrieben, bei denen die Zieltextsorte variiert wurde (Stadtler et al., 2014).

Dabei wurden Zusammenfassungen und Argumentationen spezifisch fokussiert und eine Differenz mit Rekurs auf das Dokumentenmodell angenommen (s. Teilkap. 4.1.2.3 und 4.4.1.1): Zusammenfassungen erfordern eher die Bildung von integrierten mentalen Modellen (bzw. anderen (Vor-)Formen: Mischmodell bzw. separate Repräsentationen ohne Zitationen). Demgegenüber ist ein vollständiges Dokumentenmodell mit Dokumentenknoten auf der Intertextmodell-Ebene eher bei Argumentationen annehmbar. Beide Zieltextsorten elizitieren mithin mutmaßlich variierende Integrationsleistungen mit höheren Anforderungen bei den Inferenzleistungen in argumentativen Aufgaben (intertextuelle Konfliktbenennung in expressiven Aufgaben zum einen und Quelle-Inhalt-Verbindungen über Referenzierungen zum anderen). Als Indikatoren für die Leistungen wurden zwei unverbundene Maße genutzt: Zweiseitigkeit der Darstellung der Konfliktthemen sowie die Anzahl der Referenzierungen. Die Ergebnisse der Studie zeigen, dass in der Gruppe der argumentierenden Probandinnen und Probanden mehr zweiseitige Konflikte berichtet wurden und mehr Referenzierungen in den Texten enthalten waren. Die Kombination beider Maße wurde nicht analysiert. Eine weitere Studie (Kobayashi, 2014), in welcher die Positionierungen für die intertextuellen Konflikte in Diskurssynthesen ausgewertet und Sourcingmaße als mögliche Prädiktoren genutzt wurden, ergab zweierlei. Erstens kamen explizite Nennungen von Metadaten zu selten vor, um sie als eigene Prädiktoren zu verwenden. Zweitens waren andere Metadaten, die allgemein erfasst worden waren, in Regressionsanalysen nicht prädiktiv.

- *Bei den unmittelbar erfassten Metadaten-Inhalt-Verbindungen, die sensibler wirken für Zweite-Hand-Einschätzungen bei den Begründungen, ergeben sich theoriekonforme Zusammenhänge:* Die Nutzung von Metadaten für Positionierungen bei kontroversen Themen über inhaltlich gehaltvolle Metadaten ist nicht in allen Studien vorfindbar (Scharrer et al., 2019, Studie 1). Erst bei sehr deutlich erkennbaren Differenzen in den Metadaten nutzen Personen diese Metadaten verstärkt für eine inhaltliche Positionierung (Scharrer et al., 2019, Studie 2) bzw. verwenden in offenen Antworten mehr Informationen von vertrauenswürdigen Websites (Kammerer & Gerjets, 2014a). Es liegen außerdem Hinweise darauf vor, dass Personen bei kontroversen Themen jene Dokumente stärker zitieren, die sie für vertrauenswürdiger gehalten haben und bei denen sie metadatenbezogene Gründe wie die Dokumentenart oder die Autorinnen und Autoren angegeben haben (Bråten et al., 2015). Andere Hinweise stammen aus Experimentaldaten, welche aus einer leicht geänderten Perspektive ausgewertet wurden, nämlich indem ausgehend von den Metadaten geprüft wurde, ob intertextuelle Konflikte thematisiert werden (Stadtler, Scharrer, Brummernhenrich & Bromme, 2013). Demnach berichten Personen inhaltliche Konflikte häufiger vollständig mit zwei divergierenden Positionen, wenn die Metadaten darauf hinweisen, dass zwei Personen mit

höherer Expertise konfligierende Ansichten vertreten. Dieser Effekt trat bei Dokumenten mit Quellen geringerer Expertise nicht auf. Konflikte werden demnach dann berichtenswerter, wenn der inhaltliche Dissens mit höherer Vertrauenswürdigkeit der Dissensursachen gekoppelt ist.

Woran es bei der spezifischen Nutzung der Metadaten im Zusammenhang mit den intertextuellen Konflikten derzeit mangelt, sind vor allem Studien, welche untersuchen, in welchem Ausmaß und wie konsistent Personen tatsächlich auf die spezifischen Metadaten zurückgreifen, um bei sehr konkreten und lokalen intertextuellen Konflikten im Sinne makropropositionaler Divergenzen die Kohärenz über die Explikation der Quellen als Konflikten zu realisieren.

Fokus 3: Konflikterklärung über die Interpretation von Metadaten

Der dritte und letzte Fokus verfolgt das Erkenntnisinteresse, auf welcher Basis Leserinnen und Leser Metadaten dazu verwenden, das Zustandekommen von inhaltlichen Konflikten in wissenschaftlichen Kontexten zu erklären. Kennzeichnendes Merkmal der hier versammelten Studien ist, dass in ihnen die Konfliktursachen mehrdimensional über geschlossene Skalen erfasst wurden. Neben epistemischen Kriterien wurden auch sourcingbezogene Dimensionen erfasst. Die Ergebnislinie verläuft deutlich durch die Studien: *Personen weisen eine Sensitivität auf, welches Metadatum sie für ihre Erklärung von intertextuellem Dissens nutzen.* Das illustriert eine Reihe von Experimenten, in denen Aussagen in sehr kurzen Texten systematisch in ihren Metadaten variiert werden. So wird die wissenschaftliche Expertise über die Berufserfahrung (kurz oder lang) oder Aussagen zum Wissensstand der Quelle zum jeweiligen Gegenstand operationalisiert. Aufrichtigkeit lässt sich über die Zugehörigkeit zu Institutionen der öffentlichen Hand oder private Unternehmen variieren.

- Die Ergebnisse der Studien verweisen darauf, dass bei unterschiedlicher *Expertise* der Quellen das Zustandekommen über Kompetenzdifferenzen von Personen in der Wissenschaft erklärt wird (Gottschling et al., 2020; Stadtler, Scharrer & Bromme, 2013). Außerdem sind Personen weniger unsicher in ihren Urteilen zum Dissens, wenn eine Person mit höherem Wissen eine der beiden divergierenden Aussagen geäußert hatte (Rouet, Saux et al., 2021, Experiment 1).
- Ähnliches gilt für die *Aufrichtigkeit*: Divergieren Quellen-Paare in dieser Hinsicht, werden Unterschiede der Aussagen eher auf Motive von Wissenschaftlerinnen und Wissenschaftlern zurückgeführt (Gottschling & Kammerer, 2021; Gottschling et al., 2019, 2020; Thomm et al., 2015, Studie 2).
- Es gibt *Aufgabenspezifika bei den Erklärungsmustern*. Dies legen die bislang nur vereinzelt vorliegenden Studien nahe. Ihnen zufolge sind metadatenbasierte

Erklärungsmuster in gesellschaftswissenschaftlichen Themen (im Vergleich zu naturwissenschaftlichen Themen) aus Sicht der Teilnehmerinnen und Teilnehmer besser dazu geeignet, als Dissensursache zu gelten (Thomm et al., 2017). Dies deutet darauf hin, dass Personen domänenspezifische Erklärungsmuster bemühen.

Die Auflösung der Konflikte als finaler Schritt in Modellen wie dem „Content-Source Integration Model“ (Braasch & Scharrer, 2020; Stadtler & Bromme, 2014) wirkt damit im Licht der experimentellen Daten in einem guten halben Dutzend Studien eindeutig. Zugleich ist als Limitation hervorzuheben, dass in diesen Studien ganz gezielt und mit hoher Salienz Metadaten variiert wurden, was aus einer experimentellen Logik erklärbar ist. Wie Personen unter ökologisch validen Umständen und mit einer erhöhten Anzahl von Dokumenten Konfliktursachen mit Metadaten untermauern und angemessen interpretieren, ist derzeit noch zu wenig empirisch erforscht. Allerdings bildet nach ersten Studien mit ex post erhobenen Sourcingmaßen in offen gehaltenen Tests das Nutzen von Metadaten für Konfliktauflösungen eher die Ausnahme (Kobayashi, 2014).

4.3 Determinanten und interindividuelle Unterschiede

Die Multiple-Documents-Literacy-Forschung hat sich in Theorie und Empirie mit verschiedenen Einflussfaktoren auf den Aufbau komplexer mentaler Modelle zu beschäftigen begonnen. Drei von ihnen werden an dieser Stelle fokussiert: das Vorwissen zum Gegenstand ([4.3.1](#)), die inhaltlichen Überzeugungen zum Thema ([4.3.2](#)) sowie die epistemischen Überzeugungen ([4.3.3](#)). Es handelt sich um in der Forschung dominierende Konstrukte (Espinas & Chandler, 2024).

4.3.1 Thematisches Vorwissen

Das thematische Vorwissen gilt als wichtige Bedingung für das Inferieren. Denn wenn Personen zu wenig Vorwissen haben, misslingt es ihnen, schlussfolgernde Beziehungen zwischen einzelnen Propositionen herzustellen (Noordman et al., 2015). Anders ausgedrückt bedarf die Realisation potenzieller und sinnhafter Zusammenhänge zwischen Informationen und nicht zuletzt die Klärung des Verhältnisses von Propositionen mittels komplexer kognitiver Mustererkennungen einer soliden Vorwissensbasis (Alexander & Disciplined Reading and Learning Research Laboratory, 2020). In diesem Teilkapitel werden die Zusammenhänge von Vorwissen und kombinierten Repräsentationen von mindestens zwei Informationseinheiten betrachtet. Dies betrifft erstens die intertextuellen Inferenzen ([4.3.1.1](#)) und zweitens die Quelle-Inhalt-Verbindungen ([4.3.1.2](#)).

4.3.1.1 Intertextuelle Inferenzen und Konflikte

Produkte

Produktbezogene Maße zu Inferenzen dominieren die Forschung bei rezeptiven Maßen (Primor & Katzir, 2018), und sie sind zusätzlich häufig in den Studien mit Vorwissenstests zum Einsatz gelangt. Aus bi- und multivariaten Analysen in diesen Studien sind zwei Ergebnislinien relativ eindeutig zu identifizieren:

- *Hohes Vorwissen korreliert mit den und prädiziert die Leistungen des intertextuellen Inferierens in intertextuellen Verifikationsaufgaben.* Je höher das erfasste Vorwissen in Studien war, desto mehr intertextuelle Inferenzen konnte die Person korrekt als solche in intertextuellen Verifikationsaufgaben erkennen, was viele Einzelstudien immer wieder in Korrelationsanalysen berichten (Bråten & Strømsø, 2006b, 2010a, 2011; Gil et al., 2010a, Experiment 1, partiell auch Experiment 2 in Abhängigkeit der Aufgabenstellung; Gil et al., 2010b; Hagen et al., 2014; List, 2021; List, 2023b, Studie 3; Mateos et al., 2016). Teils war das Vorwissen in regressionsanalytischen Auswertungen unter Berücksichtigung weiterer Prädiktoren dazu geeignet, Varianz aufzuklären (Bråten & Strømsø, 2010b; Bråten et al., 2009; Davis et al., 2017; Incognito & Tarchi, 2024; Strømsø & Bråten, 2009; Strømsø et al., 2008, 2010); zudem wurden positive pfadanalytische Beiträge des Vorwissens belegt (Karimi, 2018; Lee & List, 2023a). Das bedeutet: Selbst wenn weitere unabhängige Variablen rechnerisch berücksichtigt wurden, leistete das Vorwissen eigene und vor allem positive Beiträge. Diese Ergebnisse aus einer großen Zahl von Studien mit positiven Zusammenhängen von Vorwissen und Verifikationsaufgaben werden von solchen kontrastiert, in denen es keine statistischen Zusammenhänge von Vorwissen und Inferenzleistungen gab. Solche Studienergebnisse bilden im Vergleich jedoch eher die Ausnahme (Braasch, Bråten, Strømsø & Anmarkrud, 2014; Castells et al., 2023; Salmerón et al., 2010). Eine andere Form von Testverfahren, nämlich mit offenen intertextuellen Fragen zu gelesenen Texten, unterstützt die Befunde zu den geschlossenen Testformaten zum Teil: Wer ein höheres Vorwissen hat, beantwortete diese Fragen korrekter (Castells et al., 2022; Mason et al., 2010a). Da viele solcher offenen Formate von Tests mit Diskurssynthesen erfasst wurden, sei auf weitere Befunde hier nur verwiesen, die jedoch speziell bei den Inferenzen nicht immer eindeutig sind (s. Tabelle 29 und dort „Lokale Kombinationen“ in Teilkap. [5.3.2.1](#)).
- *Höheres Vorwissen korreliert mit besseren Erinnerungen an inhaltliche Konflikte bei inhaltlich divergierenden Dokumenten.* Einen Sonderfall bzw. einen Ausschnitt der intertextuellen Inferenzen stellen die erkannten intertextuellen Konflikte dar. Aufgaben, die mittels Verifikationsaufgaben die Detektion von

Dissens zwischen Dokumenten beurteilen ließen, kamen in Studien mit sich inhaltlich widersprechenden Dokumenten zum Einsatz. Die Leistung, die solche Tests messen, besteht darin, korrekt im Gedächtnis zu repräsentieren, dass und in welchen Aussagen die gelesenen Dokumente inhaltlich divergierten. Je höher das Vorwissen der Testpersonen war, desto mehr Konflikte hatten sie korrekt detektiert und memoriert (Braasch et al., 2021; Delgado et al., 2020; Kammerer et al., 2019; Stang Lund et al., 2017, 2019). Selbst in Studien, in denen die statistische Signifikanz beim Zusammenhang von Verifikationstest-Ergebnis und Vorwissen verfehlt wurde, zeigte sich dieses Muster in der erwähnten Richtung (Stadtler, Scharrer, Brummernhenrich & Bromme, 2013; Stadtler et al., 2014).

Zusammengefasst scheint ein höheres Vorwissen das korrekte Schlussfolgern in Tests mit binären Antwortkategorien zu erleichtern. Besonders konsistent wirkt dieses Befundmuster bei intertextuellen Inferenzen, die per Verifikationsaufgaben operationalisiert wurden (Demir et al., 2024).

Prozesse

Intertextuelle Integrationsprozesse sind im Vergleich zu Produktmaßen weniger stark untersucht worden (Barzilai et al., 2018; Primor & Katzir, 2018). Eine Folge davon ist, dass die Effekte des Vorwissens damit ebenfalls weniger stark erforscht wurden. Die Ergebnisse sind demnach fragmentarisch, was erklären mag, dass sie teils inkonsistent sind.

- *Bei nachträglich erfassten Vorgehensweisen bestanden teilweise positive Zusammenhänge mit den berichteten intertextuellen Integrationsstrategien.* Mit einer Skala, welche das intertextuelle Integrieren misst (Bråten & Strømsø, 2011), lassen sich direkt nach einer intertextuellen Aufgabe Integrationsprozesse valide erfassen (Hagen et al., 2014; List, Du, Wang & Lee, 2019, Studie 2). Die wenigen Studien, welche Vorwissenstests und diese Skala einsetzten, fanden schwache bis mittlere Korrelationen – wer mehr zum Thema wusste, demonstrierte ein tendenziell mehr intertextuelle Inferenzen suchendes Vorgehen (Haverkamp et al., 2024; List, 2021; Tarchi, 2019). Selbst in Studien, in denen es keine statistisch signifikanten Zusammenhänge gab, wiesen die Korrelationskoeffizienten in die gleiche Richtung (Bråten & Strømsø, 2011; Hagen et al., 2014).
- *Studien mit Lautdenk-Protokollen zeichnen kein klares Bild.* Die Ergebnisse sind insgesamt heterogen; die Divergenz in der Methodik, den Samples, den Fragestellungen und Stimuli mögen hierfür ursächlich sein. Einige Studien ermittelten keine Korrelationen zwischen Vorwissen und Integrationsprozessen in Lautdenk-Protokollen (Kammerer et al., 2015, 2021; Wolfe & Goldman,

2005). Teils zeigten sich diese Zusammenhänge nur bei einzelnen Texten (Iordanou et al., 2020). Teils ergaben sich Hinweise auf Gruppenunterschiede dergestalt, dass es nur Personen mit hohem Vorwissen gelang, intertextuelle Inferenzen beim Lesen zu bilden, während Personen mit geringem oder mittlerem Vorwissen dazu weniger in der Lage waren (McCrudden et al., 2021).

- *Andere Prozessindikatoren sind spärlich untersucht; wenn sie es sind, ergibt sich kein eindeutiges Muster.* Nicht-reaktive Verfahren der Erfassung von Prozessen sind kaum genutzt worden. Teils wurden sie erhoben, aber nicht spezifisch mit dem Vorwissen gemeinsam analysiert (s. z. B. List & Lin, 2023). Dort, wo dies in den Blick genommen wurden, waren mitunter jedoch keinerlei Zusammenhänge eruierbar. In der Studie von Hagen et al. (2014) konnte beispielsweise das Forschungsteam keinerlei Zusammenhänge zwischen verschiedenen Variablen der Notizen und dem Vorwissen ermitteln.

Die Prozessperspektive wirft aus Sicht der Empirie *summa summarum* gegenwärtig mehr Fragen auf, als sie Antworten gibt. Das gilt insbesondere im Vergleich zur Produktperspektive, die zu einer paradoxen Gesamtanmutung führt: Positive Effekte des Vorwissens lassen sich derzeit eher bei den Produkten abbilden, aber anhand der wenigen Empirie zu den Prozessen weder deutlich nachvollziehen noch erklären. Diese Uneindeutigkeit aus empirischer Sicht bei gleichzeitig hoch veranschlagter theoretischer Bedeutung des Vorwissens ist kein Unikum der Integrationsleistungen (Simonsmeier et al., 2022). In diesem Buch tritt die Uneindeutigkeit des Einflussfaktors Vorwissen ebenfalls mehrfach auf, nämlich bei den Aufgabenprodukten (s. Teilkap. [5.3.2](#)) und bei der Selbstregulation ([6.3.3](#)).

4.3.1.2 Quelle-Inhalt-Verbindungen (Produktperspektive)

Bislang wenig untersucht ist, ob Personen die Verbindungen von Aussagen und ihrer jeweiligen Quelle besser memorieren, wenn sie ein höheres inhaltliches Vorwissen haben. Für die Operationalisierung dieser Fähigkeiten in rezeptiven Tests werden Aussagen präsentiert, die dann ihrem Ursprung korrekt zugeordnet werden müssen. Empirisch ergaben sich mittlere positive Zusammenhänge: Personen mit höherem Vorwissen fällt es laut Korrelationsanalysen leichter, Inhalte und Dokumente korrekt miteinander kombiniert mental zu repräsentieren (Delgado et al., 2020; List, 2021; Strømsø et al., 2010). Dieser Zusammenhang bestand allerdings nicht in allen Studien (Barzilai & Eshet-Alkalai, 2015; Lee et al., 2024).

In Ausnahmen der Experimentalstudien wurde das Vorwissen nebst weiteren Prädiktoren genutzt, um Varianz aufzuklären. Eine pfadanalytische Auswertung in der Studie von Stang Lund et al. (2019) ergab sowohl direkte als auch indirekte Effekte. Demnach waren die korrekten Quelle-Inhalt-Verbindungen den Personen präsenter, weil sie mehr Vorwissen hatten (direkter Effekt) und weil sie mehr inhaltliche Konflikte erkannt hatten, die ihrerseits bei höherem Vorwissen

korrekter detektiert worden waren (indirekter Effekt). Damit ist insgesamt die Rolle des Vorwissens aufseiten der Produkte als förderlich annehmbar, teils auch über mediiierende Zusammenhänge, bei den Prozessen gestaltet sich die Forschungslage als unbefriedigend.

4.3.2 Thematische Überzeugungen

Thematische Überzeugungen werden – analog zum Vorwissen – als interne Ressource für Validierungsprozesse beschrieben, welchen eine verstehensunterstützende Funktion beim Aufbau mentaler Modelle zukommt. Dies wird als Validieren bezeichnet, also als Bestätigung potenziell zu integrierender Propositionen in interne mentale Modelle (Richter, 2015; Singer, 2019). Ein Modell, welches dieses Validieren beschreibt, ist das „RI-Val“-Modell (*Resonance, Integration, Validation*; O’Brien & Cook, 2016b). Dieses Modell ist eine Erweiterung des Construction-Integration-Modells (s. Teilkap. 4.1.2.1) und geht davon aus, dass die drei genannten Prozesse wellenförmig überlappend stattfinden. Mit *Resonanz* ist eine automatisch ablaufende Aktivierung von Gedächtnisinhalten durch gelesene Inhalte gemeint, die dazu führt, dass ausreichend aktivierte Inhalte im Arbeitsgedächtnis zur weiteren Verarbeitung verfügbar stehen. Im Arbeitsgedächtnis werden die Informationen dann potenziell mit den aktuell verfügbaren propositionalen Repräsentationen von Texten über Inferenzen miteinander verbunden, also *integriert*. Der hier wichtige dritte Prozess des *Validierens*, der automatisch abläuft, ist ein Vergleich der jeweils neuen Informationen mit den Inhalten der Repräsentation des Gelesenen, aber auch Gedächtnisinhalten hinsichtlich der inhaltlichen Überlappung und Widerspruchsarmut. Dies wird als „Kohärenzschwellenwert“ bezeichnet (s. dazu schon Teilkap. 2.1.2 – nicht zu verwechseln mit dem „Kohärenzstandard“), welcher je nach Leseziel variiert. Entscheidend ist bei RI-Val, dass sich diese Prozestrias kontinuierlich wiederholt und bei jedem neuen Textsegment vollzieht. Für die Validierung, also den Abgleich, ist entscheidend, dass Informationen, die kongruent sind, leichter verarbeitet werden und zu kohärenten Repräsentationen führen (van den Broek & Helder, 2017). Demgegenüber sind inkongruente bzw. neue Informationen verarbeitungsaufwändiger, weil sie z. B. Veränderungen der mentalen Repräsentation oder der Wissensbestände erfordern (Richter & Singer, 2017) bzw. zu lesestrategischen Prozessen der epistemischen Elaboration zum Erreichen hoher Kohärenzstandards führen (Richter & Maier, 2017).

Die Validierung von eingehenden Informationen im Leseprozess kann anhand verschiedener Benchmarks erfolgen (Isberner & Richter, 2014), hier stehen vor allem die thematischen Überzeugungen im Vordergrund, die eher als automatisch ablaufende Heuristiken zu denken sind (McCrudden, 2020; Richter & Maier, 2017). Thematische Überzeugungen definieren Wolfe und Griffin (2017,

S. 296) als *Position zum Wahrheitsgehalt einer Proposition*. Es geht also nicht um die Wahrheit als solche, sondern darum, was Personen subjektiv für wahr halten. Solche Überzeugungen zum Wahrheitsgehalt bezeichnen Bräten und Strømsø (2020, S. 142) als „ontische Überzeugungen“. Sie grenzen davon die „epistemischen Überzeugungen“ ab, um die es im nachfolgenden Teilkapitel [4.3.3](#) gehen wird.

Validierungsprozesse auf der Basis eigener ontischer Überzeugungen hat das „Two Step Model of Validation“ von Richter und Maier (2017) beschrieben und dabei auf den Effekt der eigenen Überzeugungen vor allem in der ersten Phase fokussiert. Diesem Modell nach dienen eigene Überzeugungen zu Themen als interner Evaluationsstandard, haben aber eine ambivalente Rolle. Der Hintergrund ist, dass die raschen, automatisch ablaufenden Aktivierungsprozesse von Langzeitgedächtnisinhalten (in diesem Fall: den ontischen Überzeugungen) zu einer schnelleren und tieferen Verarbeitung von Propositionen führen. Dies wird als „Text-Überzeugungskonsistenz-Effekt“ bezeichnet. Die Ambivalenz des „Text-Überzeugungskonsistenz-Effekts“ besteht darin, dass er nicht nur rasche Einschätzungen erlaubt, sondern bei kontroversen Sachverhalten hinderlich ist. In diesem Fall erschweren die unkorrigierten Effekte der ontischen Überzeugungen, dass die nicht-überzeugungskonsistenten Propositionen berücksichtigt werden, und unterminieren dadurch eine balancierte, mehrperspektivische Repräsentation (Richter et al., 2020).

Ehe im Folgenden produkt- und prozessbezogene empirische Perspektiven eröffnet werden, ist noch ein darstellerischer Hinweis nötig. Integrationsleistungen werden in der Forschung verschiedentlich operationalisiert, z. B. mittels lesebezogener Aufgaben mit geschlossenen Antworten und über offene Aufgaben (Primor & Katzir, 2018). Die letztgenannten Leistungen sind in Diskurssynthesen in lokalen Kombinationen, teils auch in Textstrukturmaßen oder der Textqualität betrachtet worden und werden in Teilkapitel [5.3.3](#) vertieft. An dieser Stelle werden hingegen nur lesebezogene Befunde behandelt, die sich zudem auf rein textuelle Verstehensleistungen stützen.

4.3.2.1 Produkte

Ein neueres, aber relativ konsistentes Befundmuster besteht darin, dass die Repräsentationsebenen des Gelesenen unterschiedlich sensitiv für den „Text-Überzeugungskonsistenz-Effekt“ sind. Eine propositionale, eher oberflächliche Repräsentation, die meist über Erinnerungsleistungen operationalisiert worden ist, ist weniger anfällig für Verzerrungen als inferenzbasierte tiefere Repräsentationen, bei denen die Leserinnen und Leser Informationen miteinander kombinieren müssen. Dies speist sich aus Befundlinien, die sich studienübergreifend extrahieren lassen:

- *Inferenzbasierte Leistungen, die intertextuelles Leseverstehen und damit tieferes Leseverstehen messen, fallen bei überzeugungskompatiblen Texten besser aus:* Personen bilden Inferenzen also besser bei Inhalten, die ihren Überzeugungen entsprechen. Dieser Effekt wurde sowohl für intratextuelle Inferenzen nachgewiesen (Abendroth & Richter, 2020, 2021, 2023; Karimi & Ashkani, 2023; Karimi & Richter, 2021a, 2021b, 2023a, 2023b; Maier & Richter, 2013, 2014; Pilotek et al., in Druck) als auch für die Übernahme von überzeugungskonsistenten Inhalten in offenen Aufgaben (Karimi, 2024). Es gibt zudem Hinweise darauf, dass Personen eher ausgewogen balancierte Inferenzen herzustellen in der Lage sind, wenn sie moderat ausgeprägte Einstellungen aufweisen (Kardash & Scholes, 1996), was wiederum dafür spricht, dass Einstellungen eher zu polarisierend einseitigen bzw. selektiven Verstehensleistungen führen. Denn das Problem, welches sich aus diesen Befunden ergibt, ist: Passen textuelle Aussagen und ontische Überzeugungen nicht zusammen, ist ein fragmentarisches Textverstehen erwartbar.
- *Bei der Erinnerung an singuläre Inhalte (Repräsentation der Textbasis) besteht kein Text-Überzeugungskonsistenz-Effekt:* Die Studienergebnisse lassen sich hierzu zweigeteilt bündeln. Erstens gibt es Befunde, die für einen umgekehrten Effekt der Konsistenz von Überzeugungen und Textinhalten sprechen: Erwachsene Personen konnten präzisere Erinnerungsleistungen an Inhalte erbringen, die aus Texten stammten, welche inhaltlich nicht den eigenen Überzeugungen entsprachen (Lescarret, Le Floch et al., 2023, Studie 1; Maier & Richter, 2013, 2014). Zweitens liegen ebenfalls Studien vor, die keine Effekte feststellen konnten und nach denen also weder die Überzeugungen pro noch contra mit besseren Erinnerungen an Inhalte korrespondierten (Abendroth & Richter, 2020; Bohn-Gettler & McCrudden, 2018, Experimente 1 und 2; Karimi & Richter, 2021a, 2021b, 2023a, 2023b; Lescarret, Le Floch et al., 2023, Studie 2; Strømsø et al., 2017). Das gilt insbesondere für makropropositionale Inhalte von Texten (Kardash & Scholes, 1995).

4.3.2.2 Prozesse

Die Prozessdaten sind ein Desiderat, doch es ergeben sich drei wichtige Ergebnislinien. Die erste betrifft die Lesezeiten, die zweite Prozesse und Strategien aus Lautdenk-Protokollen und die dritte die komplexen Zusammenhänge von Aufgaben, Prozessen und Produkten. Im Einzelnen sind die Ergebnislinien wie folgt:

- *Der Leseprozess verlangsamt sich an den Stellen, die eine überzeugungsinkonsistente Aussage enthalten:* Einige Studien ergaben längere Lesezeiten bei nicht-überzeugungskonsistenten Inhalten (Maier & Richter, 2014; Wertgen & Richter, 2023, Experiment 1), die sich in Experimenten mit Eye-Tracking bei jenen Sätzen zeigten, welche die überzeugungsinkonsistenten Aussagen

enthielten. Personen lasen diese Aussagen länger, was besonders auf die Argumente im Vergleich zu unterstützenden Details zutraf (Abendroth & Richter, 2023). Teilweise bestanden jedoch auch keine Differenzen in der Lesezeit der Sätze, die den Überzeugungen wider- oder entsprachen (Bohn-Gettler & McCrudden, 2018, Experiment 1).

- *Personen wenden teils andere Strategien an, je nachdem ob sie Texte lesen, welche zu ihren ontischen Überzeugungen passten.* Dabei sind die Unterschiede in den Strategien bezogen auf die reinen Häufigkeiten der Strategieeinsätze bzw. auch andere Leseprozesse in den Studien mit Lautdenk-Protokollen eher gering, sondern zeigen sich eher punktuell. Anders gesagt: In vielen Prozessen bzw. Strategien haben sich keine statistisch signifikanten Unterschiede nachweisen lassen. Doch es gibt erste Hinweise, dass je nach Verhältnis von Überzeugungen und Textinhalt Differenzen bestehen: *Texte, die mit der eigenen Überzeugung konfliktieren*, wurden in Lautdenk-Protokollen stärker kritisiert (Bohn-Gettler & McCrudden, 2018, Experiment 2; Kardash & Howell, 2000), partiell nutzten Personen auch mehr Verbindungen zwischen vorherigen Sätzen und den Aussagen, die ihrer ontischen Überzeugung widersprachen, was als Zeichen einer Widerlegungsstrategie interpretierbar ist (Bohn-Gettler & McCrudden, 2018, Experiment 2). *Überzeugungskonsistente Texte* hingegen wurden stärker metakognitiv in Bezug auf das sich aufbauende Leseverstehen überwacht (Kardash & Howell, 2000) und positiver evaluiert (Bohn-Gettler & McCrudden, 2018, Experiment 2). In anderen Studien war ein Muster, dass Personen weniger Prozesse des Überwachens, der inferenzbasierten Kohärenzbildung und des kritischen Evaluierens bei jenen Texten in Lautdenk-Protokollen geäußert hatten, die sie später in schriftlichen Zusammenfassungen einseitig darstellten (Iordanou et al., 2020).
- *Die gleichen Leseprozesse können je nach Polung der Überzeugungen zum Thema unterschiedliche Effekte auf Verstehensleistungen haben:* Mitunter wurde geprüft, ob und welche differenziellen Zusammenhänge zwischen Aufgabe, Prozessen und Produkt bestehen. Dies illustriert die mit diversen Prozessmaßen (Lesezeiten und Lautdenk-Protokolle) operierende Studie von Maier und Richter (2016), welche bislang eine der sehr wenigen ist, die diese Zusammenhänge so genau prüften. Drei Befunde sind an dieser Stelle hervorhebenswert: Ein erster Zusammenhang besteht in der relativen Zuwendung zu überzeugungskonsistenten bzw. -inkonsistenten Texten in Abhängigkeit von der Aufgabenstellung. So führen oberflächenorientierte Aufgaben wie das Zusammenfassen zu längeren Lesezeiten überzeugungskompatibler Texte im Vergleich zu überzeugungsinkompatiblen Texten, während epistemisch orientierte Aufgaben wie das Verfassen von Argumentationen eher ausgewogene Lesezeiten nach sich ziehen. Diese ausgewogenere Lesezeit – und das ist ein zweiter Effekt – hing mit einer Strategie zusammen, nämlich dem Prüfen der Aussagen aus überzeugungsinkonsistenten Texten. Wer ausgewogen beide

Texte las, evaluierte stärker jene Textinhalte, die der eigenen Überzeugung entgegenliefen. Die Effekte auf die Inferenzleistung waren unterschiedlich und fielen diametral aus. Dieser dritte Effekt lautete: Für Inferenzen zu den überzeugungsinkonsistenten Texten war dieses Prüfen positiv, bei den überzeugungskonsistenten Texten hingegen negativ. Ein und derselbe Prozess, der auftragsspezifisch über die medierende Effekte der Lesezeiten unterschiedlich ausfällt, kann also förderlich oder hinderlich sein.

Die Prozesse gestalten sich in Abhängigkeit von den ontischen Überzeugungen der Leserinnen und Leser anders. Die Personen verarbeiten überzeugungsinkonsistente Informationen anscheinend langsamer und kritischer, was sich teils negativ in den Verstehensmaßen bemerkbar macht.

4.3.3 Epistemische Überzeugungen

Personen tragen verschiedene Arten von Überzeugungen in sich, die nicht alle inhaltlich auf ein Thema bezogen sind. Darauf machen Bråten und Strømsø (2020) aufmerksam, indem sie die ontischen von den epistemischen Überzeugungen trennen. Während sich die erstgenannten auf den Wahrheitsgehalt von Aussagen beziehen und damit der Plausibilitätsprüfung dienen, erstrecken sich die epistemischen Überzeugungen darauf, welche subjektiven Annahmen Personen dazu haben, was Wissen ist und wie es zustande kommt (Bråten, Britt et al., 2011). Solche Annahmen werden unter anderem als zentrale Bestandteile der „persönlichen Epistemologie“ (Hofer, 2001) oder auch als „epistemische Kognitionen“ (Chinn & Rinehart, 2016b) bezeichnet. Hofer (2001, S. 355) definiert ihr Konstrukt „Epistemologie“ als originär philosophisches Konstrukt zur Beschaffenheit und Rechtfertigung menschlichen Wissens und nutzt die psychologisch-bildungswissenschaftliche Forschungsliteratur für eine Operationaldefinition persönlicher Epistemologien als *Denken und Überzeugungen der Lernenden in Bezug auf Wissen als Produkt und Wissen als Prozess*, welche einige bzw. sämtliche der folgenden Elemente umfassen: Überzeugungen über die Definition, Konstruktion, Beurteilung, Ursprünge und Genese von Wissen. Im Zentrum stehen damit die Überzeugungen, welche im Feld der Leseforschung als „epistemische Überzeugungen“ bezeichnet werden.

Die Überzeugungen sind kein einheitliches Konstrukt, was sich an mindestens zwei Stellen zeigt.

- Erstens variieren die *Bezugsgrößen bzw. die Generik und Spezifik*. Epistemische Überzeugungen können sich auf sehr unterschiedliche Bezugspunkte erstrecken (Muis et al., 2006), was mit den verschiedenen Arten von Wissen korrespondiert (Alexander & Schoute, 2022; Anderson et al., 2001). Diese

Bezugspunkte lassen sich theoriebasiert hierarchisch modellieren, und es gibt momentan verschiedene Granularitätsgrößen in der Empirie, wie es die nachfolgende Aufzählung aus Studien aus der Multiple-Documents-Literacy-Forschung illustriert. Epistemische Überzeugungen können sich *a) auf Wissen allgemein* (Davis et al., 2017; Karimi & Atai, 2014; Karimi & Richter, 2021a; Lang et al., 2020, 2021; List, 2021; List, Oaxaca et al., 2021; Mateos et al., 2016; Rukavina & Daneman, 1996), *b) auf das Wissen in Domänen* (Bråten, Ferguson, Strømsø & Anmarkrud, 2013; Ferguson & Bråten, 2013; Ferguson et al., 2013; Kammerer et al., 2021; Mahlow et al., 2022; Mason et al., 2010a; Strømsø et al., 2016; Trevors et al., 2017; Wiley et al., 2020) und *Themengebiete* darin, zum Beispiel Klimawandel (Bråten & Strømsø, 2010b; Bråten et al., 2008; Gil et al., 2010b; Muis et al., 2015; Strømsø & Bråten, 2009; Strømsø et al., 2008), genetisch veränderte Organismen (Muis et al., 2021), Gesundheitseffekte von Nahrungsmitteln (Barzilai & Ka'adan, 2017), ADHS (Bråten & Strømsø, 2006a) sowie Meerwasserentsalzung (Barzilai & Eshet-Alkalai, 2015; Barzilai et al., 2015), und *c) auf spezifische aufgabenbezogene Situationen* erstrecken, indem beispielsweise Lautdenk-Protokolle (Bråten, Ferguson et al., 2014; Cho, Woodward & Li, 2018; Lang et al., 2021; Mason et al., 2010b, 2011) oder aufgabenbezogene Interviews (Barzilai & Zohar, 2012; Ioannou & Iordanou, 2020, Studie 2; Iordanou et al., 2020; Mason et al., 2010a) kodiert wurden. Zudem wurden in der Forschung noch einmal quer dazu eine Perspektive eröffnet, in *d)* der eine kontextuelle Variable berücksichtigt wurde, indem spezifisch *das Internet adressierende Überzeugungen* fokussiert wurden, wobei die Struktur dieser Überzeugungen Eigenheiten aufweist, die sie nicht völlig kompatibel mit anderen Instrumenten macht (Kammerer et al., 2013, 2015; Kammerer & Gerjets, 2012a). Inzwischen bestehen Tendenzen, internetspezifische epistemische Überzeugungen mit ähnlichen Skalen zu kombinieren, wie sie bei domänenspezifischen Fragebögen zum Einsatz gekommen sind (Bråten, Brandmo & Kammerer, 2019).

- Zweitens wurden bislang *unterschiedliche Taxonomien* vorgelegt, deren Unterschiedlichkeit sich daraus ergibt, ob es sich um Fokussierungen auf Dimensionen oder auf Entwicklungsverläufe handelt (Hofer, 2001). Diese Unterschiede reflektieren die Forschungshistorie: Zunächst standen die Verläufe und Entwicklungen von epistemischen Überzeugungen im Vordergrund, Fragen des Systems von Überzeugungen und der Trennbarkeit von Überzeugungen bildeten den Gegenstand einer weiteren, später aufgekommenen Forschungslinie (Hofer, 2016). Dementsprechend variieren die Perspektiven: Entwicklungsmodelle (s. Teilkap. [4.3.3.2](#)) konturieren idealtypische phasenhafte Verläufe des Erwerbs von elaborierteren Überzeugungen und beschreiben die Ausprägungen von epistemischen Überzeugungen. Dimensionale Modelle ([4.3.3.1](#)) klassifizieren und differenzieren verschiedene, durchaus

parallel vorkommende und in individuellen Konstellationen ausgeprägte Überzeugungen.

Die unterschiedlichen Zugänge zur Modellierung von epistemischen Überzeugungen schlagen sich in unterscheidenden Ergebnissen und Ergebnislinien nieder. In Bezug auf die Multiple-Documents-Literacy-Forschung überwiegen die Studien in der Tradition dimensionaler Taxonomien. Darum beginnt die Darstellung mit ihnen.

4.3.3.1 Zugang 1: epistemische Überzeugungen in der Tradition der dimensional Taxonomien

Das hochdynamische Forschungsfeld zu epistemischen Überzeugungen mit seinem Ausschnitt, der Multiple-Documents-Literacy-Forschung, wird in einem mehrschrittigen Vorgehen in diesem Teilkapitel behandelt. Zu Beginn eröffnet eine theoriebasierte Systematik verschiedener (Teil-)Dimensionen den Gegenstandsbereich. Es folgt eine Betrachtung jeder einzelnen Teildimension, also ein konstruktbezogener Zugang. In einem letzten Darstellungsschritt werden Studienergebnisse präsentiert, die auf der Kombination von verschiedenen epistemischen Überzeugungen bei Gruppen von Personen basieren. Eine Zusammenfassung schließt das Teilkapitel ab.

Eine Taxonomie von zwei Haupt- und ihren jeweiligen beiden Teildimensionen

Der erste hier vorzustellende Strang in Theorie und Empirie widmete sich der Frage nach der Dimensionalität in den epistemischen Überzeugungen. Erkenntnisleitend ist also, wie viele Arten von Überzeugungen sich theoretisch beschreiben und empirisch erforschen lassen. Aufgrund der hohen Heterogenität in relativ zeitgleich vorgelegten Modellierungen mit divergierenden Bezeichnungen bei zugleich teils großen inhaltlichen Überlappungen haben Hofer und Pintrich (1997) eine Theoriesynopse vorgelegt. Sie destillierten aus verschiedenen, auch entwicklungsbezogenen Taxonomien zwei sogenannte Hauptdimensionen der epistemischen Überzeugungen nebst zwei ihnen untergeordneten Teildimensionen (s. Tabelle 21, v. a. Teil A). Als Hauptdimensionen werden Merkmale von Wissen („Nature of Knowledge“) und die Genese bzw. Mechanismen von Wissensherstellung („Nature of Knowing“) postuliert. Im Falle der Merkmale des Wissens werden die Teildimensionen *Einfachheit* und *Bestimmtheit* des Wissens als Facetten angenommen. Erstgenannte Teildimension betrifft die Komplexität bzw. Einfachheit von Bestandteilen von Wissen. Die zweitgenannte Teildimension bildet ab, ob Wissen fixiert und starr oder veränderlich und dynamisch ist. Die Genese des Wissens ist ebenfalls in zwei Teildimensionen untergliedert. Eine davon betrifft die *Begründbarkeit von Aussagen als Wissen*, die zweite die *soziale*

Quelle von wissensbezogenen Aussagen. Jede dieser vier Teildimensionen ist als Kontinuum vorstellbar, auf dem sich die jeweilige Person mit ihren epistemischen Überzeugungen abbilden lässt (s. Tabelle 21, Teil A). Genau dies ist – im Verbund verschiedener Teildimensionen – auch bei den Entwicklungsmodellen der Fall.

Hofers und Pintrichs (1997) Taxonomie bzw. ihre Grundlogik ist bis heute einflussreich (Schraw, 2013) und wurde teils noch erheblich ausdifferenziert, unter anderem wurde die Frage nach Domänenspezifität in Bezug auf die Komplexität und Offenheit der Herstellung von wissensbezogenen Aussagen aufgeworfen (Dimension Merkmale von Wissen); vor allem aber wurde in der Hauptdimension Genese von Wissen die zentrale Frage in den Fokus gerückt, wie Personen begründen, was sie warum für Wissen halten (Chinn et al., 2011; Greene, Azevedo & Torney-Purta, 2008).

In der Multiple-Documents-Literacy-Forschung wurden die Impulse zu den epistemischen Überzeugungen ebenfalls aufgenommen und die epistemischen Überzeugungen als Einflussfaktor für interne metakognitive Standards beim Bearbeiten von Aufgaben modelliert (List & Alexander, 2019). Den ersten umfassenden Vorschlag zur Verbindung der Forschungsgebiete von epistemischen Überzeugungen und Verstehensprodukten bei multiplen Dokumenten haben Bråten, Britt et al. (2011) vorgelegt. Sie haben dabei das Dokumentenmodell primär aus einer Produktperspektive mit den epistemischen Überzeugungen angereichert. Ihr Modell nimmt epistemische Überzeugungen als Einflussfaktor auf und spezifiziert, welche Teildimensionen der Überzeugungen die mentale Repräsentation beider Ebenen mutmaßlich beeinflussen (s. Tabelle 21, v. a. Teil B). Zusätzlich wurde auch partiell noch das Aufgabenmodell (s. Teilkap. [2.1.1](#)) betrachtet, um die Zusammenhänge zu beschreiben, die – dem damaligen Zeitpunkt und der damaligen Forschungslage geschuldet – in ihrer Vollständigkeit und Fokussierung auf Aufgabenmodell, Prozesse und Produkte im Original teils stark variieren. Die Darstellung in Tabelle 21 fällt deshalb unterschiedlich differenziert aus.

Tabelle 21: Vergleich vierer Teildimensionen epistemischer Überzeugungen

Vergleichsaspekt	Hauptdimension „Nature of Knowledge“ (Merkmale von Wissen)		Hauptdimension „Nature of Knowing“ (Genese von Wissen)	
A) Allgemeine Informationen				
Name der Teildimension	Einfachheit des Wissens		Bestimmtheit des Wissens	
Charakterisierung	Subjektive Annahmen zur Struktur von Elementen des Wissens	Subjektive Annahmen zur Veränderlichkeit von Wissen	Subjektive Annahmen zur Herstellung von Aussagen als Wissen	Subjektive Annahmen zur sozialen Herkunft von wissenschaftlichen Aussagen
	Pol Komplexität: Wissen als untereinander hochgradig verknüpfte Konzepte vs. Pol Einfachheit: Wissen als Sammlung isolierter Fakten	Pol Veränderung: Wissen ist vorläufig und veränderlich vs. Pol Konstanz: Wissen ist absolut und unveränderlich	Pol Überprüfbarkeit über epistemisch verlässliche Quellen: Wissen entsteht regelgeleitet und über kritischen Vergleich mehrerer Quellen* vs. Pol rein subjektive Überprüfbarkeit: Wissen ist entweder von anderen oder dem eigenen ‚gesunden Menschenverstand‘ abhängig*	Pol Expert:innen als Quelle: Wissen liegt außerhalb der eigenen Person bei Autoritäten, von wo es übertragen wird vs. Pol Selbst als Quelle: Wissen ist Ergebnis der Ko-Konstruktion von eigener Person und anderen Personen/Quellen
B) Spezifische Anbindung an Bestandteile des Dokumentenmodells				
Mutmaßlich stärkerer Bezug zur Ebene des Dokumentenmodells	Integriertes mentales Modell (wegen komplexer Inferenzen zur Erreichung von lesebezogenen Zielen)	Intertextmodell (wegen Rekonstruktion der verschiedenen Perspektiven aus Dokumenten)	Integriertes mentales Modell (wegen Vollständigkeit von inhaltlichen Perspektiven)	Intertextmodell (wegen Berücksichtigung aller fachlich geeigneten Quellen)

Vergleichsaspekt	Hauptdimension „Nature of Knowledge“ (Merkmale von Wissen)	Hauptdimension „Nature of Knowing“ (Genese von Wissen)
Annahmen zum Effekt auf integriertes mentales Modell mit kontrastiver Kurzcharakterisierung bzgl. Polen	Je nach Ausprägung im Kontinuum haben Leser:innen und Leser auf intertextuelle Kohärenz fokussierte oder eher intratextuelle Aufgabenmodelle, nutzen stärker oder weniger intertextuelle Integrationsprozesse und bauen vollständigere bzw. isolierte integrierte mentale Modelle auf. • <i>Pol Komplexität</i>: Aufgabenmodell fokussiert auf vollständige und metakognitiv gesteuerte Integration im b) sich anschließenden Prozess und c) auf kohärente Repräsentation als Produkt • <i>Pol Einfachheit</i>: Aufgabenmodell fokussiert auf Akkumulation mit entsprechend oberflächenorientierten Strategien als Prozesse und fragmentarischer Repräsentation als Produkt	Je nach Ausprägung im Kontinuum suchen Personen mehr oder weniger intensiv nach Aussagen von Expert:innen und gewichten diese höher oder geringer, was zu gewichteten, tendenziell tiefer verarbeiteten oder ungewichteten, eher oberflächlichen Repräsentationen führt. • <i>Pol Expert:innen als Quelle</i>: Prozesse ausgerichtet an der (intertextuellen) Suche nach Wissen und Vergleich von Quellen mit Expertise, Produkt beinhaltet vor allem Elemente aus Expert:innenwissen • <i>Pol Selbst als Quelle</i>: weniger Suche nach Expert:innenwissen beim Aufbau des Verstehensprodukts, welches undifferenzierter und eher oberflächlich verarbeitete Inhalte beinhaltet

Vergleichsaspekt	Hauptdimension „Nature of Knowledge“ (Merkmale von Wissen)	Hauptdimension „Nature of Knowing“ (Genese von Wissen)
Annahmen zum Effekt auf Intertextmodell mit kontrastiver Kurzcharakterisierung bzgl. Polen	Je nach Ausprägung im Kontinuum werden Dokumente mit komplexeren bzw. simpleren Darstellungen unterschiedlich beurteilt und im Verstehensprozess kognitiv berücksichtigt. • Pol <i>Komplexität</i>: evaluative Prozesse, die Quellen und Dokumente mit vereinfachten Darstellungen als weniger plausibel erachten, und solche Quellen/Dokumente in der mentalen Repräsentation peripher berücksichtigen – vice versa bei Quellen/Dokumenten mit komplexeren Darstellungen • Pol <i>Einfachheit</i>: evaluative Prozesse, die Quellen und Dokumente mit komplexeren Darstellungen als weniger plausibel erachten, und solche Quellen/Dokumente in der mentalen Repräsentation peripher berücksichtigen – vice versa bei Quellen/Dokumenten mit simpleren Darstellungen	Je nach Ausprägung im Kontinuum gibt es Unterschiede in Aufgabenmodell, Prozessen und Produkten bei Stellenwert, Suche, Akzeptanz und Repräsentation von eigener Person und anderen Quellen mit Expertise. • Pol <i>Expert:innen als Quelle</i>: Aufgabenmodell mit dem Ziel, Expert:innen-Perspektiven zu verstehen, mit einer starken Zuwendung zu externen Ressourcen, positiver Evaluation von Expert:innen-Aussagen und stärkerem Differenzieren von Quellen, kognitives Repräsentieren mit Betonung von Dokumenten mit Expert:innen-Perspektiven • Pol <i>Selbst als Quelle</i>: weniger Suche nach Expert:innenwissen beim Aufbau des Verstehensprodukts, welches undifferenzierter und eher oberflächlich verarbeitete Inhalte beinhaltet
Überblick über Empirie	Je nach Ausprägung im Kontinuum verstehen Leser:innen und Leser den Aufgabenzweck als Suchen von singulären und fixen bzw. diversen Zugängen, suchen fixierte bzw. dynamische Zugänge zum Thema und repräsentieren im Produkt eindeutige oder vorläufige Zugänge. • Pol <i>Veränderung</i>: Aufgabenmodell sieht Zweck des Lesens als Exploration von Perspektiven und Evolvierung von verschiedenen Sichtweisen vor, darum im Prozess Fokus auf intertextuelle Verbindungen und Quelle-Inhalt-Verbindungen, Produkte mit vernetzter Repräsentation von Breite und Vielfalt von Perspektiven • Pol <i>Konstanz</i>: Aufgabenmodell beschränkt sich auf Finden einer einzigen wahren Antwort und nicht-klarungsbedürftigen Lösung, deshalb Prozesse, die nach dem Finden dieser Antwort/Lösung stoppen, statt andere Quellen zu konsultieren, was zu einem kognitiven Produkt mit eindeutiger Lösung führt	Je nach Ausprägung im Kontinuum absolvieren Leser:innen und Leser mehr oder weniger Sourcingprozesse, die dann in vernetzte und evaluativ angeordnete Dokumentenketten münden oder in fragmentarische Intertextmodelle. • Pol <i>Überprüfbarkeit über epistemisch verlässliche Quellen</i>: Fokus auf verschiedene und kritische Sourcingprozesse der Einschätzung von Quellen und generelle stärkere Zuwendung zu Metadaten, mehr Repräsentationen von Metadaten und deren Evaluation im Dokumentenketten • Pol <i>retin subjektive Überprüfbarkeit</i>: wenig Sourcingprozesse, geringere Repräsentationen von Metadaten im Dokumentenketten, tendenziell fragmentarische Intertextmodelle
Überblick über Empirie	Tabelle 22	Tabelle 23
		Tabelle 24
		Tabelle 25

(Quelle der Darstellung: angereicherte Informationen auf der Basis von Bräten, Britt et al., 2011;

Legende: * Pole im Original anders bezeichnet, aber in der Zeit nach der Publikation wurde die Dimensionalität in der Forschung weiter bearbeitet)

Teildimension 1: Einfachheit des Wissens (Produktperspektive)

Die Ergebnisse zu den einzelnen Teildimensionen weisen insgesamt eine große Heterogenität in der Operationalisierung der Konstrukte auf und wurden zudem in den Studien mit teils sehr unterschiedlichen Indikatoren der Verstehensleistungen erfasst. Um der Heterogenität sinnvoll zu begegnen, wird in den nachfolgenden Teilkapiteln der Forschungsstand synoptisch in Tabellenform dargestellt. Dabei werden positive und negative Effekte ebenso wie Nulleffekte in einzelnen Spalten ausgewiesen. Positive Effekte meinen, dass ein statistisch positiver Zusammenhang in Korrelations- bzw. Regressionsanalysen gefunden wurde, welcher der gepolten Richtung aus Tabelle 21 entspricht. Negative Effekte bezeichnen das Gegenteil, also einen erwartungswidrigen Befund. Bei den Nulleffekten wurden keine linearen Zusammenhänge ermittelt.

Mit diesen darstellerischen Vorbemerkungen lässt sich das Ergebnismuster aus Tabelle 22 als relativ konsistent bezeichnen: Wer der Auffassung ist, dass Wissen komplex ist, erbringt in der Tendenz bessere Verstehensleistungen. Befunde zu Indikatoren des Intertextmodells und zu den Prozessen liegen aktuell nicht bzw. kaum vor.

Tabelle 22: Überblick zum Zusammenhang der Teildimension „Einfachheit des Wissens“ mit Produktmaßen aus Studien der Multiple-Documents-Literacy-Forschung

Ebene des Dokumentenmodells	Positive Effekte	Nulleffekte	Negative Effekte
Integriertes mentales Modell	• Intertextuelle Inferenzen (Bråten & Strømsø, 2010a; Mateos et al., 2016; Muis et al., 2015, Studien 1 und 2; Strømsø et al., 2008) bzw. Inferenztest (Davis et al., 2017) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domäne Geschichte) • Kritisches Denken in Diskurssynthesen (Muis et al., 2021) • Integrationsleistungen in offenen Fragen (Rukavina & Daneman, 1996)	• Anzahl von zweiseitigen Konfliktberichten (Stadtler et al., 2014) • Erinnerung an konfligierende Themen (Konfliktverifikation) (Stadtler et al., 2014) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domäne Physik) • Balanciertheit von Argumenten in offener Aufgabe (Kammerer et al., 2013) • Entscheidung für eine objektiv angemessenere Option (Kammerer et al., 2015)	
Intertextmodell		Anzahl Referenzierungen (Stadtler et al., 2014)	

Teildimension 2: Bestimmtheit des Wissens (Produktperspektive)

Die Überzeugungen zum Wissen als unveränderliche oder sich wandelnde Größe wurden vergleichsweise häufig, aber bislang ausschließlich für Indikatoren des integrierten mentalen Modells erforscht (s. Tabelle 23).

Tabelle 23: Überblick zum Zusammenhang der Teildimension „Bestimmtheit des Wissens“ mit Produktmaßen aus Studien der Multiple-Documents-Literacy-Forschung

Ebene des Dokumenten-modells	Positive Effekte	Nulleffekte	Negative Effekte
Integriertes mentales Modell	• Intertextuelle Inferenzen (Bråten & Strømsø, 2010b; Mateos et al., 2016; Strømsø & Bråten, 2009; Strømsø et al., 2008) • Evaluation wissenschaftlicher Kontroversen (Lang et al., 2020, 2021) • Kritisches Denken in Diskurssynthesen (Muis et al., 2021)	• Intertextuelle Inferenzen (Bråten & Strømsø, 2010a; Gil et al., 2010b; Muis et al., 2015, Studien 1 und 2) bzw. Inferenztest (Davis et al., 2017) • Erinnerung an konfligierende Themen (Konfliktverifikation) (Stadtler, Scharrer, Brummernhenrich & Bromme, 2013; Stadtler et al., 2014, 2020) • Anzahl von zweiseitigen Konfliktberichten (Stadtler, Scharrer, Brummernhenrich & Bromme, 2013; Stadtler et al., 2014, 2020) • Intratextuelle Inferenzen: überzeugungskonsistent (Karimi & Richter, 2021a) • Anzahl Switches in Diskurssynthesen (Gil et al., 2010b) • Gesamttransformationsindex (Gil et al., 2010b) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domänen Geschichte und Physik) • Anzahl von Inhalten objektiver Webseiteninhalte in schriftlichen Antworten (Kammerer & Gerjets, 2012a)	• Intratextuelle Inferenzen: überzeugungsinkonsistent (Karimi & Richter, 2021a)
Intertext-modell		Anzahl Referenzierungen (Stadtler et al., 2014)	

Die Forschungslage ist insgesamt uneindeutig, was sich bei den intertextuellen Inferenzen im oberen Teil der Tabelle 23 besonders verdeutlichen lässt. Neben einigen positiven Zusammenhängen gab es sehr viele Null-Ergebnisse. Der auffälligste Befund ist demnach, dass die Studien erstens grosso modo keine linearen Zusammenhänge bei verschiedenen Indikatoren des integrierten mentalen Modells erbracht haben. Zweitens fehlt es an Studienergebnissen, die den von Bräten, Britt et al. (2011) angenommenen Zusammenhang von epistemischen Überzeugungen und Bestandteilen des Intertextmodells substanziell untersucht und unterfüttert hätten. Drittens besteht ein Mangel an Studien mit einem Fokus auf die Prozesse (mit der Ausnahme der Eye-Tracking-Studie von Lang et al., 2020).

Auch wenn die Zusammenhänge allgemein unklar wirken, so gibt es doch Ergebnisse aus Studien, die spezifische Zusammenhänge nahelegen.

- *Überzeugungen zur Bestimmtheit wirken mit anderen Überzeugungen zusammen.* Einen differenzierten Befund ergab die Studie von Karimi und Richter (2021a). Sie kamen zum Ergebnis, dass die Überzeugung, Wissen sei fixiert und unveränderlich, nicht durchgängig mit Leseverstehensleistungen zusammenhängt. Sie gingen darum gezielt auf die Inferenzleistungen bei einzelnen Texten ein und ermittelten zweierlei: Erstens gab es Effekte nur bei den Inferenzleistungen in ihrem Variablenset, und zweitens betraf dies nur einen Teil. Forschungsparadigmatisch prüfte die Studie den Text-Überzeugungskonsistenz-Effekt (s. Teilkap. 4.3.2), also ob Verstehensleistungen von den thematischen Überzeugungen abhängen. Tatsächlich ergab sich dieser Effekt zuungunsten der überzeugungsinkonsistenten Inhalte. Es erwies sich als besonders verstehenshinderlich, wenn die Personen stärker ausgeprägte Überzeugungen bei der Bestimmtheit des Wissens aufwiesen. Epistemische und ontische Überzeugungen interagierten demnach: Wer Wissen für fixiert und unveränderlich hält, dem gelingt es schlechter, Inferenzen zu bilden, die den eigenen thematischen Überzeugungen zuwiderlaufen.
- *Überzeugungen zur Bestimmtheit betreffen spezifische Verstehensleistungen.* Zu diesem Resultat gelangte die Studie von Mahlow et al. (2022). Das Forschungsteam prüfte in vertiefenden Analysen, welche von vier im generischen Test erforderlichen Teilleistungen mit verschiedenen Facetten der domänenspezifischen epistemischen Überzeugungen statistisch interagierten. Sie stellten fest, dass die Leistungen, intertextuelle Aussagen zu integrieren, dann besser waren, wenn Personen der Auffassung waren, dass Wissen in der Domäne Physik veränderlich ist. Bei anderen Teilleistungen und bei den Bestimmtheitsüberzeugungen in der Domäne Geschichte ließ sich dieser Effekt nicht nachweisen.

Es wirkt damit gesamthhaft so, dass die genaueren Zusammenhänge für die Teildimension der Bestimmtheit des Wissens derzeit nicht ausreichend geklärt sind, da sich möglicherweise Moderationseffekte ergeben, die noch zu wenig verstanden sind und mehr Forschung erfordern.

Teildimension 3: Begründung des Wissens (Prozess- und Produktperspektiven)

Die Teildimension „Begründung des Wissens“ ist in dem Gesamt aller vier Teildimensionen eine hervorstechende Ausnahme. Der Grund dafür ist, dass sie einer theoriebasierten und -getriebenen Veränderung unterworfen ist (Perspektive 1), welche dazu geführt hat, dass die empirische Forschung das ursprünglich alleinige Kontinuum zugunsten mehrerer Kontinua aufgegeben hat und mit Produktmaßen in Verbindung gebracht hat (Perspektive 2). Zu guter Letzt hat die Forschung damit begonnen, Prozesse in ihrem Konnex mit den Überzeugungen zu erforschen und den Zusammenhang von Überzeugungen zur mehrdimensionalen Begründung des Wissens, diversen Prozessen und deren Ergebnissen in Produkten zu bestimmen (Perspektive 3).

Perspektive 1: Theoretische Aufwertung und Ausweitung der Binnendimensionierung

Die Dimensionalität von einzelnen und gut trennbaren Überzeugungen ist in der Forschung längst nicht ausreichend geklärt, was gerade für die Teildimension „Begründung des Wissens“ gilt (s. dazu im Detail Buehl, 2008, und Greene, Azevedo & Torney-Purta, 2008). Dies mag mit der Heterogenität in puncto Inhalt zusammenhängen. Tatsächlich fällt bei der Beschreibung des Kontinuums in dem Modell von Bräten, Britt et al. (2011) auf, dass sich in dem Kontinuum zwischen den von ihnen ursprünglich als „adaptiv“ bzw. „fehlangepasst“ bezeichneten Polen sehr unterschiedliche Begründungen zur Genese des Wissens finden. Diese beinhalten internal basierte Prüfungen auf Plausibilität ebenso wie den Rekurs auf externale Quellen. Diese Schwachstelle adressiert die Kritik von Greene, Azevedo und Torney-Purta (2008). Sie basiert auf philosophischen Perspektiven, welche knapp ausgedrückt die Begründungen als mindestens zweidimensional annehmen, nämlich zum einen basierend auf Begründungen der eigenen Person und zum anderen basierend auf Aussagen von anderen, nämlich epistemischen Autoritäten, also Personen mit hoher Sachkenntnis. Diesen Grundgedanken, mehrere Dimensionen anzunehmen, haben wenig später Chinn et al. (2011) weiter ausformuliert. Sie postulieren aus einer philosophischen Warte eine erhebliche Ausweitung von Wissensquellen und damit den verschiedenen Begründungen über diese Wissensquellen und plädieren stark dafür, verschiedene Wissensquellen und die sich auf sie beziehenden epistemischen Überzeugungen als parallele und interagierende Konstrukte anzunehmen.

Ein solcher mehrdimensionaler Zugang liegt einer Re-Modellierung dessen zugrunde, was als elaborierte epistemische Überzeugungen gefasst werden kann (Bromme et al., 2008). Solche Überzeugungen sind demnach nicht nur mehrdimensional, sondern auch adaptiv. Bromme et al. (2008) gehen deshalb davon aus, dass es auf einen kontextsensitiven und flexiblen Umgang mit den epistemischen Überzeugungen ankommt, der dezidiert Mischverhältnisse nicht nur zulässt, sondern die Situationsadäquanz und nicht zuletzt die eigenen Wissensbestände inkludiert. Wer zum Beispiel wenig Wissen hat, wäre demnach weniger dazu in der Lage, Aussagen auf ihre epistemische Plausibilität hin zu überprüfen und kann sich darum eher auf Begründungen über multiple Quellen und/oder Aussagen von Expertinnen und Experten stützen. Umgekehrt kann dies bei hohen eigenen Wissensbeständen einer inversen Konstellation unterliegen. Diese Logik entspricht der Differenzierung von Erste- und Zweite-Hand-Einschätzungen (s. Teilkap. [4.2.2](#)) und verknüpft sie mit den epistemischen Überzeugungen zur Genese des Wissens.

Die aus philosophischen Impulsen stammende veränderte Sicht auf epistemische Überzeugungen zur Begründung des Wissens hat sich in der Leseforschung niedergeschlagen. So wurde die Mehrdimensionalität dieser Teildimension sowohl in Lautdenk-Protokollen erfasst (Ferguson et al., 2012) als auch in einen domänenspezifischen Fragebogen mit eigenen Skalen überführt (Ferguson et al., 2013). Diese Arbeiten waren besonders ertragreich für die Multiple-Documents-Literacy-Forschung, die bei der Teildimension „Begründung des Wissens“ inzwischen die differenziertesten und zahlreichsten Erkenntnisse im Vergleich zu anderen Teildimensionen hervorgebracht hat. Diese auffällig starke Forschungsaktivität sorgte für ein vielschichtiges und reichhaltiges Bild von Zusammenhängen einzelner Facetten der epistemischen Überzeugungen mit diversen Produktmaßen (darunter auch solchen zum Intertextmodell) und zusätzlich mit Prozessindikatoren.

Die Produktmaße sind verschiedentlich in den Blick der Forschung geraten. Die Darstellung in Tabelle 24 trennt zwischen den Indikatoren des integrierten mentalen Modells im oberen Teil und denen des Intertextmodells im unteren Teil. Vertikal unterscheidet die Darstellung noch die verschiedenen Facetten der epistemischen Überzeugungen, nämlich die Logik der Trichotomie gemäß Bråten, Ferguson, Strømsø und Anmarkrud (2013) und – ganz zuoberst – noch in einer generischen Gesamtskala zu Begründungen. Die studienübergreifende Verteilungsmuster bei Indikatoren des *integrierten mentalen Modells* lauten:

- *Generische Begründungen* haben sich nicht klar als günstig erwiesen, denn die positiven Zusammenhänge sind ähnlich häufig wie Nulleffekte.
- Die *Begründungen zu multiplen Quellen* erweisen sich mehrheitlich günstiger für das Verstehen von multiplen Texten.
- Bei den *epistemischen Autoritäten als Begründungsgrundlage* sind eher Nulleffekte typisch.
- Begründungen, welche über die Aktivitäten der *eigenen Person* erfolgen, weisen eher einen ungünstigen Zusammenhang mit dem Verstehen multipler Dokumente auf.

Damit lassen sich die vier Konstrukte profilieren hinsichtlich ihrer Zusammenhänge, die mehr oder weniger stark mit der Ausprägung der Verstehensleistungen einhergehen – und das in verschiedene Richtungen. Insgesamt zeichnet sich jedoch ab, dass die theoretisch begründbare Ausdifferenzierung empirisch eingeholt wird. Bei den Indikatoren des *Intertextmodells* ist die Forschungsbasis freilich zu spärlich und die Verteilung der Effekte auf die Zellen in der Tabelle 24 noch zu wenig dazu geeignet, klare Muster zu abstrahieren.

Tabelle 24: Überblick zum Zusammenhang den Facetten der Teildimension „Begründung des Wissens“ mit Produktmaßen aus Studien der Multiple-Documents-Literacy-Forschung

Ebene des Dokumentenmodells	Positive Effekte	Nulleffekte	Negative Effekte
Integriertes mentales Modell			
a) <i>Begründung (generisch)</i>	• Intertextuelle Inferenzen (Muis et al., 2015, Studien 1 und 2; Strømsø & Bråten, 2009) • Kritisches Denken in Diskurssynthesen (Muis et al., 2021) • Balanciertheit von Argumenten in offener Aufgabe (Kammerer et al., 2013)	• Intertextuelle Inferenzen (Bråten & Strømsø, 2010b; Gil et al., 2010b; Strømsø et al., 2008) • Anzahl Switches in Diskurssynthesen (Gil et al., 2010b) • Gesamttransformationsindex (Gil et al., 2010b)	
b) <i>Multiple Quellen</i>	• Intratextuelle Inferenzen: überzeugungsinkonsistent (Karimi & Richter, 2021a) • Abdeckung von Kausalfaktoren in Diskurssynthesen (Wiley et al., 2020, Experimente 2 und 3) bzw. von wichtigen Inhalten (Kammerer et al., 2021) • Explizite sprachliche Hinweise auf Corroboration in Diskurssynthesen (Wiley et al., 2020, Experimente 2 und 3) • Intertextuelles Integrieren in offenen Fragen (Bråten, Ferguson, Strømsø & Anmarkrud, 2013) bzw. argumentativen Diskurssynthesen (Bråten, Ferguson et al., 2014) • Entscheidung für eine objektiv angemessene Option (Kammerer et al., 2015)	• Intratextuelle Inferenzen: überzeugungskonsistent (Karimi & Richter, 2021a) • Intertextuelles Integrieren in offenen Fragen (Bråten, Anmarkrud et al., 2014; Strømsø et al., 2016, Majoritäten- und Minoritätensample) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domänen Geschichte und Physik) • Kombiniertes Maß als Indikator für integriertes mentales Modell (List, 2021) • Erinnerung an korrekte Argumente aus gelesenen Texten (Braasch, Bråten, Britt et al., 2014)	

Ebene des Dokumentenmodells	Positive Effekte	Nulleffekte	Negative Effekte
c) <i>Epistemische Autoritäten</i>	• Intertextuelles Integrieren in argumentativen Diskurs-synthesen (Bråten, Ferguson et al., 2014) bzw. in offenen Fragen (Strømsø et al., 2016, Minoritätensample)	• Intertextuelles Integrieren in offenen Fragen (Bråten, Ferguson, Strømsø & Anmarkrud, 2013; Strømsø et al., 2016, Majoritätensample) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domänen Geschichte und Physik) • Abdeckungsgrad von wichtigen Inhalten in offenen Aufgaben (Kammerer et al., 2021) • Erinnerung an korrekte Argumente aus gelesenen Texten (Braasch, Bråten, Britt et al., 2014)	
d) <i>Eigene Person</i>		• Intertextuelles Integrieren in argumentativen Diskurs-synthesen (Bråten, Ferguson et al., 2014) bzw. in offenen Fragen (Strømsø et al., 2016, Majoritätensample) • Entscheidung für eine objektiv angemessene Option (Kammerer et al., 2015) • Erinnerung an korrekte Argumente aus gelesenen Texten (Braasch, Bråten, Britt et al., 2014)	• Intertextuelles Integrieren in offenen Fragen (Bråten, Ferguson, Strømsø & Anmarkrud, 2013; Strømsø et al., 2016, Minoritätensample) • Gesamtscore Verstehen multipler Dokumente (Mahlow et al., 2022, Domänen Geschichte und Physik) • Abdeckungsgrad von wichtigen Inhalten in offenen Aufgaben (Kammerer et al., 2021)
Intertextmodell			
a) <i>Multiple Quellen</i>	• Anzahl von Referenzierungen in Diskurs-synthesen (Bråten, Ferguson et al., 2014; Wiley et al., 2020, Experiment 3) • Anzahl Quelle-Inhalt-Verbindungen in Diskurssynthesen (Bråten, Ferguson et al., 2014)	• Anzahl von Referenzierungen in Diskurs-synthesen (Wiley et al., 2020, Experiment 2) • Kombiniertes Maß als Indikator für Intertextmodell (List, 2021)	

Ebene des Dokumentenmodells	Positive Effekte	Nulleffekte	Negative Effekte
b) <i>Epistemische Autoritäten</i>	Anzahl expliziter Referenzierungen in Diskurssynthesen (Bråten, Ferguson et al., 2014)	Anzahl Quelle-Inhalt-Verbindungen in Diskurssynthesen (Bråten, Ferguson et al., 2014)	
c) <i>Eigene Person</i>		- Anzahl expliziter Referenzierungen in Diskurssynthesen (Bråten, Ferguson et al., 2014) - Anzahl Quelle-Inhalt-Verbindungen in Diskurssynthesen (Bråten, Ferguson et al., 2014)	

Die teils unterschiedlichen Zusammenhänge in den Studien mit den Indikatoren des integrierten mentalen Modells wirft die Frage auf, ob es zusätzliche Moderationseffekte gibt. Dafür gibt es Hinweise: Die bereits bei der Teildimension „Bestimmtheit des Wissens“ zur Sprache gekommene Studie von Karimi und Richter (2021a) hatte auch bei der Dimension „Begründung über multiple Quellen“ einen Moderationseffekt zum Ergebnis. Dieser betraf Inferenzleistungen bei den überzeugungsinkonsistenten Texten. Die Testpersonen unterschieden sich darin, wie gut sie überzeugungsinkonsistente Inhalte verstanden hatten: Je stärker sie der Auffassung waren, Wissen sei über multiple Quellen abzustützen, desto mehr überwandten sie den Effekt der eigenen Überzeugungen.

Zusammenfassend sind also verschiedene Facetten der Teildimension empirisch nachweisbar, deren Zusammenhänge mit Verstehensleistungen besonders deutlich – und theoriekonform in Bezug auf das Modell von Bråten, Britt et al. (2011) – beim integrierten mentalen Modell zutage treten. Wichtig ist, dass die Überzeugung, Aussagen müssten mit multiplen Dokumenten und Wissensquellen verifiziert werden, mit einem stärkeren intertextuellen Verstehen korrespondiert. Dies wirft die Frage auf, welche Prozesse dafür verantwortlich sind.

Perspektive 3: Prozesse

Der Zusammenhang von Prozessmaßen und epistemischen Überzeugungen ist bisher ausgesprochen selten erforscht worden. Die bisher vorgelegten Studien fokussieren auffällig die Teildimension „Begründung des Wissens“ und in ihr die Begründungen über multiple Quellen. Es lassen sich drei Ergebnislinien konturieren:

- *Die zeitliche Zuwendung zu Dokumenten als unspezifisches Maß für Prozesse bringt bislang nur annäherungsweise Erkenntnisse bei der Teildimension „multiple Quellen“, die bei genauerer Analyse aussagekräftiger wirken.* Die Studien zeichnen bezogen auf die reine Lesezeit ein unklares Bild: Wer stärker ausgeprägte Überzeugungen zur Begründung über multiple Quellen hatte, verbrachte mehr Zeit mit dem Lesen (Bråten, Anmarkrud et al., 2014), doch dieser Effekt wurde nicht durchgängig belegt (List, 2021). Die beiden genannten Studien verwendeten ähnliche Vorgehensweisen, sodass unklar ist, warum diese Differenz in den Resultaten zustande kommt. Besser erklärbar wird dieses diskrepante Ergebnis, wenn nicht die reine Zeit als unspezifisches Prozessmaß berücksichtigt wird, sondern genauer fokussiert wird, wie viel Zeit mit welchen Dokumenten Personen verbracht haben. Das haben Kammerer et al. (2015) in ihrer Studie mit einer simulierten Internet-Umgebung überprüft. Gegenstand war eine Recherche zur Krankheit Morbus Bechterew, zu der die Personen eine begründete Entscheidung hinsichtlich einer Therapie treffen sollten. Die fingierte Suchmaschinenumgebung erlaubte es, dank Logfiledaten genau nachzuvollziehen, welche Website wie lange konsultiert wurde. Das anwählbare Set von Dokumenten setzte sich zu gleichen Teilen aus kommerziellen, privaten und objektiven Internetdokumenten zusammen. Daneben wurden drei Konstrukte aus den epistemischen Überzeugungen erfasst: a) jene zur Begründung über multiple Quellen, b) jene zur eigenen Person (hier: kritisches Evaluieren aufgrund eigenen Wissens) und schließlich c) „generelle internetbezogene Überzeugungen“, die das Internet als Wissensquelle betreffen und Korrektheit sowie Detailliertheit von Informationen vor allem aus Expert:innen-Sicht zum Gegenstand der Begründung machen. Regressionsanalysen mit allen drei Überzeugungen als unabhängige Variablen und der verbrachten Zeit auf den drei Arten von Websites als abhängige Variablen ergaben differenzielle Effekte: Stärker ausgeprägte allgemeine Internet-Überzeugungen zum einen und zur Begründung über multiple Quellen zum anderen gingen mit längeren Verweildauern auf den objektiven Websites und geringeren Lesezeiten bei den subjektiven Websites einher. Zudem lasen Personen umso kürzer die kommerziellen Websites, je mehr sie der Begründung über mehrere Quellen zugestimmt hatten. Genau das gegenteilige Muster ergab sich bei den kommerziellen Websites und der Begründung über die eigene Person: Hier lasen die Teilnehmerinnen diese Websites umso länger, je stärker sie dieser Teildimension zugestimmt hatten. Damit spiegeln diese Befunde wider, dass Personen rein von der Zuwendungszeit zu verschiedenen belastbaren Websites das am günstigsten wirkende Muster aufwiesen, wenn sie Wissensaussagen für abstützungsbedürftig mit mehreren Quellen hielten.

- *Es gibt unterschiedliche Muster beim Zusammenhang von spezifischeren Prozessmaßen und epistemischen Überzeugungen.* Gleich zwei Studien fanden in retrospektiv geschilderten Vorgehensweisen zum intertextuellen Integrieren pfadanalytische positive Vorhersagen der Überzeugung zur Begründung des Wissens über multiple Quellen (Bråten, Anmarkrud et al., 2014; List, 2021). Dies verweist auf eine Übereinstimmung von Überzeugungen und berichteten Strategien, deren Schnittmenge das textvergleichende bzw. den Aufbau eines integrierten mentalen Modells abzielende Lesen darstellt. Anders gestaltet sich hingegen der Befund von Kammerer et al. (2021). In dieser Untersuchung wurden drei Facetten aus der Teildimension „Begründung des Wissens“ erfasst: a) multiple Quellen, b) epistemische Autoritäten und c) eigene Person. Als Prozessvariablen wurden aus Lautdenk-Protokollen die Äußerungshäufigkeiten zum Sourcing sowie zum textvergleichenden Lesen (Corroboration) extrahiert. In Regressionsanalysen wurden diese Frequenzen der Prozesse als abhängige Variablen genutzt, als unabhängige Variablen kamen neben den drei Facetten der epistemischen Überzeugungen noch das Vorwissen hinzu. Die Ergebnisse ergaben keine Effekte der Überzeugungen zu den multiplen Quellen, dafür aber spezifische Varianzaufklärungen der anderen beiden Überzeugungen. Beim Sourcing erwies sich die Überzeugung, Wissen stamme von Expertinnen und Experten – also epistemischen Autoritäten –, als positiv prädiktiv. Bei dem Prozess Corroboration war hingegen die Überzeugung, Wissen müsse von der eigenen Person geprüft werden, ein negativer Prädiktor: Je stärker Personen dieser Überzeugung waren, desto weniger Corroboration-Aktivitäten waren nachweisbar. Damit zeichnet diese Studie ein deutlich anderes Bild als die anderen beiden, die mit einem anderen Design operierten. Gleichwohl ist ein allgemeines Muster abstrahierbar: Die Prozesse des textvergleichenden Lesens bzw. der Bemühungen, intertextuelle Bezüge aktiv herzustellen, hängen mit den epistemischen Überzeugungen zusammen.
- *Die Zusammenhänge von epistemischen Überzeugungen, Prozessen und Produkten sind bislang wenig untersucht und stellen sich aktuell empirisch widersprüchlich dar.* In zwei Studien (Bråten, Anmarkrud et al., 2014; List, 2021) wurden Mediationsanalysen in Pfadmodellen vorgelegt; beide Studien fokussierten auf die Begründung des Wissens über multiple Quellen. Beide Studien verwendeten unterschiedliche Indikatoren für das intertextuelle Verstehen, und beide Studien kommen zu unterschiedlichen Befunden. Die erste Untersuchung (Bråten, Anmarkrud et al., 2014) ermittelte einen Mediationseffekt: Die positiven Effekte der Überzeugungen wurden über die aufgewandte Zeit für das Lesen zum einen und die selbstberichteten intertextuellen Integrationsstrategien vermittelt. Es gibt also keine direkten Zusammenhänge, stattdessen sind unspezifische und spezifische Prozessmaße Mediatoren. Diesen Zusammenhang konnte die zweite Studie nicht replizieren

(List, 2021). Nicht nur bestanden keine Zusammenhänge zwischen Lesezeit und Überzeugungen, zudem waren die durchaus vorhandenen Effekte der intertextuellen Strategien nicht statistisch absicherbare Mediatoren der Überzeugungen. Diese unklaren Zusammenhänge sind auch an anderen Stellen der Forschung, nämlich beim selbstregulierten Lernen beobachtet und werfen die Frage nach dem Status der epistemischen Überzeugungen als Einflussquelle für ein lernstrategisches Vorgehen auf (s. Teilkap. [6.3.2](#)).

Bei den epistemischen Überzeugungen zur Begründung des Wissens sind zwar die meisten Studien mit Blick auf Prozessdaten vorgelegt worden. Gleichzeitig ist die Forschungslage eher heterogen. Es existieren zwar Zusammenhänge, aber diese lassen sich nicht durchgängig nachweisen. Insbesondere die Studien mit unspezifischen Prozessmaßen verweisen auf die Notwendigkeit, die Bezugspunkte des Lesens genauer in den Blick zu nehmen. Außerdem wirkt es nötig, allgemein mehr als eine Dimension der epistemischen Überzeugungen zu fokussieren, da mitunter spezifische Effekte zu erwarten und teils bereits zu beobachten sind.

Teildimension 4: Ursprung des Wissens (Produktperspektive)

Die Teildimension „Ursprung des Wissens“ ist eine in der Forschung zum Verstehen multipler Dokumente vernachlässigte Dimension. Das mag sich daraus erklären, dass sich durch die Ausdifferenzierung in der Teildimension „Begründung des Wissens“ Überschneidungen ergeben haben, welche Expertinnen und Experten sowie die eigene Person als Quelle des Wissens als eigenständige Kontinuen modellieren und nicht als Endpunkte auf einem einzigen Kontinuum. Gleichwohl liegen vereinzelte Studien vor, die mit der Teildimension „Ursprung des Wissens“ operierten. Wie aus Tabelle 25 hervorgeht, erbrachten diese Studien keine klaren Hinweise zu positiven oder negativen, sondern überwiegend zu keinerlei statistisch signifikanten Effekten für Merkmale des integrierten mentalen Modells. Zudem sind Zusammenhänge mit den Prozessen und die Effekte auf das Intertextmodell mangels Forschung unklar.

Tabelle 25: Überblick zum Zusammenhang der Teildimension „Ursprung des Wissens“ mit Produktmaßen aus Studien der Multiple-Documents-Literacy-Forschung

Ebene des Dokumentenmodells	Positive Effekte	Nulleffekte	Negative Effekte
Integriertes mentales Modell	• Intertextuelle Inferenzen (Muis et al., 2015, Studie 1)	• Intertextuelle Inferenzen (Gil et al., 2010b; Muis et al., 2015, Studie 2; Strømsø et al., 2008) • Anzahl Switches in Diskurssynthesen (Gil et al., 2010b) • Gesamttransformationsindex (Gil et al., 2010b) • Kritisches Denken in Diskurssynthesen (Muis et al., 2021)	
Intertextmodell			

Personenbezogene Unterschiede

Der bisherigen Darstellung lag die Logik zugrunde, die Konstrukte bzw. Variablen einzeln und in ihrem Zusammenhang mit Leistungs- und auf weit geringerer Datenbasis mit Prozessdaten zu betrachten. Eine solche variablenzentrierte Darstellung bildet jedoch nur einen Ausschnitt. Ein anderes Vorgehen besteht darin, dass Personengruppen mit verschiedenen Ausprägungen der epistemischen Überzeugungen voneinander unterschieden werden. Dieser personenzentrierte Ansatz basiert darauf, dass erstens (Extrem-)Gruppenvergleiche auf der Basis von Mediansplits und darauffolgender Gruppenbildung (Bråten et al., 2008) bzw. (sehr) hohen oder (sehr) niedrigen Werten auf statistisch eindimensional modellierten, inhaltlich aber mehrere theoretische Konstrukte abdeckenden Skalen (Bråten & Strømsø, 2006a; Mason et al., 2010a) erfolgen. Eine zweite Möglichkeit, Gruppen zu vergleichen, setzt darauf, dass über Clusteranalysen Personengruppen ermittelt werden, die sich in der Ausprägung individueller Merkmale empirisch ähneln und zugleich von anderen Gruppen unterscheiden. Diese Clusteranalysen werden oftmals mit mehreren Konstrukten durchgeführt. Als Grundlage fungierten sowohl epistemische Überzeugungen als Personenmerkmale (Ferguson & Bråten, 2013; Trevors et al., 2017) als auch verhaltensbasierte Indikatoren auf der Basis von Lautdenk-Protokollen (Mason et al., 2010b, 2011) und retrospektiven Interviews (Mason et al., 2010a). Einen Sonderfall bildet schließlich, drittens, ein Vorgehen mit Ex-post-Daten aus offenen Fragen, bei denen Personengruppen nach ihren epistemischen Überzeugungen gemäß der Antworten ordinal gereiht wurden und die höchstmögliche Ausprägung in den Kodierungen die Gruppenzugehörigkeit bestimmte (List, Oaxaca et al., 2021).

Bei den Verstehensleistungen waren folgende Kombinationen aus dem ersten Ansatz mit dem Vergleich von Extremgruppen günstig:

- hohe Ausprägungen bei der „Einfachheit des Wissens“ zugunsten der Komplexität von Wissen in Verbindung mit hohen Ausprägungen auf der Skala „Quelle des Wissens“ zugunsten der Expertinnen und Experten als Quelle, denn in dieser Merkmalskombination erzielten die Personen bessere Ergebnisse im intertextuellen Inferenzverifikationstest (Bråten et al., 2008);
- hohe Ausprägungen in einer Gesamtskala mit vier epistemischen Überzeugungen, wobei der Lernerfolg mit offenen Fragen nach einer Internet-Recherche erfasst wurde (Mason et al., 2010a);
- hohe Ausprägungen auf einer Gesamtskala zu allgemeinen epistemischen Überzeugungen, die neben den beiden Teildimensionen zu den Merkmalen des Wissens auch noch weitere Konstrukte enthält, die gemeinhin nicht als epistemische Überzeugungen gelten (Bråten & Strømsø, 2006a, Test: intertextueller Inferenzverifikationstest) – in dieser Studie wurden zudem praktisch bedeutsame, aber aufgrund des kleinen Samples statistisch nicht-signifikante Gruppendifferenzen bei den retrospektiv berichteten Prozessen ermittelt: Personen mit stärker ausgeprägten Überzeugungen elaborierten Inhalte mehr und waren metakognitiv aktiver.

Damit wirkt es im Licht dieser Studien so, als gehe insbesondere die Überzeugung, Wissen sei komplex statt einfach, mit besseren Inferenzleistungen auf der Basis mehrerer gelesener Texte einher. Ein weiteres Muster ist, dass normativ positiv zu wertende Ausprägungen in den jeweiligen einzelnen Dimensionen mit höheren Leistungen korrespondieren.

Der zweite Forschungszweig der Personenvergleiche kommt zu Ergebnissen, welche insbesondere die Bedeutung der Teildimension „Begründung des Wissens“ unterstreichen. Die Ergebnisse sind im Einzelnen:

- *Auch wenn sich Personengruppen unterscheiden lassen, scheinen einzelne Teildimensionen besonders wichtig zu sein, vor allem zur Begründung des Wissens.* Dieses Muster bezieht sich in gleich drei Studien auf die Begründung des Wissens. Hohe Verstehensleistungen erzielte in den Studien von Mason et al. (2010b, 2011) jene Gruppe von Personen, in der von den vier Dimensionen epistemischer Überzeugungen besonders stark diejenige zur Begründung ausgeprägt war. Dieses Muster unterfütterte die Analyse von Mason et al. (2010a). In dieser Studie waren nicht die gefundenen Cluster signifikant mit dem Lernerfolg nach einer Internet-Recherche verbunden, sondern alleinig die Ausprägung, wie elaboriert die Sekundarschuljugendlichen die Auswahl von Websites mit Begründungen des Wissens untermauert hatten. Diesen allgemeinen Befund differenzieren Trevors et al. (2017) aus. Sie nutzten

das trichotome Instrumentarium mit den drei Skalen zur Begründung des Wissens (a) multiple Quellen, b) epistemische Autoritäten und c) eigene Person). Sie ermittelten drei Cluster, und der entscheidende Befund bestand darin, dass dasjenige Cluster die besten Verstehensleistungen bei einem Transfertest erbracht hatte, welches die am geringsten ausgeprägten Überzeugungen zur eigenen Person aufwies.

- *Epistemische Überzeugungen zur Begründung des Wissens und Vorwissen in ihrer Kombination korrespondieren mit Verstehensunterschieden.* Anders als in anderen Studien stützte sich eine norwegische Studie nicht nur auf die drei verschiedenen Facetten der Begründungen des Wissens, sondern nutzte für die Clusteranalyse zusätzlich das Vorwissen (Ferguson & Bråten, 2013). Das Leseverstehen, welches mit offenen Fragen erfasst wurde, war in den drei vorgefundenen Clustern unterschiedlich ausgeprägt. Am geringsten war es in jener Gruppe, welche sich vor allem durch ein geringes Vorwissen und stark ausgeprägte Überzeugungen zur eigenen Person als Wissensbegründung auszeichnete. Die beiden anderen Gruppen mit höherem Vorwissen und höheren Werten bei den multiplen Quellen und epistemischen Autoritäten als Begründungsgrundlage für Wissen hatten erheblich bessere Verstehensleistungen erzielt. Das bedeutet: Die Gruppenzugehörigkeit allein auf der Basis der epistemischen Überzeugung kann ergänzt werden um weitere Indikatoren der internalen Ressourcen, welche dabei helfen, Personengruppen zu profilieren.
- *Personen lassen sich in ihren epistemischen Überzeugungen speziell zum Lernen mit multiplen Dokumenten unterscheiden und weisen dabei Eigenheiten in ihrem Vorgehen und in ihrem Verstehen auf.* So lassen sich mittels offener Fragen am Ende von komplexen Aufgaben und damit ex post Gruppen von Personen mit spezifischen epistemischen Überzeugungen bilden, die sich in ihren Vorgehensweisen und in ihren Produktmaßen analysieren lassen. Diese Forschungslinie ist noch sehr jung, aber die einzige hier anführbare Studie von List, Oaxaca et al. (2021) ist so aufschlussreich, dass ihr der Kasten „Illustrierendes Beispiel 19“ gewidmet ist. In dieser Studie erwies es sich partiell von Vorteil, wenn Personen eine Überzeugung hatten, die sich als Variante der Begründung des Wissens über das (kritische) Nutzen multipler Quellen bezeichnen lässt.

Illustrierendes Beispiel 19

Unterscheiden sich Personen mit unterschiedlichen epistemischen Überzeugungen darin, wie sie multiple digitale Dokumente nutzen?

A) Was war das Erkenntnisinteresse?

Die Studie widmete sich der Frage, wie Prozesse und Produkte des intertextuellen Integrierens sich in Gruppen mit verschiedenen Überzeugungen zum Lernen mit multiplen Dokumenten unterscheiden (List, Oaxaca et al., 2021).

B) Wer wurde untersucht?

Die Teilnehmerinnen und Teilnehmer der Studie waren 72, im Durchschnitt 19 Jahre alte Studierende aus den USA. Fünf Sechstel waren Frauen, zu sieben Achteln bestand das Sample aus europäischen Amerikanerinnen und Amerikanern. Die Studierenden stammten aus bildungswissenschaftlichen Studiengängen, die Mehrheit der Testpersonen befand sich im ersten Jahr ihres Studiums.

C) Welche Instrumente kamen zum Einsatz?

In der Untersuchung wurden Prozesse sowie leistungsbezogene und dispositionale Merkmale erfasst. Die *Überzeugungen* als Basis der Gruppenbildung wurden am Schluss mittels einer offenen Frage ermittelt, nachdem die Studierenden die Aufgabe absolviert hatten. Gegenstand der Frage war eine Begründung der Frage, warum es wichtig sei, mittels mehrerer Dokumente zu lernen. Die Antworten wurden inhaltsanalytisch kodiert und 70 Personen auf dieser Basis einer von vier Gruppen zugewiesen: a) hohe Informationsmenge erreichen (möglichst viele Informationen sammeln bzw. abdecken) (3 Prozent), b) persönliches Verstehen eines Sachverhaltes aufbauen (25 Prozent), c) Mehrperspektivität durch Informationen aus verschiedenen Standpunkten erkennen (32 Prozent) und d) evaluierendes Vergleichen durch dokumentenübergreifendes Prüfen auf Verzerrungen und Vertrauenswürdigkeit (38 Prozent).

Die Prozesse in einer digitalen Bibliothek mit fünf schriftlichen Dokumenten und einem Video nebst anwählbaren Zusatzinformationen wurden mittels Logdaten erfasst. Dadurch wurden Lesezeiten, Anzahl der aufgerufenen Dokumente sowie Anzahl der Zusatzinformationen pro Dokument nicht-reaktiv erhoben.

Die *Produktmaße* bezogen sich auf eine Diskurssynthese, die einen Forschungsbericht als Entscheidungsgrundlage für Politikerinnen und Politiker zum Mindestlohn als Gegenstand erforderte. Diese Textprodukte wurden mehrdimensional ausgewertet. Erstens wurden sie entlang der Typologie zu Dokumentenmodellen klassifiziert. Zweitens wurden intertextuelle Integrationen gezählt.

D) Was waren die Hauptbefunde?

Mittels Varianzanalysen wurden die metrischen Daten und mittels Chi-Quadrat-Tests die Verteilungen der Produktmaße gemäß Dokumentenmodell-Typologie überprüft, die unabhängige Variable bildete die Zugehörigkeit zu einer der vier Gruppen von epistemischen Überzeugungen. Das zentrale Ergebnismuster bestand darin, dass die Gruppe der Studierenden aus der Gruppe des evaluierenden Vergleichens in vielen, aber nicht sämtliche Variablen überlegen waren. Dieser Vorteil, und das ist ein zweites Befundmuster, manifestierte sich im Kontrast zur Gruppe, welche die Mehrperspektivität als bedeutend für das Lernen aus multiplen Dokumenten angeführt hatte.

Bei den *Prozessmaßen* zeigten sich diese Vorteile resp. Nachteile in einem Maß aus den insgesamt drei möglichen: der Zahl zusätzlich konsultierter Informationen, die optional anwählbar waren. Bei den *Produktmaßen* betraf dies alle beiden Indikatoren. Zum einen unterschieden sich die beiden Gruppen bei der Anzahl der intertextuellen Inferenzen

nahezu um den Faktor 3. Zum anderen waren die Diskurssynthesen in zwei der möglichen Ausprägungen der Dokumentenmodelle über- bzw. unterrepräsentiert. Testpersonen, deren epistemische Überzeugungen die Multiperspektivität fokussierten, hatten häufiger Forschungsberichte verfasst, in denen die Inhalte der Dokumente mit Metadaten versehen waren und blockweise dargestellt waren (separate Repräsentationen mit Zitationen; 40 vs. 15 Prozent). Prononcierter war die Differenz zwischen den beiden Gruppen bei den Dokumentenmodellen als höchstmöglich zu erreichender Kategorie. Studierende der Gruppe mit der epistemischen Überzeugung, dass Lernen aus multiplen Dokumenten wegen der kritischen Evaluation wichtig sei, hatten überwiegend Dokumentenmodelle aufgebaut (56 vs. 7 Prozent).

E) Diskussion

Personen lassen sich danach unterscheiden, welche epistemischen Überzeugungen sie zum Lernen mit multiplen Dokumenten haben. Auch wenn dies nur mit einer einzelnen Frage in der Studie erfolgte, erlauben die inhaltsanalytisch ausgewerteten Antworten eine Typologie mit vier Ausprägungen. Viele der auf Differenzen in Prozess- und Produktmaßen abzielenden Analysen erbrachten keine Differenzen zwischen allen vier Gruppen. Gleichwohl lassen sich zwei Gruppen mit prima vista ähnlich wirkenden Überzeugungen konturierend gegenüberstellen. Personen, welche die Mehrperspektivität betonten, neigen zu einer eher oberflächenorientierten Verarbeitung, welche sich in weniger stark integrierten Produkten widerspiegelt. Demgegenüber haben Personen mit einer eher tiefenorientierten Prüfung verschiedener Dokumente eher mehr Informationen konsultiert, mehr Inferenzen in den Produkten gebildet und eine stärker integrierte Darstellung der Thematik realisiert. Tiefergehende Analysen der eigentlichen Prozesse mit anderen Verfahren und Zugängen würden dabei helfen, die logfilebasierten Auswertungen anzureichern.

Zusammenfassung

Die empirische Forschung zu den epistemischen Überzeugungen als einer potenziellen Quelle interindividueller Unterschiede in den dominant untersuchten Produkten (des integrierten mentalen Modells) und teils den Prozessen im Umgang mit multiplen Dokumenten hat sich in den vergangenen Jahren stark verändert. Die ursprünglich vier Teildimensionen umfassende Trennung wurde überwiegend aufrechterhalten, aber auch in einer Teildimension, nämlich bei der Begründung des Wissens, ausdifferenziert. Die Forschungsstände und die Eindeutigkeit divergieren je nach Dimension, übergreifen lässt sich dies festhalten:

- Personen, welche Wissen für komplex halten, erbringen bessere Verstehensleistungen.
- Dasselbe gilt für Personen, die bei der Begründung des Wissens eher auf epistemische Autoritäten wie Expertinnen und Experten vertrauen bzw. Wissen fürs abstützungsbedürftig aus verschiedenen Quellen halten. Dieses Muster

ist sowohl in variablen- als auch personenzentrierten methodischen Zugängen gut abgestützt. Als deutlich weniger günstig gilt es, wenn Personen – zumal bei komplexen (naturwissenschaftlichen) Themen – allzu sehr sich selbst als Richtmaß für die Entscheidung von Wissensaussagen erachten.

- Bei der Bestimmtheit des Wissens und dem Ursprung des Wissens ist auffällig, dass sie gemäß der bisherigen Forschung eher unverbunden mit den Leistungen waren.

Damit wirkt es so, als wären die beiden Facetten „Einfachheit des Wissens“ und die unterteilbare „Begründung des Wissens“ empirisch besonders gut geeignet, interindividuelle Unterschiede im bei intertextuellen Integrationsleistungen auf der Basis multipler Dokumente zu erklären.

4.3.3.2 Zugang 2: epistemische Überzeugungen in der Tradition von Entwicklungsmodellen

Wie schon im Teilkapitel [4.3.3.1](#) wird die nachstehende Darstellung von einer theoretischen Basis eröffnet, die dabei hilft, die taxonomischen Zugänge in einer Entwicklungslogik fortzuführen. Das Drei-Phasen-Modell, das zu Beginn dargestellt wird, bildet einen Rahmen, der die Ergebnisse verortbar werden lässt. Die Ergebnisse werden sodann zweigeteilt präsentiert, nämlich zunächst aus einer Produkt- und danach in einer Prozessperspektive.

Ein Drei-Phasen-Modell der Entwicklung epistemischer Überzeugungen

Der zweite Zugang zu den epistemischen Überzeugungen beruht auf einer Entwicklungsperspektive, bei der angenommen wird, dass es zu substanziellen Veränderungen dazu kommt, wie Personen Wissen subjektiv erklären. Phasenmodelle unterscheiden sich in variierenden Detaillierungsgraden und unterschiedlicher Zahl der Phasen darin, welche markanten Wechsel sich ergeben (Hofer & Pintrich, 1997). Diese Veränderungen betreffen zum einen das Zusammenspiel von einzelnen Überzeugungen und zum anderen die Ausprägung bei den Teildimensionen epistemischer und ontologischer Überzeugungen. Diese geänderte Nomenklatur geht auf Greene, Azevedo und Torney-Purta (2008) und ihr philosophisch angereiches Entwicklungsphasenmodell bzw. ihre Modellsynopse zurück.

Das Plädoyer von Greene, Azevedo und Torney-Purta (2008) besteht darin, genauer zu differenzieren, wie sich „ontologische Überzeugungen“ und epistemische Überzeugungen in einem (domänenspezifischen) Makrophasen-Modell entwickeln (s. Tabelle 26). Mit den ontologischen Überzeugungen sind sich verändernde Überzeugungen zu den Merkmalen des Wissens gemeint, die sich auf die Teildimensionen Einfachheit und Bestimmtheit aus der Hauptdimension „Merkmale des Wissens“ beziehen (s. Teilkap. [4.3.3.1](#)) und in dem Makrophasen-Modell

zu einem einzigen Konstrukt kombiniert werden. Der entscheidende Gedanke ist, dass Personen im Lauf ihrer Entwicklung idealtypisch eine Veränderung vollziehen, indem sie Komplexität und Veränderlichkeit von Wissen (an-)erkennen. Dieser Wechsel vollzieht sich im Übergang von der ersten zur zweiten Phase, indem Personen aus einem „*Realismus*“ genanntem Stadium austreten. Kennzeichnend ist hierfür, dass Personen Wissen nicht mehr für eine Kopie der Realität halten. Dies ist in der ersten Phase noch der Fall, deren wesentliches Merkmal ist, dass Wissen als solches nicht strittig zu sein scheint, sondern von unverrückbaren Fakten abhängt. Entsprechend sind Begründungen des Wissens als epistemische Überzeugungen in dieser Phase schlichtweg für das Individuum nicht nötig. Es hält sowohl die eigene Person als auch epistemische Autoritäten für belastbar.

Tabelle 26: Darstellung eines allgemeinen Entwicklungsmodells der epistemischen Überzeugungen

Vergleichsdimension	1) Realismus	2) Dogmatismus/ Skeptizismus	3) Rationalismus
A) Allgemeine Informationen			
Ontologische Überzeugungen (Natur des Wissens, Kombination von Bestimmtheit und Einfachheit des Wissens)	- Wissen als stabile Akkumulation von Konzepten - Individuelle Ausprägungen Richtung Pole Konstanz bzw. Einfachheit	- Wissen als komplex und veränderlich - Individuelle Ausprägungen Richtung Pole Veränderung und Komplexität	- Wissen als komplex und veränderlich - Individuelle Ausprägungen Richtung Pole Veränderung und Komplexität
Epistemische Überzeugungen (Genese und vor allem Begründung des Wissens)	- Wissen weder subjektiv noch von anderen notwendigerweise begründungsnotwendig - Hohe individuelle Ausprägung bei beiden Überzeugungen eigene Person und epistemischen Autoritäten	- Wissen stark personenabhängig, sei es an äußere Quellen gebunden (<i>Dogmatismus</i>) oder hochgradig subjektiv (<i>Skeptizismus</i>) <i>Dogmatismus</i>: hohe Ausprägungen bei epistemischen Autoritäten und gleichzeitig geringe Ausprägungen bei der eigenen Person <i>Skeptizismus</i>: hohe Ausprägungen bei der eigenen Person und gleichzeitig geringe Ausprägungen bei epistemischen Autoritäten	- Abwägung und Prüfen von verschiedenen Begründungen, ehe Aussagen als Wissen akzeptiert werden - Mittlere Ausprägungen bei verschiedenen Facetten der Begründungen, diese im Zusammenspiel wichtig

Vergleichs- dimension	1) Realismus	2) Dogmatismus/ Skeptizismus	3) Rationalismus
B) Empirische Belege aus der Multiple-Documents-Literacy-Forschung*			
1) Personen- zentrierte Auswertungen		• Geringere Fähigkeit zur kritischen Integration von konfligierenden Websites (Barzilai & Zohar, 2012; Gruppe Absolutismus) • Eher einseitige Zusammenfassungen mit My-Side-Bias zu zwei Texten (Iordanou et al., 2020; Gruppen Absolutismus und Multiplismus) • Geringeres Leseverstehen (Ioannou & Iordanou, 2020, Studie 2; Gruppe Absolutismus)	• Höhere Fähigkeit zur kritischen Integration von konfligierenden Websites (Barzilai & Zohar, 2012; Gruppe Evaluatismus) • Eher zweiseitige Zusammenfassung mit überwundenem My-Side-Bias zu zwei Texten (Iordanou et al., 2020; Gruppe Evaluatismus) • Höheres Leseverstehen (Ioannou & Iordanou, 2020, Studie 2; Gruppe Evaluatismus)
2) Variablen- zentrierte Auswertungen		• Negative Effekte auf Verstehen der Perspektiven multipler Autor:innen (Barzilai & Eshet-Alkalai, 2015; Skalen Absolutismus und Multiplismus) • Keine Effekte auf thematische Integrationen in drei kurzen offenen Fragen (Barzilai & Ka'adan, 2017; Skalen Absolutismus und Multiplismus) • Keine Effekte auf Integrationsmaße bei schriftlichen Argumentationen (Barzilai & Eshet-Alkalai, 2015; Skalen Absolutismus und Multiplismus) bzw. keine Zusammenhänge mit drei verschiedenen Produktindikatoren des Integrierens in Argumentationen (Barzilai et al., 2015; Skalen Absolutismus und Multiplismus)	• Positive Effekte auf Verstehen der Perspektiven multipler Autor:innen (Barzilai & Eshet-Alkalai, 2015; Skala Evaluatismus) • Positive Effekte auf thematische Integrationen in drei kurzen offenen Fragen (Barzilai & Ka'adan, 2017; Skala Evaluatismus) • Keine Effekte auf Integrationsmaße bei schriftlichen Argumentationen (Barzilai & Eshet-Alkalai, 2015; Skala Evaluatismus) bzw. keine Zusammenhänge mit drei verschiedenen Produktindikatoren des Integrierens in Argumentationen (Barzilai et al., 2015; Skala Evaluatismus)

(Quelle: eigene Darstellung, im oberen Teil basierend auf Greene, Azevedo & Torney-Purta, 2008, S. 152–155;

* Legende: Die Typologie von Greene, Azevedo & Torney-Purta, 2008, unterscheidet sich in einigen Aspekten hinsichtlich jener in der empirischen Forschung dominanten Vier-Phasen-Modellierung von Kuhn & Weinstock, 2002 (Realismus, Absolutismus, Multiplismus, Evaluatismus) sodass der Dogmatismus sensu Greene, Azevedo & Torney-Purta, 2008, nicht direkt gleichsetzbar ist mit dem Absolutismus gemäß Kuhn & Weinstock, 2002, und Multiplismus sensu Greene, Azevedo & Torney-Purta, 2008, nicht direkt gleichsetzbar ist mit dem Skeptizismus gemäß Kuhn & Weinstock, 2002)

Die ontologischen Überzeugungen unterliegen einer krisenhaften Veränderung beim Übergang in die nächste Phase, in der Individuen erkennen, dass es legitime und legitimierbare Sichtweisen auf das Wissen gibt. Wissen ist damit nicht mehr absolut und einfach, sondern Gegenstand der Prüfung. In der zweiten Phase sieht das Modell zwei Muster bei der Ausprägung epistemischer Überzeugungen vor, die stark gegenläufig sind. Auf der einen Seite ist der „*Dogmatismus*“ anzuführen, dass die ontologisch begründbare Unsicherheit des Wissens alleinig über externe Quellen, nämlich die epistemischen Autoritäten, absicherbar wirkt. Auf der anderen Seite – und genau das Gegenteil davon – ist der „*Skeptizismus*“ eine Variante des Umgangs. Hier wird Wissen für eine zuvorderst subjektiv abstützbare Größe gehalten, sodass Personen epistemischen Autoritäten eher misstrauen.

In der letzten Phase, die nicht alle Personen erreichen und welche dezidiert erst für Erwachsene erreichbar wirkt, findet eine Entwicklung statt, in der verschiedene Begründungen zulässig, aber auch überprüfungsnotwendig wirken. Wissensbezogene Aussagen sind für Personen in der Phase des „*Rationalismus*“ erst akzeptabel, wenn sie abstützbar sind. Diese Abstützbarkeit kann sich auf verschiedene Begründungen beziehen, darum gibt es keine universalen und exklusiven Ansprüche auf den Geltungsanspruch und auf die Begründbarkeit des Wissens.

Das skizzierte Modell nimmt damit die Trennung der beiden Hauptdimensionen der epistemischen Überzeugungen mit teils geänderter Nomenklatur auf und beschreibt, wie sich Überzeugungen zu Merkmalen des Wissens (ontologische Überzeugungen) und dessen Begründung (epistemische Überzeugungen) ausdifferenzieren. Die grundsätzliche Entwicklungsdynamik ist von einer höheren Komplexität geprägt, die aber je nach Phase unterschiedlich stark ausfällt und bei den epistemischen Überzeugungen nach einer eher apodiktischen Auffassung in der zweiten Phase elaborierteren Mischverhältnissen weicht. Die Dynamik geht von einem rigiden und simplifizierenden ontologisch geprägten Set von Überzeugungen in Phase 1 aus. Diese Rigidität ist auch kennzeichnend für Phase 2, betrifft nun aber Extrempole bei den epistemischen Überzeugungen, während die ontologischen Überzeugungen bereits differenzierter sind. Diese Entwicklung hin zu elaborierteren Überzeugungen in puncto der epistemischen Überzeugungen wird erst in Phase 3 realisiert.

Produkte

Bevor die Ergebnisse, die für die Produkte auf breiterer Datenbasis vorliegen, gebündelt werden, ist es noch nötig, auf die Differenz in der Operationalisierung der epistemischen Überzeugungen hinzuweisen. Einige der Studien ordneten Personen gemäß ihren Testleistungen, z. B. dem metakognitiven Wissen (Barzilai & Zohar, 2012; Ioannou & Iordanou, 2020, Studie 2; Iordanou et al., 2020) einer Gruppe zu, welche einer Phase gemäß dem Entwicklungsmodell von Kuhn und

Weinstock (2002) entsprach. Dies folgt einer *personenzentrierten Perspektive*. Andere Forschungsteams verwendeten eigene Skalen für die drei Ausprägungen bei den Phasen und wählten damit einen *variablenzentrierten Zugang* (Barzilai & Eshet-Alkalai, 2015; Barzilai & Ka'adan, 2017; Barzilai et al., 2015). Diese Differenz nimmt die Darstellung der Ergebnisse im unteren Teil (B) der Tabelle 26 gezielt auf, denn es gibt Unterschiede in der Art, wie die Phasenmodell-Logik operationalisiert wurde:

- *Bei den personenzentrierten Auswertungen zeigen sich konsistente Vorsprünge in Produktmaßen auf der Basis konfligierender Dokumente für Mitglieder aus der Gruppe „Evaluatismus“.* Die Zahl der Studien ist gering, doch eint sie im Design, dass sich die zu lesenden Dokumente als Grundlage der Produktmaße einander inhaltlich widersprachen. Personen, die aufgrund ihrer offenen Antworten der jeweils höchsten Stufe im Entwicklungsmodell zugeordnet wurden (Evaluatismus), erbrachten in den verschiedenen Produktmaßen bessere Leistungen. Bemerkenswert ist: Personen aus der Gruppe „Multiplismus“ sind in diesen Studien teils so marginal vorgekommen, dass sie nicht ausgewertet wurden.
- *Bei den variablenzentrierten Auswertungen gibt es nur tendenzielle und nicht durchgängige positive Effekte der Skala „Evaluatismus“.* Die Personen verstehen zwar tendenziell Perspektiven besser und können auch kurze Fragen beantworten, aber die Leistungen insbesondere beim schriftlichen Argumentieren werden von den epistemischen Überzeugungen nicht tangiert.

Mögen sich die Ergebnisse in ihrer Eindeutigkeit je nach personen- oder variablenzentrierten Analysefokus anders darstellen, so gibt es doch eine Konvergenz der Ergebnisse. Sie besteht darin, dass erst bei stärker ausgeprägter Fähigkeit, aus einer evaluatistischen Warte Wissen anzuerkennen, auch bessere Integrationsleistungen erbracht werden. Dies steht in Einklang mit den Aussagen aus dem Drei-Phasen-Modell von Greene, Azevedo und Torney-Purta (2008), nach welchem das Verstehen konfligierender und komplexer Themen dann erwartbar wirkt, wenn sowohl ontologische als auch epistemische Überzeugungen Komplexität anerkennen. Gleichwohl ist auffällig, dass ausgerechnet beim Argumentieren diese Zusammenhänge bisher nicht beobachtet wurden.

Prozesse

Die Prozesse sind bislang ein Desiderat, aber immerhin drei Studien liegen vor, zwei von ihnen basieren auf personenzentrierten Auswertungen. Die Ergebnisse sind:

- *Personen, die gemäß Phasenmodellen weiter fortgeschritten sind, zeigen überlegene lesestrategische Vorgehensweisen bei Texten zu gesellschaftswissenschaftlichen Themen.* So ergab ein Gruppenvergleich mit Sechstklässlern (Ioannou & Iordanou, 2020, Studie 2), dass diejenigen Schülerinnen und Schüler mit höherer Entwicklungsstufe (Gruppe Evaluatismus) sich in ihrem Vorgehen gemäß Lautdenk-Protokollen von einer anderen Gruppe (Absolutismus) unterschieden. Diese Differenzen betrafen ein stärker auf Organisationsleistungen (Unterstreichen und mündliches Zusammenfassen) basierendes sowie das Klären von Wortbedeutungen abzielendes Vorgehen zugunsten der erstgenannten Gruppe. Zudem lasen sie länger und wiederholten weniger oberflächlich Inhalte. In eine ähnliche Richtung geht ein Befund aus einer zweiten Studie (Iordanou et al., 2020). Laut dieser sind ebenfalls lediglich Studierende aus der Gruppe Evaluatismus dazu in der Lage, die Inhalte desjenigen Textes einer Prüfung zu unterziehen, der ihren eigenen Überzeugungen widerspricht. Damit weisen beide Studien auf eine tiefere Verarbeitung beim Lesen hin, die auf die Personen aus der spätesten Entwicklungsphase limitiert ist.
- *Bei den Sourcingaktivitäten sind nur teilweise theoretisch erwartbare Zusammenhänge nachweisbar.* Dieses Ergebnis fußt auf einer Studie mit Studierenden, deren Lautdenk-Protokolle spezifisch für das Sourcing ausgewertet wurden (Barzilai et al., 2015). In der Studie wurden die drei Skalen Absolutismus, Multiplismus und Evaluatismus mit der Häufigkeit verschiedener Sourcingaktivitäten korreliert. Nur ein Ergebnis erwies sich als theoriekonform: Personen mit höheren Werten auf der Skala Multiplismus betrieben weniger Sourcing bei zwei von drei Kategorien (Aufmerksamkeit auf Metadaten sowie Verbinden von Metadaten mit Inhalten). Bei den anderen beiden Skalen gab es keine statistisch überzufälligen Zusammenhänge.

Die Ergebnisse zeigen damit, dass die Ergebnisse analog zu den Produktmaßen eher bei den personenzentrierten Zugängen deutlicher zu sein scheinen.

4.3.4 Abschlussbemerkung

Dieses Teilkapitel hat sich einem Merkmal von Personen gewidmet, welches nicht nur in der Bildungsforschung im Allgemeinen (Hofer, 2016; Schraw, 2013) und der Leseforschung im Besonderen (Afflerbach et al., 2013; Alexander et al., 2012; Bräten, Britt et al., 2011) an Bedeutung gewonnen hat: den epistemischen Überzeugungen. Solche subjektiven wissensbezogenen Annahmen werden als mehrdimensional, also facettenreich, modelliert. Sie existieren in verschiedenen Ausprägungen und Granularitätsgrößen in Personen, unterliegen einer zeitlichen Veränderung und werden als Kontinuum einzelner Facetten mit mehr oder

weniger positiven Folgen für den Aufbau komplexer mentaler Repräsentationen beschrieben und erforscht.

Die Essenz des gegenwärtigen Stands ist, dass vor allem Überzeugungen zu Wissensmerkmalen (teils als ontologische Überzeugungen bezeichnet) und zur Begründbarkeit des Wissens (epistemische Überzeugungen) aus theoretischer und empirischer Sicht wichtig sind. Je komplexer und differenzierter diese Überzeugungen sind, als desto verstehensförderlicher gelten sie. Zwei Dimensionen bzw. Teildimensionen stechen empirisch hervor, wobei dies vor allem für intertextuelle Integrationsleistungen gilt: die Einfachheit bzw. richtiger: die Komplexität des Wissens zum einen und die Begründungen des Wissens zum anderen. Diese Begründungen des Wissens haben in der vergangenen Dekade – ausgehend von markanten Anreicherungen und Ausweitungen möglicher Wissensquellen aus philosophischer Sicht (Chinn & Rinehart, 2016b; Greene, Azevedo & Torney-Purta, 2008) und im Kontext einer teils scharf geäußerten Kritik an rein psychologischen Modellierungen aus philosophischer Warte (Holma & Hyytinen, 2015) – eine auffällige Zuwendung in der Forschung erfahren. Die Überzeugung, dass Wissen über mehrere Quellen abstützbar sein muss (nicht zuletzt wegen seiner Veränderlichkeit und Komplexität), hat relativ konsistente Zusammenhänge mit dem Verstehen kontroverser Inhalte, die in multiplen Dokumenten dargestellt wurden, in der empirischen Forschung erbracht. Allerdings ist über die Effekte auf Bestandteile des Intertextmodells noch wenig bekannt, Gleiches gilt für die Zusammenhänge von Prozessen und Überzeugungen.

4.4 Probleme

Probleme beim Integrieren sind im Bereich der Multiple Documents Literacy bekannt und bereits andernorts ausführlich dargelegt worden (Philipp, 2020a, S. 171–189). Um die inhaltliche Überschneidung bewusst gering zu halten, erfolgt mit Bezug auf die Integrationsleistungen eine Engführung auf den Aufbau von Dokumentenmodellen zum einen ([4.4.1](#)). Zum anderen sind auch die Gestaltung von multimodalen Texten mit deskriptionalen und depiktionalen Inhalten einen Blick wert, denn teils speisen sich Probleme bei der Integration von Propositionen daraus, dass die Arbeitsgedächtnisnutzung erschwert werden ([4.4.2](#)).

4.4.1 Problemfeld 1: mangelnde Kohärenzleistungen beim Aufbau umfassender (Dokumenten-)Modelle

Das Ziel beim Integrieren besteht darin, Kohärenz herzustellen, sei es lokal bei geringen Mengen an Text (Absatzebene), sei es global auf der Ebene ganzer Texte, sei es auf Ebene mehrerer Dokumente, also intertextuell. In diesem Teilkapitel

werden intertextuelle Verthebensleistungen fokussiert, nämlich enggeführt auf das Dokumentenmodell (4.4.1.1) und auf ihm ähnelnde komplexe intertextuelle Integrationsleistungen (4.4.1.2).

4.4.1.1 Dokumentenmodelle

Im Teilkapitel 4.1.2.3 wurde als Verstehensprodukt das „Dokumentenmodell“ vorgestellt, das auf der Basis von inhaltsbezogenen Kohärenzbildungsprozessen (Integrationsprozessen) sowie Sourcingprozessen eine kohärente mentale Repräsentation von multiplen Dokumenten als idealisiertes und implizit erwünschtes Produkt darstellt (Rouet & Britt, 2022). In einem Set von Studien wurden studentische Diskurssynthesen theoriegeleitet und oftmals mit holistischem Zugang dahingehend analysiert, inwieweit sie Dokumentenmodellen entsprechen. Dabei fungierte die theoriebasierte Taxonomie von Britt et al. (1999) sowie Perfetti et al. (1999) als Basis, die inhaltliche Integrationsleistungen und die Quelle-Inhalt-Verbindungen miteinander zu kombinieren und vier Ausprägungen zu unterscheiden.

- 1) Geringe Integration, teils Sourcing: „*Separate Repräsentationen*“ sind nicht-intertextuelle Integrationen, sondern sequenziell aneinandergereihte intra-textuelle Inhalte. Diese sind ohne oder mit Zitationen im Sinne von Quelle-Inhalt-Verbindungen ausgewiesen worden.
- 2) Höhere Integration, kein Sourcing: „*Mischmodelle*“ gehen darüber hinaus: Sie enthalten inhaltliche Kombinationen, weisen aber nicht aus, aus welchen Dokumenten diese Inhalte stammten.
- 3) Höhere Integration, angemessenes Sourcing: Um als „*Dokumentenmodell*“ zu gelten, mussten die Texte die wichtigsten Inhalte der gelesenen Dokumente abdecken und über Referenzierungen die Quellen von Aussagen angegeben sein.
- 4) Hohe Integration, sehr ausgiebiges Sourcing: Die höchste Stufe bildeten die „*Modelle mit allen Dokumentenetiketten*“. In ihnen wurden alle Inhalte mit der jeweiligen Quelle verbunden, sie geben also am präzisesten Auskunft über Inhalte und deren Ursprung. Diese Modelle wurden aber kaum kodiert resp. empirisch vorgefunden.

Studien, die solche taxonomischen Kodierungen vorgenommen haben, sind noch recht selten, aber die vorhandenen Ergebnisse lassen sich in Tabelle 27 verdichtet darstellen. In zwei Studien wurde die Taxonomie von mentalen Modellen nicht mit schriftlichen Aufgaben, sondern als grafische Variante operationalisiert (List, 2019; List, Du & Lee, 2021b); List (2019) erfasste ebenfalls schriftliche Maße, auf die sich die Ergebnisdarstellung in Tabelle 27 konzentriert. Die restlichen Studien in Tabelle 27 operierten ausschließlich mit schriftlichen Texten.

Tabelle 27: Verteilung von überwiegend schriftlichen Antworten beim intertextuellen Integrieren in Taxonomien, welche in der Tradition des „Dokumentenmodells“ stehen

Studie (N)	Separate Repräsentation	Mischmodell	Dokumentenmodell	Modell mit allen Dokumentenetiketten	Rudimentäre, unvollständige Modelle
Hartmann et al. (2021) (N = 87)	39 %	13 %	36 %	0 %	13 %
Lee et al. (in Druck) (N = 92*)	24 % (m. Zit.: 5 %; o. Zit.: 18 %)	8 %	63 %	–	5 %
List (2019) (N = 52*)	38 % (m. Zit.: 25 %; o. Zit.: 13 %)	23 %	21 %	–	17 %
List (2023a) (N = 136) °	42 % (m. Zit.: 7 %; o. Zit.: 35 %)	29 %	12 %	–	18 %
List, Du & Lee (2021b) (N = 33)	30 % (m. Zit.: 9 %; o. Zit.: 21 %)	39 %	30 %	–	–
List, Du, Wang & Lee (2019, Studie 1) (N = 137)	44 % (m. Zit.: 30 %; o. Zit.: 14 %)	20 %	36 %	1 %	–
List, Du, Wang & Lee (2019, Studie 2) (N = 29)	55 % (m. Zit.: 38 %; o. Zit.: 17 %)	21 %	24 %	0 %	–
List, Oaxaca et al. (2021) (N = 69*)	41 % (m. Zit.: 19 %; o. Zit.: 22 %)	20 %	39 %	–	–
Gesamt (N = 635; gewichteter Mittelwert über alle Studien)	39 %	20 %	32 %	0 %	8 %

(Legende: m. Zit. = mit Zitationen, o. Zit. = ohne Zitationen; Abweichungen bei den Summen ergeben sich durch Rundungsfehler; * Abweichungen von den in den Primärstudien berichteten Samplegrößen und den Prozentangaben ergeben sich durch Summierungen der eindeutig kodierten Antworten in den Primärstudien; – = in der Studie nicht kodiert; ° Daten in der Originalpublikation nicht benannt, am 20.6.2023 von der Autorin auf Nachfrage bereitgestellt; Abweichungen von 100 Prozent sind rundungsfehlerbedingt)

Aus der Darstellung in der Tabelle 27 lassen sich drei Muster abstrahieren:

- *Große Heterogenität:* Die Daten aus den acht Studien weisen bei den prozentualen Anteilen der Kategorien eine erhebliche Spannweite auf, die sich deutlich in den separaten Repräsentationen zeigt: Zwischen 24 und 55 Prozent lagen die Werte. Bei den Mischmodellen unterschieden sich die Quoten mit einer Spannweite von 8 bis 39 Prozent um den Faktor 5, bei den Dokumentenmodellen variierten sie zwischen einem knappen Achtel und fast zwei Dritteln. Lediglich beim Modell mit allen Dokumentenetiketten gab es aufgrund der Seltenheit seines Auftretens kaum Variabilität.

- *Geringer Anteil von Dokumentenmodellen:* Wegen der Heterogenität der Befunde wurden zuunterst in der Tabelle 27 die Anteile über alle Studien hinweg ausgewiesen. Demnach bilden Mischmodelle und separate Repräsentationen zusammengenommen fast drei Fünftel aller kodierten Textprodukte, wobei die separaten Repräsentationen die am häufigsten vorkommende Kategorie darstellten. Dokumentenmodelle waren hingegen in nur jedem dritten Aufgabenprodukt kodierbar.
- *Hohe Anteile von rudimentären mentalen Modellen:* In vier der Studien wurden sonstige Kategorien gebildet, auf die bis zu einem guten Sechstel der Aufgabenprodukte entfielen. Dies waren entweder Lösungen mit Aussagen aus nur einem Dokument bzw. reine Vorwissensbestände (Hartmann et al., 2021; List, 2019, 2023a) oder zu kurz, um sie überhaupt kodieren zu können (List, 2023a).

Das einzelstudienübergreifende Muster besteht also darin, dass es den Testpersonen eher selten gelingt, die Dokumenteninhalte angemessen inhaltlich zu kombinieren und mittels Metadaten auszuweisen, woher diese Inhalte stammen.

4.4.1.2 Dokumentenmodell-verwandte Maße

Nicht direkt das Dokumentenmodell sowie dessen Vorstufen als theoriegeleitete Klassifikation gemäß Britt et al. (1999) bzw. Perfetti et al. (1999), sondern andere und verwandte Maße wurden in weiteren Studien verwendet. In diesen bislang sporadisch durchgeführten Untersuchungen wurden Maße der inhaltlichen Integration und der Quelle-Inhalt-Nennungen kombiniert. Drei dieser Studien kommen zu folgenden Ergebnissen:

- Einen Sammelindex namens „Integration von Inhalten und Quellen“ nutzten Schuster et al. (2023) in einer Experimentalstudie aus der Domäne Geschichte. Die Testpersonen lasen drei Texte und sollten danach erklären, warum es zu dem historischen Ereignis gekommen war. Die schriftlichen Antworten wurden nach drei Variablen kodiert: dem Abdeckungsgrad der zentralen historischen Ereignisse, der Anzahl der kontrovers eingeschätzten Sachverhalte (intertextuelle Konflikte) und schließlich der Zahl der expliziten Quelle-Inhalt-Verbindungen. Diese drei Dimensionen bzw. Bestandteile, die das Autorinnen- und Autorenteam explizit als Indikator für Dokumentenmodelle bezeichnet, wurden mit leicht unterschiedlichen Gewichtungen zu einer Skala kombiniert. Von der potenziell zu erreichenden Gesamtpunktzahl erzielten die drei Gruppen zwischen 20 bis 27 Prozent; die Kontrollgruppe wies die geringsten Werte auf.

- In einer Studie, in der die Integrationsleistungen von Studierenden auf der Basis von zwei Videos oder Dokumenten erfasst wurden (List, 2018), kam eine fünfstufige Skala zum Einsatz. Sie bezog sich auf kurze schriftliche Antworten, in denen die Testpersonen Inhalte zum Artenschutz aus den beiden Videos bzw. Dokumenten nutzen sollten. Die Antworten wurden danach kodiert, ob die Probandinnen und Probanden a) nur eines der Dokumente bzw. der Videos nutzten, b) beide verwendeten, sie aber linear hintereinander reihten (separate Repräsentation), c) generisch integrierten (Mischmodell), d) partiell integrierten (implizite Metadatenutzung resp. mit nur wenigen Details) oder e) komplett integrierten. Die Mittelwerte der Skala entsprachen den Integrationsleistungen zwischen einer separaten bzw. einer partiell integrierten Lösungsvariante.
- In einer neueren Studie (List, Du & Lee, 2021b) verwendete man eine vierstufige Skala, mittels derer die Integrationsleistungen in der Logik des Dokumentenmodells erfasst wurden. In dieser Studie wurden die Testpersonen gebeten, eine offene Frage zu kontroversen Sachverhalten schriftlich zu beantworten. Mittels der Skala wurde klassifiziert, ob Personen die Hauptaussagen korrekt intertextuell kombiniert, ob sie diese Hauptaussagen mit Details unterfüttert und ob sie die Dokumente über Quelle-Inhalt-Verbindungen mit den Inhalten verbunden hatten. Den Skalen-Mittelwerten nach demonstrierten die Personen nur Leistungen auf der Stufe der intertextuellen Inferenzen der Hauptideen und damit nur auf der Ebene des integrierten mentalen Modells bzw. entsprachen ihre Leistungen am ehesten einem Mischmodell.

Die Ergebnisse dieser Studien und jene zu den Kodierungen von längeren Antworten in der Dokumentenmodell-Taxonomie verdeutlichen, dass umfassende mentale Repräsentationen, die in schriftlichen bzw. teils in grafischen Aufgabenprodukten wiedergegeben wurden, nur in geringen Anteilen vorkamen. Dabei zeichnen sich musterhaft zwei Schwierigkeitsfelder ab: Zum einen besteht ein Hindernis darin, *textübergreifend Inhalte inferierend zu integrieren*, was sich in den großen Anteilen der separaten Repräsentationen manifestiert (s. ähnlich auch Teilkap. 5.4.2.1). Zum anderen liegt die Schwierigkeit vor, *Metadaten systematisch mit den Inhalten zu verbinden*. Dies illustrieren die geringen Skalenmittelwerte in den Studien mit Dokumentenmodell-ähnlichen Instrumenten und die Anteile von Mischmodellen, separaten Repräsentationen ohne Zitationen und rudimentären mentalen Modellen in der Kategorie „Rudimentäre, unvollständige Modelle“ aus Tabelle 27.

4.4.2 Problemfeld 2: kognitiv ungünstig gestaltete Texte/ Dokumente aus Sicht der verstehenserschwerenden Arbeitsgedächtnisbelastung

Anders als sonst in den anderen Teilkapiteln zu eher problematischen individuellen Merkmalen scheint hier ein rein externes Merkmal im Vordergrund zu stehen. Selbst wenn sich dieses Teilkapitel tatsächlich außerhalb der lesenden Person befindlichen, die Texte selbst betreffenden Merkmalen widmet, so geschieht dies mit Blick auf ein internes Merkmal, nämlich die Funktionsweise des Arbeitsgedächtnisses. Speziell die Forschung zum integrierten Text-Bild-Verstehen bzw. zum multimedialen Lernen (s. Teilkap. [4.1.2.2](#)) hat die Funktionsweise des Arbeitsgedächtnisses adressiert. Dabei gilt das Arbeitsgedächtnis nicht nur als Ort der aktiven Bedeutungskonstruktion, sondern auch als limitierender Faktor wegen seiner Begrenztheit, für eine kurze Zeitspanne wenige Informationen zur Verfügung zu halten.

Das Arbeitsgedächtnis wird an anderer Stelle in diesem Band vertieft und ausführlicher dargestellt, nämlich als Teil der exekutiven Funktionen im Kapitel zur Selbstregulation (s. Teilkap. [6.1.3.2](#)). Gleichwohl ist das Arbeitsgedächtnis bedeutend bei den Integrationsleistungen und verdient darum Erwähnung. Der Grund liegt in dessen Funktionsweise: Das Arbeitsgedächtnis kann nur wenige Information für lediglich kurze Zeit verarbeiten, doch auf diese Informationen in ihrer Komplexität kommt es für das Lernen und das Leseverstehen an (Sweller, 2020). Darum stellen Theorien wie jene zur *kognitiven Belastung* (Sweller et al., 2019), auf denen auch die im Teilkapitel [4.1.2.2](#) dargestellten beiden Modelle aufbauen, die möglichst effektive Nutzung der limitierten Arbeitsgedächtniskapazitäten in ihr Zentrum (Camp et al., 2022). Für das Lesen sind hier für das reine Verstehen zwei Aspekte bedeutend, die für die sog. *lerngegenstandsinhärente kognitive Belastung*, welche sich a) aus der Komplexität des Gegenstands selbst (als Element-Interaktivität bezeichnet) im Sinne von Zahl und Vernetzungsnotwendigkeit von Informationen und b) dem beim Integrieren helfenden Vorwissen der lesenden Person ergibt (Sweller et al., 2011). Diese Art der kognitiven Belastung wird also als Wechselspiel von externalen und internalen Merkmalen konzeptualisiert und gilt als essenziell für jegliches Verstehen von Texten. Entsprechend fokussiert vor allem die reichhaltige Forschung zum multimedialen Lernen die Frage, wie insbesondere komplexe Gegenstände möglichst verstehensförderlich gestaltet werden können bzw. wie sie nicht unnötig komplexer werden sollten (Mayer, 2021; Mayer & Moreno, 2003; Noetel et al., 2022). Als Gegenstück dazu adressiert die Forschung die Frage, welche lesendenseitigen Strategien dabei helfen, suboptimal konstruierte Texte bzw. Text-Bild-Kombination trotzdem zu verstehen (Castro-Alonso et al., 2021).

Die Theorie der kognitiven Belastung greift als – ursprünglich zur optimalen Gestaltung von Lernmaterialien verwendet – diesen Gedanken insofern auf,

als sie eine Variante der kognitiven Belastung, also der Beanspruchung des Arbeitsgedächtnisses durch die Verarbeitung von Informationen, als ungünstig annimmt. Die *sachfremde kognitive Last* (extraneous cognitive load) führt gemäß der Theorie zur kognitiven Belastung dazu, dass Personen einen zusätzlichen Verarbeitungsaufwand zu leisten haben, der sich nicht als verstehensförderlich erweist, weil er anderweitig benötigte Ressourcen für sich beansprucht (Kalyuga, 2011). Darum lassen sich aus der Theorie der kognitiven Belastung Prinzipien ableiten, wie eine solche Über-Last im Allgemeinen und speziell eine solche, die aus einer vermeidbaren, ungünstigen Gestaltung von Texten bzw. Text-Bild-Kombinationen resultiert, verhinderbar sind. Inzwischen liegen vielfältige Experimentalstudien vor, die in Metaanalysen ausgewertet wurden und sich zu Prinzipien verdichten lassen (s. Tabelle 28).

Tabelle 28: Überblick über drei Prinzipien zur Minimierung der sachfremden kognitiven Last beim Lernen mit Inhalten aus Bild-Text-Kombinationen

Prinzip	Beschreibung	Anwendungssituationen	Exemplarische Metaanalyse
Kohärenzprinzip	Zusatzinformationen wie ablenkende Details aus Texten und Bildern entfernen	besonders geeignet für Leserinnen und Leser mit geringem Vorwissen oder geringer Arbeitsgedächtniskapazität, komplexere Lektionen und dort, wo Personen von sich aus nur wenig kognitives Engagement aufbringen	Rey (2012), Sundararajan & Adesope (2020)
Signalisierungsprinzip	Hinzufügen von Hinweisen, um die Organisation des Materials hervorzuheben	besonders geeignet für Leserinnen und Leser mit geringem Vorwissen, komplexere Inhalte und wenn die Hervorhebung die Text-Bild-Integration unterstützt	Alpizar et al. (2020), Richter et al. (2016), Schneider et al. (2018), Xie et al. (2017)
Kontiguitätsprinzip	Text neben den entsprechenden Grafiken platzieren (1) räumliche Kontiguität) bzw. zeitgleich präsentieren (2) zeitliche Kontiguität)	1) besonders geeignet für komplexere Lektionen und wenn Lernende die Strategien zum Selbstmanagement von separierten Bestandteilen nicht beherrschen 2) besonders geeignet für Leserinnen und Leser mit geringen Vorkenntnissen, für komplexe Inhalte ohne Temposteuerung, wenn der Text gesprochen und nicht schriftlich dargeboten wird und wenn die Wörter und Grafiken als größere Segmente präsentiert werden	1) Ginns (2006), Schroeder & Cencki (2018, 2020) 2) Ginns (2006)

(Quelle der Darstellung: Fiorella & Mayer, 2022, S. 195, sowie Noetel et al., 2022)

Die drei Prinzipien aus der Tabelle 28 lauten:

- *Zusatzinformationen („verführerische Details“) minimieren das Verstehen* (Kohärenzprinzip, s. Teilkap [4.4.2.1](#)).
- *Texte und Text-Bild-Kombinationen ohne Hinweise auf verstehensnotwendige Elemente sind schwerer zu verstehen* (Signalisierungsprinzip, [4.4.2.2](#)).
- *Text-Bild-Kombinationen mit räumlich oder zeitlich weit auseinanderliegenden Informationen werden schlechter verstanden als solche mit räumlich bzw. zeitlich nahen Informationen* (Kontiguitätsprinzip, [4.4.2.3](#)).

Diese drei Prinzipien bzw. welche Auswirkungen es hat, wenn diese Prinzipien nicht beachtet werden, werden im Folgenden anhand jener Metaanalysen entfaltet, die bereits in Tabelle 28 in der Spalte rechts aufgeführt sind.

4.4.2.1 Verletzungen des Kohärenzprinzips

Das Kohärenzprinzip (Fiorella & Mayer, 2022) setzt daran an, dass Personen dann besser lernen, wenn Texte bzw. Text-Bild-Kombinationen keine redundanten Informationen enthalten – solche Informationen werden auch beim gelingenden Verstehen mittels Makroregeln von der lesenden Person reduzierend getilgt. Zu den überflüssigen Informationen zählen unwesentliche Informationen (wie zusätzliche Details) oder interessante, aber wenig verstehensdienliche Informationen (verführerische Details). Aus diesen beiden Arten und wie sich Texte optimieren lassen, sind vor allem die „verführerischen Details“ (*seductive details*) und ihre kohärenzerschwerende Wirkung untersucht worden

Der Ausdruck „verführerische Details“ geht auf Garner et al. (1989) zurück, die damit mikropropositionale Informationseinheiten meinten, die zwar interessant sind, für den Aufbau globaler Kohärenz allerdings keinen Beitrag leisten. Der Begriff wird unterschiedlich verwendet, manche Autorinnen und Autoren limitieren ihn ausschließlich auf Informationen aus zu lesenden Texten selbst, andere bezeichnen damit jegliche Form von Stimulus, der aus der Umwelt kommen kann, und wiederum andere nehmen eine mittlere Position ein, indem sie jegliche interessanten, aber mit dem Ziel des Lesens bzw. Lernens nicht verbundenen Informationen als verführerische Details deklarieren (Rey, 2012).

Theoretische Erklärungsversuche

Aus theoretischer Sicht werden verschiedene Ursachen für den ungünstigen Effekt verführerischer Details angenommen. Dabei gibt es keine universelle Kausalerklärung für den Effekt, da dies unterkomplex wäre und da sich sowohl theoretische als auch empirische Belege für mehrere Erklärungen anbringen lassen. Rey (2012) nennt vier theoriebasierte Hypothesen.

- *Hypothese zur Überlastung des Arbeitsgedächtnisses*: Diese Hypothese besagt, dass aufgrund der limitierten Kapazitäten des Arbeitsgedächtnis durch die verführerischen Details jene Kapazitäten beansprucht werden, die für die essenzielle Verarbeitung benötigt würden, also für Prozesse, um die Komplexität des Inhalts angemessen zu repräsentieren.
- *Ablenkungshypothese*: Gemäß dieser Hypothese führen Salienz und leichte Verarbeitbarkeit von verführerischen Details dazu, dass die Aufmerksamkeit sich auf sie richtet statt auf die essenziellen Informationen, die abstrakterer Natur oder schwerer verständlich sein können.
- *Interferenzhypothese*: Diese Hypothese geht davon aus, dass verführerische Details die falschen Schemata aus dem Langzeitgedächtnis aktivieren, die zu einem inadäquaten Aufbau einer internalen Repräsentation führen. Diese Inadäquanz besteht darin, dass Details dominieren.
- *Hypothese zur Kohärenzunterbrechung*: Diese Hypothese basiert auf der Annahme, dass der mentale Aufbau von Kohärenz beim Lesen optimaler erfolgt, wenn der Übergang von Hauptideen untereinander in Texten möglichst wenig unterbrochen wird, was verführerischen Details attestiert wird.

Wie aus der knappen Aufzählung oben ersichtlich ist, liegen die Hypothesen nicht als Set einander ausschließender Annahmen vor, vielmehr sind sie komplementär. Die darin postulierten Mechanismen wirken mutmaßlich zudem interaktiv und müssen für Personengruppen und weitere Merkmale spezifiziert werden. Denn für die Forschung zu den Effekten verführerischer Details auf das Lernen aus Texten und aus Text-Bild-Kombinationen ist kennzeichnend, dass die Befunde heterogen ausfallen (Alexander, 2019; Mayer, 2019b), was partiell methodischen Einflüssen sowie Merkmalen der verführerischen Details geschuldet ist. Allerdings sind die Zusammenhänge vermutlich komplexer, als es das gegenwärtig formulierte Kohärenzprinzip nahelegt.

Empirische Befunde (zu Produkten)

Zwei Metaanalysen (Rey, 2012; Sundararajan & Adesope, 2020) haben einige empirische Muster herausgearbeitet, die darauf verweisen, welche Merkmale mit stärker ausgeprägten Minderleistungen im Verstehen zusammenhängen. Die metaanalytischen Hauptergebnisse waren vergleichbar: Sowohl Erinnerungsleistungen als auch Transferleistungen fielen in den Gruppen schlechter aus, deren Bild-Text-Kombinationen bzw. Texte verführerische Details enthielten. Diese Metaanalysen konnten teils empirische Merkmale extrahieren, die mit stärkeren oder schwächeren experimentübergreifenden Effekten verführerischer Details korrespondierten. Stärkere negative Effekte bestanden,

- wenn es zeitliche Restriktionen bei der Lern- und Leistungsphase gab (Rey, 2012; Sundararajan & Adesope, 2020);
- wenn die Darbietung in gedruckten Medien statt digital erfolgte (Sundararajan & Adesope, 2020);
- wenn die verführerischen Details in Grafiken enthalten waren, während diese Effekte bei schriftlichen Texten erheblich geringer ausfielen (Rey, 2012), besonders negativ war der Effekt in Diagrammen (Sundararajan & Adesope, 2020);
- wenn verführerische Details statisch vorlagen, also die gesamte Zeit präsent waren, statt nur vorübergehend aufzutauchen (Sundararajan & Adesope, 2020);
- wenn die Personen über wenig Vorwissen in der jeweiligen Wissensdomäne verfügten (Sundararajan & Adesope, 2020).

Das Kohärenzprinzip macht darauf aufmerksam, dass es ungünstig ist, wenn Personen Bild-Text-Kombinationen, aber auch einzelne Texte verstehen sollen, welche Informationen enthalten, die eine gewisse Interessantheit aufweisen, aber in der Makrostruktur von Texten eine untergeordnete Rolle spielen (sollten). Dafür gibt es verschiedene theoretische Erklärungsansätze, welche darauf hinauslaufen, dass verführerische Details den Aufbau von kohärenten internalen Repräsentationen erschweren, weil sie Aufmerksamkeit auf sich lenken, was lückenhafte oder fehlgeleitete Repräsentationen nach sich zieht. Empirisch gesehen zeigen metaanalytisch ausgewertete Experimentalstudien, dass bei wenig Zeit in der Rezeption von vor allem analogen und Bilder enthaltenen externen Repräsentationen dauerhaft zugängliche verführerische Details negative Auswirkungen haben. Das gilt insbesondere für Personen mit geringem Vorwissen.

4.4.2.2 Verletzungen des Signalisierungsprinzips

Das Signalisierungsprinzip (manchmal auch als „Hinweisprinzip“ bezeichnet), bezieht sich darauf, dass Menschen dann bessere Verstehensleistungen bei Bild-Text-Kombinationen bzw. bei Texten erzielen, wenn diese explizite Hinweise enthalten, um die Aufmerksamkeit auf bestimmte Elemente bzw. auf deren Organisation zu lenken (van Gog, 2022). Damit fokussiert das Signalisierungsprozess primär das Auswählen von Informationen als kognitivem Prozess, welches es gezielt zu unterstützen und zu lenken versucht.

Je nach Modalität der zu verarbeitenden externalen Repräsentationen gibt es Unterschiede in dem, wie das Signalisierungsprinzip umgesetzt werden kann (van Gog, 2022):

- Im Falle *sprachlich kodierter Informationen* sind Einfügungen von Sätzen bzw. Aussagen, welche die Organisation des Textes oder spezifischer Hinweise

für die lesende Person, ihre Aufmerksamkeit auf einzelne Elemente zu lenken, eine Möglichkeit. Eine weitere Variante besteht darin, über farbige oder sonstige typografische Hinweise bei schriftlichen Texten bzw. über Betonungen bei mündlichen Texten, subtilere Signale zu senden.

- Bei *Bildern* können ebenfalls verschiedene Varianten des Signalisierens zum Einsatz kommen. Extrinsisch ist das Hinzufügen von Pfeilen, eher intrinsisch das Hervorheben von einzelnen Elementen durch farbiges Betonen, durch Animationen oder durch das Ausblenden nicht-fokussierter Bestandteile.
- Wenn *Bild-Text-Kombinationen* betroffen sind, so lassen sich durch gleiche Farben sowohl im Text als auch im Bild Hilfestellungen anbieten, deskriptionale und depiktionale Informationen in ihrer Korrespondenz zu erkennen.

Prozessperspektiven

Prozesse sind in den Metaanalysen bislang nur wenig untersucht worden, jene von Schneider et al. (2018) bildet eine Ausnahme. Sie fokussierte zwei Prozessmaße: die Bearbeitungsdauer und Fixationsdauern in Studien mit Eye-Tracking. Das Hinzufügen von Signalen in Text-Bild-Kombinationen

- führt zu einer *Verlängerung der Lernzeit*, das bedeutet, die Personen wenden im Vergleich zu Probandinnen und Probanden mehr Zeit auf, die keine Hervorhebungen im Lernmaterial hatten;
- erhöht die *Fixationszeiten von Inhalten* (in sogenannten „Areas of Interest“), die mit den Signalisierungen hervorgehoben haben, was als Indikator für gesteigerte Aufmerksamkeit und Verarbeitung gewertet werden kann und die verlängerte Lernzeit zum Teil erklärt.

Signalisierungen verlängern demnach die Lernprozesse, indem sie die Aufmerksamkeit der Lernenden auf jene Areale ziehen, die absichtsvoll hervorgehoben worden sind.

Produktperspektiven: kognitive und metakognitive Maße

Deutlich besser als die Prozesse sind die Ergebnisse der Prozesse untersucht worden, also die Produkte. Hier gibt es drei kognitive Bereiche, welche als abhängige Variablen in den Studien operationalisiert worden sind. Das Hinzufügen von Hinweisen in Bild-Text-Kombinationen

- führt zu verbesserten Leistungen des *Erinnerns* (Alpizar et al., 2020; Richter et al., 2016; Schneider et al., 2018; Xie et al., 2017);
- mündet in besseren Anwendungen des Gelernten in *Transferleistungen* (Alpizar et al., 2020; Richter et al., 2016; Schneider et al., 2018; Xie et al., 2017);

- senkt in geringem Maß die subjektiv *wahrgenommene kognitive Belastung* (mentaler Aufwand, wahrgenommene Schwierigkeit oder eine Kombination aus beidem; Schneider et al., 2018; Xie et al., 2017), das heißt, Personen schildern während oder nach dem Lernen die Aufgabe als leichter, es gibt also Effekte auf die metakognitiven Urteile. Besonders erwähnenswert ist die Reduktion der subjektiv eingeschätzten kognitiven Belastung insofern, als diese Reduktion substanziell mit den Lernleistungen zusammenhing, und zwar bei Erinnerungs- und bei Transferleistungen (Xie et al., 2017).

Je nach Metaanalyse wurden weitere Merkmale auf Moderatoreffekte überprüft, die sich in zwei Gruppen bündeln lassen, wobei an dieser Stelle nur die signifikanten Differenzen berichtet werden. Im Falle der personenbezogenen Merkmale erwies sich ein Merkmal teils als statistisch auffälliger Moderator, teils auch nicht. Personen mit geringerem *Vorwissen* profitierten mehr von Hinweisen als Personen mit hohem Vorwissen, bei denen keine positive Wirkung mehr nachweisbar war (Alpizar et al., 2020; Richter et al., 2016). Die kompensierende Wirkung der Hinweise tritt demnach nicht gleichermaßen für alle Personen auf: Personen mit hohem Vorwissen profitieren anscheinend nicht davon. Jedoch gibt es auch größere Metaanalysen, die keinen moderierenden Effekt des Vorwissens haben nachweisen können (Schneider et al., 2018).

Teilweise prüften die Metaanalysen noch, ob es Merkmale des Signalisierens in den Experimenten selbst gab, die besonders günstig waren. In diese zweite Gruppe von Variablen entfielen Merkmale, die sich aus der Gestaltung der Materialien und des technischen Umgangs mit ihnen speisen und welche die Erinnerungs- und Transferleistungen betreffen:

- *Dynamische Signale*, also solche, die nicht die gesamte Zeit unverändert waren, gingen mit besseren kognitiven Leistungen einher (Alpizar et al., 2020).
- Günstig erwies sich das Signalisieren, wenn die Personen die Möglichkeit hatten, *das Tempo der Bild-Text-Kombinationen* zu beeinflussen (Alpizar et al., 2020).
- *Je kürzer die Lernphasen* waren, desto positiver wirkten sich Signale in der Regel aus. Ein ähnliches Muster bestand generell bei kürzeren Experimenten, d. h. unter Zeitdruck machte sich der Effekt allgemein deutlicher bemerkbar (Alpizar et al., 2020).
- Bei *Inhalten ohne Bilder* (dafür aber mit anderen grafischen Darstellungen) war der Effekt des Signalisierens höher, als wenn die externalen Repräsentationen Bilder umfassten (Alpizar et al., 2020). Dieses Ergebnis bestätigten Schneider et al. (2018), allerdings nur für Erinnerungsleistungen. Textimmanente Hervorhebungen hatten eine stärker positive Wirkung, was vor allem auf das Kenntlichmachen der Struktur des jeweils zu lesenden Textes mittels Unterstreichungen und Überschriften zurückzuführen war.

Es scheint, dass sich Signalisierungen vor allem in Studien mit wenig Lernzeit, aber mit einer gewissen Steuerbarkeit durch die Lernenden positiv auszuwirken. Dynamische Signalisierungen mit höherer Salienz und die textuelle Makrostruktur von Texten hervorhebende Aufmerksamkeitslenkungen erwiesen sich ebenfalls vorteilhaft. Daneben ist der Vollständigkeit halber noch anzuführen, dass diverse Moderatoranalysen keine Muster hervorbrachten, was als Indikator für eine breite Anwendbarkeit des Signalisierungsprinzips spricht. Daraus ergibt sich, spezifisch bei den Transferleistungen, die ein gewisses Maß an Integrationsprozessen voraussetzen, das schwierigkeitsgenerierende Merkmal: Müssen Personen im Prozess des Lernens Wichtiges von Unwichtigem unterscheiden, dann dürfte dies mit einer erhöhten kognitiven Belastung einhergehen. Dass ausgerechnet die Lernzeit aber geringer ausfällt, wenn Personen auf nicht eigens optimierte Bild-Text-Kombinationen treffen, kann als prozessbezogener Indikator eines Problems interpretiert werden.

4.4.2.3 Verletzungen des Kontiguitätsprinzips

Das Kontiguitätsprinzip besagt, dass Lernende Inhalte besser verstehen, wenn sie auf disparate Einheiten verteilt vorliegen, aber integriert werden müssen und dafür idealerweise nah beieinander sind (Fiorella & Mayer, 2022). *Geteilte Aufmerksamkeit* ist die Ursache für das Kontiguitätsprinzip. Geteilte Aufmerksamkeit liegt dann vor, wenn Informationen auf zwei räumlich oder zeitlich voneinander getrennte Informationsquellen verteilt sind, aber es nötig ist, beide Informationen integrierend aufeinander zu beziehen. Wenn der Abstand zwischen den disparaten Quellen zu groß ist, das Prinzip der Kontiguität bzw. Nähe also verletzt wird, entsteht dadurch eine erhöhte kognitive Belastung, die sich negativ auf das Lernen auswirkt (Ayres & Sweller, 2022).

Allerdings gilt diese Erklärung als umstritten, wie es Schroeder und Cenkci (2020) anhand ausgesprochen heterogener Studienbefunde verdeutlichen: Ihrer kritischen Sichtung der empirischen Forschungsliteratur gemäß sind die Belege für eine Reduktion der kognitiven Belastung durch das Herstellen von räumlicher Nähe weder breit genug abstützbar noch plausibel genug, um erklären zu können, warum der Effekt zustande kommen soll. Sie schlagen daher, sich unter anderem auf Mayers (2022a) kognitive Theorie des Multimedia-Lernens stützend (s. Teilkap. 4.1.2.2), eine alternative Erklärung vor. Demnach vermuten Schroeder und Cenkci (2020) nicht, dass die Erhöhung der kognitiven Belastung die alleinige Ursache sei. Vielmehr sei die räumliche Nähe von aufeinander zu beziehenden Informationen darum hilfreich, weil sie Leserinnen und Lesern nahelege, dass Informationen zueinander gehören, sodass dadurch das aktive Integrieren erleichtert werde (Alemdag & Cagiltay, 2018). Schroeder und Cenkci (2020) knüpfen ihre Hypothese allerdings an eine Bedingung: Sie vermuten, dass räumliche Nähe nur dann verstehensförderlich ist, wenn es zeitliche Limitationen

beim Lesen gibt, da bei ausreichend Zeit die Leserinnen und Lesern kompensierende Vorgehensweisen anwenden können.

Bisherige metaanalytische Forschungsüberblicke bestätigen zunächst, dass das Kontiguitätsprinzip besteht. Denn sowohl zeitliche (Ginns, 2006) als auch räumliche Nähe (Ginns, 2006; Schroeder & Cenkci, 2020) verbessern Erinnerungs- und Transferleistungen. Schroeder und Cenkci (2018) stellten fest, dass Lernleistungen mit räumlich nahen Informationen besonders bei a) statischen Bild-Text-Kombinationen sowie b) inhaltlich nicht-redundanten Text- und Bild-Elementen besser ausfielen. Anders gesagt: Mussten die Inhalte aufeinander bezogen werden und blieben die einzelnen Bestandteile fixiert (waren also bspw. nicht animiert), war räumliche Nähe von Vorteil. Abgesehen von diesen beiden Effekten erwies sich der Effekt räumlicher Nähe in den Experimenten als stabiler, studienübergreifender Befund.

4.4.2.4 Schlussbemerkung

Die Prinzipien des Multimedia-Lernens und der günstigen Gestaltung von Bild-Text-Kombinationen sind für das Lesen und den Aufbau kohärenter mentaler Repräsentationen bei digitalen Dokumenten deshalb so wichtig, weil sie ursprünglich aus einem Kontext stammen, in dem es um die *Optimierung* von Texten und Lernmaterialien zum Zweck des Lernens ging (Fiorella & Mayer, 2022; Mayer & Moreno, 2003). Wenn Personen auf sehr unterschiedliche digitale Dokumente variierender Güte, Provenienz, kommunikativer Absicht und Faktennähe etc. im Internet stoßen, kann man nicht berechtigt damit rechnen, dass sie lernoptimierte Dokumente rezipieren. Vielmehr ist davon auszugehen, dass Dokumente viele Details beinhalten, möglicherweise auch gezielt solche, die von Hauptpunkten ablenken, dass sie nicht ihre zentralen Inhalte und Relationen hervorheben und dass diese Dokumente Informationen enthalten, die distal über verschiedene Dokumente verteilt sind. Oder anders formuliert: Es ist eher erwartbar, dass kontingente Dokumentensets bei realweltlichen Internet-Lesevorgängen nicht optimal hinsichtlich der kognitiven Belastung ausfallen, sondern im Gegenteil die kognitive Belastung eher sogar erhöht wird.

4.5 Fazit

Das Verstehen von external gegebenen Informationen, seien es textuelle, seien es grafische, mögen sie aus einzelnen Texten stammen oder auf mehreren Dokumenten basieren, ist oftmals das übergeordnete Ziel des Lesens. Auch wenn es in der Leseforschung dazu diverse Modelle und Theorien gibt, so wird Lesen als ein dynamisch ablaufender Prozessverbund modelliert, bei dem die lesende Person aufgrund von kognitiven Transformationsprozessen zu verschiedenartigen

Arten von internalen Repräsentationen gelangt. Die in diesem Kapitel vorgestellten Produkte des Leseverstehens eint die Architektur: Es handelt sich um Mehrebenen-Repräsentationen, bei denen sich der Grad an Transformationen von Informationen unterscheidet, wobei die Nutzung des eigenen Vorwissens betont wird, da vorwissensbasierte Schlussfolgerungen – Inferenzen – als Schlüssel für den Aufbau kohärenter Mehrebenen-Repräsentation gelten (McNamara, 2021).

Der Prozess, in dem das Inferieren eine wichtige Rolle spielt, wird als *Integrieren* bezeichnet und dient dazu, ein stabiles Netzwerk an Propositionen zu extrahieren bzw. anzureichern, um Sinnzusammenhänge auf lokaler und globaler Textebene, zwischen textuell-deskriptional und grafisch-depiktional kodierten Inhalten, aber auch über Dokumentengrenzen hinweg mit Rekurs auf Metadaten, begründbare Perspektiven und Inhalte herzustellen. Insbesondere im Zusammenhang mit dem Verstehen von Text-Bild-Kombinationen zum einen und multiplen Dokumenten zum anderen wird das Integrieren als Bedingung des angemessenen, tiefen Verstehens begründbar angenommen (Alexander & Disciplined Reading and Learning Research Laboratory, 2020; Barzilay et al., 2018; Cartiff & Greene, 2020; Rouet & Britt, 2022; Schnotz, 2022).

Als idealtypisches Produkt des Verstehens gilt das *mentale Modell*, welches eine vernetzte Repräsentation external gegebener Informationen sowie internaler Vorwissensbestände im Gedächtnis der lesenden Person darstellt und zentrale Informationen beinhaltet, die aufgrund individueller Vorwissensbestände in eine kohärente, also widerspruchssarme Repräsentation überführt worden sind. Ein solches mentales Modell lässt die Modalität seines externalen Inputs hinter sich, wofür die kognitiven Transformationsprozesse verantwortlich sind. Im Falle des Lesens multipler Dokumente wird jedoch auffällig, was bisherige Modelle des Leseverstehens auffällig außen vor gelassen haben: der Umstand, dass Dokumente (verstanden als Kombination von Informationen und Metadaten) Perspektiven auf Gegenstände einnehmen und dass sie dies innerhalb eines bedeutungsrelevanten Kontexts tun. Daher gehen neuere Modelle dezidiert auf diesen blinden Fleck ein und ergänzen Mehrebenen-Repräsentationen um neue Bestandteile, z. B. das „Intertextmodell“ mit Metadaten und der intertextuellen Struktur von Dokumenten als Rückgrat komplexerer integrierter mentaler Modelle auf der Basis von separaten, teils intra- und intertextuell kombinierten Informationen in den mentalen Modellen.

Für den Aufbau solcher mentalen Modelle ist das Integrieren als Prozessverbund essenziell. Integration wird im Zusammenhang mit dem Herstellen von Kohärenz beim Lesen multipler Dokumente (durchaus in verschiedenen Modalitäten) auf zweierlei Art in der Forschung beschrieben. Zum einen wird Integrieren als automatischer, wenig steuerbarer Aktivierungs- und Resonanzprozess modelliert, wenn die Informationen aus verschiedenen Dokumenten semantisch stark überlappen. Zum anderen gilt Integrieren als absichtsvoller, lesestrategischer Prozess der Informationsprüfung und Inferenzbildung, wenn

Informationen eher implizite intertextuelle Beziehungen aufweisen (Kurby et al., 2005; List, Du & Lee, 2021b; van den Broek & Kendeou, 2015). Die hohen Herausforderungen des schlussfolgerungsbasierten Integrierens, insbesondere bei mangelndem Vorwissen und bei komplexen Sachverhalten wie Klimawandel, Impfstoffen etc., legen es nahe, dass intertextuelles Integrieren eine lernstrategische Dimension aufweist. Lesende Personen können nicht darauf vertrauen, auf kohärente Dokumentensets zu stoßen, sondern müssen wachsam und aufmerksam ihr Verstehen überwachen, um thematische Bezüge jedweder Art zu erkennen (Cho, Afflerbach & Han, 2018; Coiro, 2015). Darum wird das Lesen multipler Internetdokumente als selbstreguliertes Lesen konzeptualisiert, bei dem das metakognitive Überwachen von Kohärenzbrüchen und das Erreichen von Kohärenzstandards typischer Bestandteil ist (s. Teilkap. [6.2.3](#)).

Die Verstehensprozesse des (selbstregulierten) Integrierens sind komplex, und eine zentrale Position weisen die als „*Inferenzen*“ bezeichneten Schlussfolgerungen auf der Basis textueller Informationen bzw. dem Vorwissen auf. Diese Inferenzen gelten als generelle Bedingung des Leseverstehens (Kendeou et al., 2016), und glückendes Inferieren wird auch für das Verstehen von multiplen Dokumenten als essenziell angenommen (Butterfuss et al., 2020). Der Grund dafür besteht darin, dass der Aufbau von Dokumentenmodellen auf einer Vielzahl von Inferenzen beruht. Auf der Ebene des *integrierten mentalen Modells* sind damit vor allem intertextuelle Inferenzen gemeint, die auf verschiedene Weisen erforscht werden. Intertextuelle Verifikationsaufgaben, darunter auch solche Aufgaben mit Fokus auf intertextuelle Konflikte werden in geschlossenen Tests erfasst. In expressiven Aufgaben werden die Inferenzen ex post klassifiziert und verschiedentlich kodiert; in Studien, die dies getan haben, wurden bislang häufiger die Prozesse untersucht. Die Hauptidee dieser Studien mit expressiven Maßen ist, dass die inhaltlichen Schlussfolgerungen vielgestaltig sind und dass dies zu einer Unübersichtlichkeit der Klassifikationen führt. In solchen Studien deutet sich bereits das Potenzial an, die intertextuellen Inferenzen aus Produkt- und Prozessperspektive noch genauer zu bestimmen. Ebenfalls ist deutlich erkennbar, dass Leistungen bei expressiven und rezeptiven Tests sich dahingehend unterscheiden, dass in expressiven Aufgaben weniger Inferenzen nachweisbar sind, als es mit forcierten geschlossenen Tests der Fall ist.

Solche Inferenzen sind mit Blick auf die zweite Ebene des Dokumentenmodells ebenfalls nötig: beim *Intertextmodell*. Auf dieser Ebene, welche der Repräsentation von Dokumenten und quellenbezogenen Metadaten dient, besteht eine Inferenzleistung im Intertextprädikat, also der Inferenz auf der Ebene ganzer Dokumente und ihrem Verhältnis untereinander. Dieses Maß ist trotz seiner bereits seit langer Zeit sowohl bei den Produktmaßen (Perfetti et al., 1999) als auch in prozessualen Modellen (List & Alexander, 2019) benannten Wichtigkeit weitestgehend unterbelichtet. Demgegenüber sind die Quelle-Inhalt-Verbindungen zwischen Metadaten und Inhalt stärker erforscht, hier deuten sich ebenfalls

Differenzen zwischen den expressiven und rezeptiven Tests an, die dem oben genannten Muster bei den intertextuellen Inferenzen entsprechen. Weiterreichende Untersuchungen, die komplexe mentale Modelle wie das Dokumentenmodell – zumal mit Prozessdaten – fokussieren, sind derzeit noch selten durchgeführt worden, doch die bisherigen empirischen Resultate machen deutlich, dass nur geringe Anteile von Personen dazu in der Lage sind, umfassende mentale Modelle über multiple Dokumente aufzubauen.

Eine der Hauptfragen der Multiple-Documents-Literacy befasst sich mit einem Kohärenzproblem und damit einem Hindernis bei den Bemühungen von Leserinnen und Lesern, intertextuelle Kohärenz zu entwickeln (Bråten & Braasch, 2018; Saux, Britt et al., 2021). Denn wenn sich Dokumente inhaltlich widersprechen, wird es für die Leserinnen und Leser erschwert, widerspruchsfreie und logische Vorstellungen über das Gelesene aufzubauen. Im Dokumentenmodell fungiert das Intertextmodell als Struktur, konfligierende Aussagen auf ihren Ursprung zurückzuführen und dies über die Quelle-Inhalt-Verbindungen zu realisieren. Darüber hinaus können Personen über die Evaluation der Quellen Glaubwürdigkeitsurteile bilden, welche sie dafür nutzen, indirekte Plausibilitätsurteile zum Inhalt zu generieren. Diese Zweite-Hand-Einschätzungen bedürfen einer gezielten Vorwissensnutzung und der Sensibilität der Leserinnen und Leser für Nuancierungen der Glaubwürdigkeitszuweisungen. Sie sind ein intensiv erforschter Bereich, in dem sich die Zuwendung zu und Nutzung von Metadaten infolge der intertextuellen Konflikte als gut belegter Kohärenzwiederherstellungsmechanismus erweist. Eine verstärkte Zuwendung zu Metadaten gilt als gut dokumentiert, ebenso die Zuweisung von Glaubwürdigkeit durch die lesenden Personen, die glaubwürdigere Quellen für belastbarer halten. Woran es jedoch mangelt, sind Studien, die prüfen, wie genau intertextuelle Konflikte – zumal in komplexen Settings mit verschiedenen ökologisch validen Dokumenten und Anwendungssituationen – aufgelöst werden. Ähnliches gilt für die inhaltlichen Plausibilitätsurteile bei Erste-Hand-Einschätzungen, die insgesamt ein noch viel größeres Desiderat darstellen. Hierin liegt ein wichtiges, noch kaum systematisch erschlossenes Forschungsfeld.

Der Aufbau komplexer interner Repräsentationen gelingt Personen unterschiedlich gut. Die Leseforschung bzw. deren Ausschnitt, die Multiple-Documents-Literacy-Forschung widmet sich darum den Fragen nach individuellen Einflussfaktoren. Drei von ihnen, die sich allesamt als interne Ressourcen für die Verstehensleistungen modellieren lassen (Britt et al., 2018; van den Broek & Kendeou, 2022), bildeten den Gegenstand der Betrachtung: erstens das inhaltliche Vorwissen zum Gegenstand, zweitens die inhaltlichen (ontischen) Überzeugungen zum Gegenstand und drittens die ontologischen und epistemischen Überzeugungen. Diese drei Merkmale wurden primär produktbezogen erforscht. Das *Vorwissen* schlägt sich in vornehmlich als Inferenzleistung operationalisierten Produktmaßen positiv nieder: Je mehr Wissen eine Person zu einem Gegenstand

hat, desto besser ist sie in der Fähigkeit, schlussfolgernd Zusammenhänge zu erkennen. Bei den Prozessen ist paradoxerweise dieses in den Produkten stabil auftretende Muster in dieser Konsistenz nicht empirisch abgesichert. Mehr noch: Je direkter der Prozess in den Blick genommen worden ist, desto unklarer wirkt der Zusammenhang.

Eine zweite Quelle differenziell ausgeprägter kognitiver Repräsentationen sind *ontische Überzeugungen*, also subjektive Annahmen einer Person dazu, welche Propositionen bei (kontroversen) Themen die Leserin bzw. der Leser für wahrer oder unwahrer hält. Im Normalfall helfen solche ontischen Überzeugungen als unbewusster Benchmark dabei, den Wahrheitsgehalt automatisch und ohne Mehraufwand zu validieren und damit Kohärenzprozesse zu unterstützen. Bei den für das Lesen im Internet typischen Leseanlässen mit konfligierenden bzw. inkonsistenten Aussagen besteht jedoch eine Ambivalenz bei der Validierung. Der Grund dafür ist der „Text-Überzeugungskonsistenz-Effekt“. Er beschreibt das Phänomen, dass Inhalte aus Texten, die zu den ontischen Überzeugungen passen, also konsistent sind, leichter verarbeitet werden und in die internale Repräsentation leichter Einzug halten. Besteht jedoch zwischen Überzeugungen und Inhalt keine Übereinstimmung, erwächst daraus das Risiko, dass ohne aktive lesestrategische Korrektur durch die lesende Person überzeugungsinkonsistente Inhalte von ihr ignoriert werden. Die Folge sind unvollständige, verzerrte Repräsentationen, welche der Komplexität des gelesenen Inhaltes nicht gerecht werden. Empirisch betrachtet ist dieser produktbezogene Effekt bei anspruchsvolleren Verstehensleistungen und nicht bei Repräsentationen einzelner Inhalte beobachtbar. Die Prozesse scheinen sich ebenfalls dahingehend zu unterscheiden, dass überzeugungsinkonsistente Texte kritischer und langsamer verarbeitet werden.

Das dritte Merkmal besteht aus einem Set verschiedener subjektiver Annahmen zu Merkmalen des Wissens (*ontologische Überzeugungen*) und zur Genese von Wissen (*epistemische Überzeugungen*). Solche Überzeugungen lassen sich taxonomisch in verschiedene Dimensionen untergliedern und in Entwicklungsmodellen als phasentypische Konfigurationen dieser Dimensionen beschreiben. Die Forschungsergebnisse, die sich vornehmlich auf ältere Personen als Studienteilnehmerinnen und -teilnehmer sowie auf Studien mit überwiegend inhaltlich konfligierenden Dokumentensets stützen, ergeben klare Muster. Günstig für verschiedene Integrationsleistungen auf der Ebene des integrierten mentalen Modells sind zum einen ontische Überzeugungen, nach denen Wissen komplex und vernetzt ist, statt aus einer Sammlung simpler und isolierter Fakten zu bestehen. Zum anderen haben sich epistemische Überzeugungen als günstig erwiesen, welche Wissen als abstützungsbedürftig auf diverse externe Quellen annehmen und weniger als eine durch die lesende Person selbst entscheidbare Angelegenheit. Dies spiegelt nicht nur die Arbeitsteiligkeit komplexer Gesellschaften mit hohen Spezialisierungsgraden in der Herstellung gesicherter Aussagen wider, sondern nimmt indirekt auch die Fehleranfälligkeit der durch rein internale Prüfprozesse

stattfindenden Validierungen auf, welche im Text-Überzeugungskonsistenz-Effekt bei den ontischen Überzeugungen die Ursache von Verzerrungen bilden. In diesem Zusammenhang verdient Erwähnung, dass die bisher nur sporadisch untersuchten Moderationseffekte von ontischen und ontologischen/epistemischen Überzeugungen vielversprechend wirken. Denn elaborierte epistemische/ontologische Überzeugungen schwächen den Text-Überzeugungskonsistenz-Effekt ab.

Komplexe und kognitiv verarbeitungsintensive Repräsentationen eines Sachverhalts, der Gegenstand mehrerer (multimedialer) Dokumente ist, aufzubauen, ist ein dynamischer Prozess. Dessen Ergebnis, welches als Mehrebenen-Konstrukt in der Theorie als „Dokumentenmodell“ beschrieben wird, setzt umfassende über Inferenzen kombinierte, aber auch separate Repräsentationen voraus. Allein seine Komplexität macht den Verstehensprozess anfällig für Fehlleistungen, die in der Forschung vor allem über die Vollständigkeit der Repräsentationen produktbezogen in den Blick geraten sind (Primor & Katzir, 2018). Dieser Zugang war auch bei dem ersten fokussierten Problembereich leitend. Studien illustrieren zweierlei: Umfassende Repräsentationen von Inhalten, wie sie das integrierte mentale Modell einfordert, sind erstens rar. Zweitens sind Kombinationen von Metadaten mit Inhalten in den Repräsentationen ebenfalls selten ermittelbar, was Probleme bei der Konstruktion von Dokumentenknoten und damit beim Aufbau eines Intertextmodells impliziert. Diese Problemlagen sind in der Forschung gut dokumentierte Muster (Philipp, 2020a).

Wenn beim Integrieren der Aufbau von adäquaten mentalen Repräsentationen das Ziel darstellt, dann ergibt sich mit Blick auf Probleme die Frage, was den Aufbau solcher Repräsentationen erschwert. Die Einflussfaktoren geringes Vorwissen, allzu starke thematische Überzeugungen sowie unterkomplexe epistemische und ontologische Überzeugungen sind bereits aus den Determinanten ableitbar. Tatsächlich hat sich zudem noch die Multimedia-Lernforschung damit befasst, welche Merkmale der Textgestaltung günstig oder ungünstig für das Verstehen sind. Die konsultierte Theorie der kognitiven Belastung sieht vor, dass jenseits der inhaltlichen Merkmale und Faktoren von Texten generische Merkmale von Gestaltungen singulärer und multipler Texte verstehens erleichternd bzw. -erschwerend sind. Die Stellgröße sind zusätzliche Lernerschwernisse durch zusätzliche kognitive Belastung, die zu einer ungünstigen Nutzung des Arbeitsgedächtnisses führen, nämlich zu dessen vermeidbarer Überlastung. Die Multimedia-Lernforschung hat diese Prämissen genutzt und Prinzipien formuliert, von denen drei im Kapitel konsultiert und mit metaanalytisch gewonnenen Erkenntnissen untermauert wurden. Ungünstig ist demnach die Präsentation von überflüssigen Informationen wie verführerischen Details, was einer Verletzung des Kohärenzprinzips entspricht. Verstehenserschwerend sind außerdem Materialien, deren Struktur nicht durch Hervorhebungen leichter erkennbar sind. Dies wird als Verletzung des Signalisierungsprinzips bezeichnet. Schließlich sind räumliche und zeitliche Distanz von Informationen, die sachlogisch

zusammengehören, ein Verstehenshindernis. Das Kontiguitätsprinzip postuliert, dass es Personen schwerer fällt, Inferenzleistungen zu erbringen, wenn es eine zu geringe Nähe dieser Informationen gibt. Dass all diese Prinzipien bei kontingenten Dokumentensets berücksichtigt werden, ist beim ökologisch validen Lesen im Internet eher unwahrscheinlich.

Das Integrieren von Aussagen aus verschiedenen Dokumenten und von Informationen verschiedener Arten wie Metadaten und Propositionen ist ein Erfordernis des digitalen Lesens. Die Verstehensleistungen von verschiedenen Dokumenten, die verschiedenste Perspektiven und Positionen offerieren, verlangt nach analytischen Fähigkeiten seitens der Leserinnen und Leser, intertextuelle Verhältnisse zu bestimmen, (inter-)textuelle Informationen mit Metadaten zu kombinieren und durch evaluative Einschätzungen wo möglich und nötig weiter kognitiv zu transformieren. Die kognitionspsychologische Leseforschung hat dies zu kartieren begonnen. Bei aller Vorläufigkeit und Unvollständigkeit der Ergebnisse ist die zentrale Botschaft ihrer Erkenntnisse, dass mit dem digitalen Lesen anspruchsvolle Leseverstehensprozesse und -produkte den Kern angemessener mehrdimensionaler Repräsentationen bilden. Digitales Lesen wird allein deshalb die Anforderungen an das lesende Individuum und dessen Lesekompetenz drastisch erhöhen.

5 Komponente 4: Aufgabenprodukt herstellen

Darum geht es

Verstehendes Lesen und der Aufbau umfassender mentaler Repräsentationen sind beim Lesen digitaler Dokumente mitunter nicht das zentrale Ziel, sondern eher Bedingungen oder Vorstufen für übergeordnete und darüber hinausgehende Absichten. Insbesondere wenn es um das Lesen zum Zweck des Lernens geht, bilden Aufgabenprodukte das Ziel oder das Zwischenziel des Lesens. Aufgabenprodukte sind kein feststehender Terminus, sondern ein loser Sammelbegriff für mündliche und schriftliche externe Repräsentationen auf der Basis internal verarbeiteter Informationen aus Dokumenten und Texten. Um sie zu erstellen, bedarf es in der Regel einer Transformation der Informationen, damit sie den Zweck jenseits des verstehenden Lesens erfüllen.

Eines dieser Aufgabenprodukte wird in diesem Kapitel fokussiert, nämlich die „Diskurssynthese“. Damit ist das Indienststellen von mehreren Dokumenten für das Erzeugen eines eigenen Textes gemeint. Diskurssynthesen bedingen als hybride Lese- und Schreibanlässe, dass Personen sich nicht nur Inhalte lesend erschließen und angemessen kognitiv repräsentieren müssen, sondern sie auch in einem schriftlichen Produkt mit einer kommunikativen Absicht darlegen können. Hierfür bedarf es mehrerer, im Verbund zusammenwirkender Prozesse, die sich in Anlehnung an eine prototypische Vorgehensweise versierter Schreiberinnen und Schreiber als „Wissenstransformation“ (*Knowledge Transforming*) bezeichnen lassen (5.1). Der Prozessverbund aus Lesen, Schreiben und metakognitiven Selbstregulationsaktivitäten beim Herstellen von Diskurssynthesen und die Zusammenhänge mit Produktmaßen wird zunehmend intensiver erforscht. Eine Folge davon ist Unübersichtlichkeit. Darum wird die Ergebnisentfaltung mehrerer Dutzend Studien mehrstufig vorgenommen. Acht Produktmaße der Diskurssynthesen werden sowohl auf unspezifische (Hinweise, dass Lese- bzw. Schreibprozesse stattfinden) als auch auf spezifische (Spezifikation des Ankerpunkts und Funktion der Prozesse) Prozessindikatoren des Lesens und Schreibens studienübergreifend bezogen. Das Resultat ist ein Katalog von mehreren Mustern, welche Lese- und Schreibprozesse nach aktuellem Kenntnisstand als vorteilhaft für Diskurssynthese-Maße gelten (5.2).

Die Fähigkeit, Diskurssynthesen herzustellen, wird immer mehr in der Forschung mit Individualmerkmalen in Verbindung gebracht. Vier kognitive Maße stehen deshalb im Zentrum der Darstellung, nämlich die inhaltlichen Wissensbestände und Überzeugungen sowie Lese- und Schreibfähigkeiten. Nicht bei allen diesen Variablen gestalten sich die Zusammenhänge jedoch sehr eindeutig, vor allem bei dem Vorwissen nicht. Demgegenüber sind Lese- und Schreibfähigkeiten positiv mit solchen Produktmaßen korreliert, die auf anspruchsvolle Leistungen der Informationstransformation auf Gesamttextebene verweisen. Von Vorteil ist zudem eine thematische Unbefangenheit, da starke inhaltliche Überzeugungen eine ausgewogene Darstellung von thematisch kontroversen Inhalten zu behindern scheinen (5.3).

Die Vielzahl von koordinationsbedürftigen Lese- und Schreibprozessen, die Personen beim Verfassen von Diskurssynthesen selbstreguliert meistern müssen, lassen Fehlanwendungen erwartbar wirken. Aus der Forschung lassen sich tatsächlich musterhafte Probleme rekonstruieren, die bis in die Tertiärbildung hinein, also noch und gerade unter Studierenden bestehen. Zwei dieser Problemfelder ergeben sich daraus, dass zum einen die Personen zu wenige der zentralen Informationen bzw. Informationseinheiten schreibend verarbeiten. Zum anderen haben sie Probleme dabei, gelesene Texte in eine adäquate, dem Schreibziel und dem Inhalt gleichermaßen gerecht werdende Struktur zu überführen. Stattdessen erfolgt in erheblichem Ausmaß eine oberflächliche, einseitige bzw. additive Behandlung des Themas – kurzum: Eine Wissenstransformation leisten die Personen in ihren Diskurssynthesen eben genau nicht (5.4).

5.1 Funktionen

Wenn Personen mit der klaren Absicht lesen, die gefundenen Informationen für einen außerhalb des Lesens liegenden Zweck zu nutzen, spricht man in der Leseforschung von einem aufgabenbezogenen oder instrumentellen Lesen. Dies war bereits ganz zu Beginn dieses Bands im Kapitel 2 und dort speziell beim relevanzbezogenen Lesen (Teilkap. 2.1.3) ein ausführlich betrachtetes Thema. Eine solche Absicht kann darin bestehen, einen eigenen mündlichen oder schriftlichen Text zu erstellen, etwa eine eigene Präsentation auf der Basis des Gelesenen oder eine schriftlich fixierte Sammlung von Gründen für und wider eine Entscheidung. Im hochschulischen und schulischen Kontext gewinnt aktuell eine Variante des Aufgabenprodukts an Bedeutung in Theorie und Empirie, welche sich in diesem Kapitel fokussiert wird: die Diskurssynthese (Nelson, 2001).

Dieser Fokus ist eine wichtige inhaltliche Einschränkung in diesem Kapitel, da es sich um eine Engführung der Darstellung auf die Diskurssynthese als nur einer Variante von schriftlichen, kontinuierlichen Aufgabenprodukten handelt (Knoch & Sitajalabhorn, 2013), die auf der Basis von mehreren Dokumenten mit einem Zielbezug entstehen. Selbstredend gibt es weitere Produkte, zumal im schulischen Kontext, etwa mündliche Vorträge, die man auf der Grundlage von Internetrecherchen herstellt (s. z. B. Bråten, Brante & Strømsø, 2019), oder digitale Concept-Maps als nicht-lineare, grafische und textuelle Informationen enthaltene externe Repräsentation (s. z. B. Hilbert et al., 2008). Dass hier der Diskurssynthese der Vorzug gegeben wird, hat damit zu tun, dass sie durch das Herstellen eines eigenen Textes – eines ihrer zentralen Merkmale –, auch das ausführliche Schreiben für das Lernen als Fähigkeitenbündel erfordert (Chen, 2019; St. Graham et al., 2020). Damit zeigt sich an ihr besonders deutlich der hybride, die Domänenspezifik verwischende Charakter dessen, was es bedeutet, verschiedene Fähigkeiten einzusetzen, um digital präsentierte Informationen gekonnt für sich zu nutzen.

5.1.1 Zielbezogene Transformationen von Informationen beim Herstellen von Diskurssynthesen

Lesen, ob mit analogen und/oder digitalen Texten, mit einzelnen Texten oder multiplen Dokumenten, erfolgt nicht immer zum Zweck, vollständige Texte vom Anfang bis zum Ende zu lesen, sondern es kann sich durchaus mit instrumenteller Absicht vollziehen (Vidal-Abarca et al., 2011; s. Kap. 2), um z. B. in geforderte oder selbstgewählte Aufgabenprodukte zu münden. Aufgabenprodukte, die auf der Lektüre digitaler Texte beruhen, können, müssen aber nicht auf das Verstehen im Sinne einer vollständigen und elaborierten internalen Repräsentation im Sinne dessen hinauslaufen, was im Teilkapitel 4.1.1 als mentales Modell beschrieben wurde. Hier nehmen List und Alexander (2019) eine hilfreiche Differenzierung vor, indem sie eine Ausführungs- und Produktionsphase beim Umgang mit multiplen Texten postulieren. Sie differenzieren dabei zwei Arten von Produkten: erstens kognitive Produkte, die sie als Produkt verschiedener zuvor stattgefundenen (lernstrategischer) Integrationsprozesse und -grade der Ausführungsphase konzeptualisieren, und zweitens die Formulierung bzw. Herstellung von externalen resp. physischen Produkten, welche die internalen Veränderungen aus der Ausführungsphase in der Produktionsphase in ein wie auch immer geartetes Produkt überführen. Mit einem Produkt sind mündliche und schriftliche Produkte gleichermaßen gemeint.

Die entscheidende Differenz, auch zu dem, was Gegenstand im Kapitel 4 war, besteht darin, dass kognitive Produkte eine externalisierte eigene Form der Repräsentation haben, welche transformative Züge der individuellen Aneignung und Verarbeitung aufweist. Hierbei unterscheiden List und Alexander (2019) entlang zweier Dimensionen danach, ob die Produkte darauf abzielen, Gelesenes möglichst vollständig oder selektiv wiederzugeben (z. B. Makropropositionen für Zusammenfassungen vs. verwendbare Informationen für Argumentationen) – dies ist die *Dimension Umfangsabdeckung* –, und inwieweit das Produkt Vorgaben hat (wie die wenig persönlichen Ausdruck zulassende Zusammenfassung ohne Gewichtungen und Wertungen oder ob Schwerpunktsetzungen wie bei einem Forschungsbericht möglich sind) – dies ist die *Dimension Personalisierung*. Beide Dimensionen sind kombinier- und kreuzbar. Der zentrale Punkt ist bei alldem: Es gibt bei den Aufgabenprodukten sowohl Grenzen als auch Freiheitsgrade, entsprechend sind die Aufgabenmodelle und die Merkmale des erwarteten Aufgabenprodukts entscheidend dafür, wie Personen die Aufgabenbearbeitung von Anfang an gestalten (s. Teilkap. 2.1.1).

Die zeitliche Reihung und Dreiteilung, die List und Alexander (2019) in eine bislang an diesem Ort nicht dargestellte a) Vorbereitungs-, b) eine Ausführungs- und c) Produktionsphase vornehmen, deckt sich mit Grundannahmen der Selbstregulationsforschung (s. dazu Teilkap. 6.1.1.3). Laut diesen Annahmen wird der Erfolg beim Herstellen eines Aufgabenprodukts dadurch ermöglicht, dass

Personen zielbasierte Entscheidungen treffen und diese mit den ihnen zur Verfügung stehenden Lernstrategien zu erreichen versuchen. Ein solches Ziel kann zum Beispiel darin bestehen, einen eigenen Text herzustellen. In der zeitlichen Reihung bzw. in der Phasenlogik mit den Spezifika hinsichtlich der Funktionen jeder Phase lässt sich jedoch noch etwas anderes wiederfinden. Denn das, was sich als lesende Aneignung von Gegenständen in der Ausführungsphase abbilden lässt, bildet den Ausgangspunkt für eine darauf folgende Herstellung des Aufgabenprodukts. Dies wird aus der Perspektive des Schreibens über multiple Dokumente als Nachdenken *über* Texte (in der Logik der Ausführungsphase) bzw. als Nachdenken *mit* Texten (in der Logik der Produktionsphase) bezeichnet, was jeweils unterschiedliche Vorgehensweisen und Prozesse zum einen und differente Produkte zum anderen impliziert (Rouet et al., 1996). Das Nachdenken über Texte hat die Funktion, sich den Gegenstand lesend anzueignen und kognitiv zu repräsentieren (s. Kap. 4), das Nachdenken mit Texten stellt die angeeigneten Inhalte gemäß Verwertungsabsicht und damit relevanzbezogen in den Dienst, um das angestrebte Produkt herzustellen (s. Teilkap. 2.2.2).

Nachdenken mit Texten und über Texte und ein verstehendes Lesen zum Zweck des Schreibens, das weist auf ein verschränktes Miteinander von verschiedenen Prozessen hin. Diese Interaktion von Prozessen des verstehenden Lesens lässt sich sowohl produkt- als auch prozessbezogen modellieren. Aus der *Produkt*perspektive ist eine sequenzielle Reihung hilfreich, welche in Abbildung 16 dargestellt ist. Sie wurde ursprünglich als Illustration dafür vorgelegt, welche Parameter zu berücksichtigen sind, wenn man im Sinne einer Diagnostik die Texte von Personen computerbasiert automatisch auswerten will, wenn diese Texte auf der Basis einer Aufgabe und basierend auf mehreren Dokumenten zu erstellen waren (Magliano et al., 2018). Die Darstellung in Abbildung 16 geht von einer zeitlich linearen Logik aus, indem sie die externalen Komponenten wie die Aufgabenstellung ganz links, die beiden exemplarisch zu verarbeitenden Dokumente und schließlich als externale, von der lesenden und schreibenden Person herzustellende Repräsentation in Form des Texts sequenziert. Dazwischen in gestrichelten Kästen dargestellt sind die internalen Repräsentationen, ihrerseits als metakognitives Aufgabenmodell (s. Teilkap. 2.1.1) und kognitiv als Dokumentenmodell (s. Teilkap. 4.1.2.3) bezeichnet. Aus der Produktwarte ist hervorzuheben, dass – ausgehend von den internal verarbeiteten und als Aufgabenmodell repräsentierten Aufgabenmerkmalen – die Personen nicht nur eine singuläre mentale Repräsentation des in den prozessierten Dokumenten enthaltenen Sachverhalts haben, nämlich im Dokumentenmodell. Vielmehr – und das zeigen die rearrangierten grafischen Elemente als Veranschaulichung für Propositionen – müssen Personen aus ihrer eigenen internalen kognitiven Repräsentation des Sachverhalts eine weitere, externale Repräsentation herstellen, welche die Aufgabenerfordernisse adäquat adressiert: ihren Text. Die Forschung hat hierzu bereits diverse Maße vorgelegt, um produktbezogene Indikatoren zu

erfassen, welche über Art, Ebene und Ausmaß propositionaler Veränderungen Auskunft erteilen (Primor & Katzir, 2018).

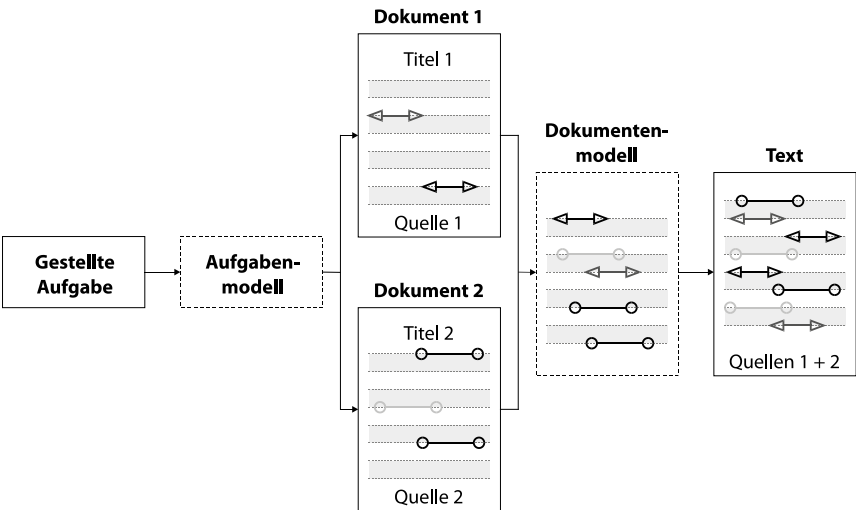


Abbildung 16: Heuristik zur Veranschaulichung von produktbezogenen Transformationen von Informationen bei der aufgabeninduzierten Herstellung eines eigenen Textes (= Diskurssynthese) mit mehreren zu verarbeitenden Dokumenten (Quelle der modifizierten Darstellung: Magliano et al., 2018, S. 503)

Zwischen Dokumentenmodell und eigenem Text und damit zwischen den Produkten, die auf dem Nachdenken über Texte zum Ersten und dem Nachdenken mit Texten zum Zweiten beruhen, vollziehen sich informatorische Transformationsprozesse, die sich in der Schnittmenge von Lese- und Schreibprozessen abspielen. Die damit verbundene Perspektive auf rezeptive und produktive Prozesse widmet sich der Frage, warum und wie die Informationen verändert werden. Eine solche Prozessperspektive ist typisch für die Forschung zu Diskurssynthesen (Mateos & Rijlaarsdam, in Druck; Nelson & King, 2023; Spivey, 1990), sie wird zudem in der Leseforschung prominenter (Rouet & Britt, 2011; Rouet et al., 2017).

Mit dem Ausdruck „Diskurssynthese“ lassen sich laut Nelson (2001, S. 379) jene Lese- und Schreibhandlungen bezeichnen, bei denen eine schreibende Person Inhalte aus verschiedenen Dokumenten integriert, um mit einer kommunikativen Absicht einen neuen eigenen Text herzustellen, welcher eine Thematik verbal behandelt und als Diskurs bezeichnet wird. Dies verlangt eine Syntheseleistung: die Kombination und Abstimmung anderer Diskurse, sprich: anderer Dokumente. Zentral hierfür sind drei sowohl das Lesen als auch das Schreiben betreffende Transformationsprozesse: Auswählen, Organisieren und Verknüpfen. Diese Art des lesend-verarbeitenden Schreibens wird international auch als „Writing from Sources“ (Cumming et al., 2016) bezeichnet, im deutschsprachigen Raum hat

sich durch jüngere schreibdidaktische Entwicklungen der Ausdruck „material-gestütztes Schreiben“ (Philipp, 2017) eingebürgert.

5.1.2 Transformationsprozesse und -logiken beim Herstellen von Diskurssynthesen

Das Erstellen von Diskurssynthesen ist gemäß dem Verständnis aus dem Teilkapitel [5.1.1](#) ein Prozessverbund, der eine vertiefende Betrachtung verdient und an dieser Stelle erfährt. Insbesondere die im Kern des Herstellens von Diskurssynthesen liegende Informationstransformation ist klärungsbedürftig. Ehe dies geschieht, sind aber fünf Merkmale von Diskurssynthesen aus einer prozessualen Perspektive erwähnenswert. Der Katalog ergibt sich aus einer Synopse der Forschungsliteratur (Nelson & King, 2023; Philipp, 2017; Primor & Katzir, 2018; Spivey, 1990; van Ockenburg et al., 2019) und lässt sich knapp folgendermaßen bündeln:

- *Informationstransformation:* Dem Ziel folgend, mit der Diskurssynthese einen eigenständigen Text auf der Grundlage mehrerer Dokumente zu erstellen, ergibt sich der Anspruch, für dieses Ziel die Informationen, Metadaten und Dokumente adäquat zu verwenden. In der Regel genügt es dafür nicht, additiv Informationen in der Sequenz der gelesenen Dokumente in die eigene Diskurssynthese zu übernehmen. Vielmehr müssen diese Informationen ausgewählt, restrukturiert, angereichert, kritisch evaluiert etc. werden und bedürfen einer linguistischen, konzeptuellen sowie rhetorischen Transformation, um das Ziel der Diskurssynthese effektiv zu erreichen.
- *Interdependenz:* Lese- und Schreibprozesse sind interaktiv und rekursiv zu modellieren, d. h. es braucht für den Erfolg bei der Aufgabenlösung beide Prozesse, und zwar in ihrem Verbund. Weder Lesen allein noch das reine Schreiben genügen dem hohen Anforderungscharakter des Herstellens einer Diskurssynthese. Stattdessen bedingen sich beide, und das Verfassen von Diskurssynthesen ist damit ein hybrider Lese- und Schreibprozess mit teils fließenden Grenzen.
- *Rekursivität:* Die Interaktivität von Lesen und Schreiben führt zu einer Kontingenz auf dem Weg zum Produkt, bei der rekursive Schleifen erwartbar sind, indem es beispielsweise zu Relektüren beim Schreiben und zu Veränderungen im entstehenden Text durch Leseprozesse kommt. Zudem ist erwartbar, dass sich der sich durch das Lesen konstituierende Gegenstand der Diskurssynthese während seiner Explikation weitere Klärungsbedürftigkeit erfährt, beispielsweise dann, wenn ein für belastbar gehaltenes Argument weniger persuasiv wirkt, als die lesende Person es zuvor gedacht haben mag.

- *Kontextabhängigkeit*: Die Prozesse des Lesens und Schreibens vollziehen sich nicht akontextuell, sondern sind allein aufgrund der Tatsache der external vorliegenden Dokumente und den in ihnen enthaltenen Informationen, die hochkontingent erst zu lokalisieren und in ihrem Verhältnis zu bestimmen sind, stark an den Kontext gekoppelt. Hinzu kommt, dass auch die Aufgabenstellung bzw. das Produkt samt seinen mehr oder minder spezifizierten Merkmalen und das ihm vorgelagerte Aufgabenmodell als Einflussvariablen fungieren. Zu berücksichtigen ist ferner, dass die entstehende Diskurssynthese selbst über Rückkopplungen und Teil des Kontexts mitbeeinflusst, welche Aktivitäten eine lesende und schreibende Person noch zu absolvieren hat.
- *Selbstregulationsnotwendigkeit*: Die Komplexität des Prozessverbunds mit- samt seiner Zielbezogenheit und grundsätzlichen Offenheit führen zu einer hohen metakognitiven Koordinationsnotwendigkeit. Darum sind planerische, überwachende und die Lese- und Schreibaktivitäten orchestrierende metakognitive Fähigkeiten nötig – also die Fähigkeit zur Selbstregulation (s. Kap. 6).

Diese fünf Merkmale aus der Liste oben hängen insofern zusammen, als sie in ihrem Gesamt dabei helfen, das aus kognitiver Perspektive aufwändige Unterfangen zu konturieren. Sie machen auf die Verschränkt- und Bedingtheit sowie die Kontingenz und Steuerungsbedürftigkeit aufmerksam, welche das Erstellen einer Diskurssynthese zu einem Aspiranten lernstrategischer Vorgehensweisen im Verbund avancieren lassen. Im Zentrum des Herstellens einer Diskurssynthese stehen die kognitiven Transformationsprozesse, die im nachfolgenden Teilkapitel spezifischer betrachtet werden.

5.1.2.1 Auswählen, Organisieren und Verknüpfen – eine Prozesstrias der Informationstransformation

Textuelle Informationstransformation ist Merkmal und Herausforderung zugleich bei dem Herstellen von Diskurssynthesen für das lesende und schreibende Individuum (Nelson, 2001). *Informationstransformation* meint, dass Leserinnen und Leser sowohl beim Lesen als auch beim Schreiben einerseits die aus den Dokumenten stammenden Informationen, andererseits die aus dem eigenen Langzeitgedächtnis abrufbaren Wissensbestände so aufeinander beziehen, dass das angestrebte Produkt erreicht wird. Um dies zu erreichen, ist ein Prozessverbund nötig. Prononciert hat diesen Verbund Spivey (1990) beschrieben. Sie geht davon aus, dass es sich um domänenspezifische Ausprägungen von drei generischen Prozessen der Informationsverarbeitung handelt, die grundsätzlich jeweils für das Lesen und Schreiben benötigt werden und dort komplementäre Funktionen erfüllen. Diese drei Prozesse sind, zuoberst für das Lesen und zuunterst für das Schreiben, in Abbildung 17 gesammelt dargestellt. Die Doppelpfeile

in der Darstellung visualisieren die zu Beginn des Teilkapitels angesprochene Interdependenz und Rekursivität; die Kontextsensitivität und metakognitive Selbstregulation spart die Heuristik aus.

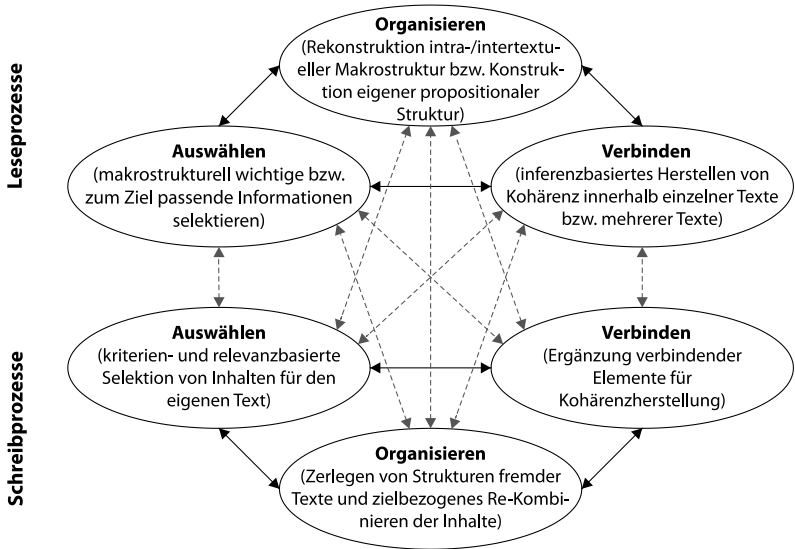


Abbildung 17: Heuristische Darstellung der Transformationsprozesse für das Herstellen einer Diskurssynthese (eigene Darstellung, basierend auf Nelson & King, 2023)

Die drei Prozesse, die beim Lesen und Schreiben als Teil des Herstellens von Diskurssynthesen benötigt werden, sind im Einzelnen:

- *Informationen auswählen.* Hierbei sind bezogen auf das *Lesen* zwei Arten grundsätzlich denkbar. Zum einen ist das Auswählen jener Propositionen damit gemeint, die in der Hierarchie des Textes als übergeordnete Makropropositionen gelten, weil sich implizit oder explizit diverse andere Propositionen im Netzwerk der Propositionen auf der Textbasis auf sie beziehen. Dies entspricht jenen Prozessen, die für die Integrationsprozesse auf der Ebene einzelner Texte bereits in Teilkapitel [4.1.2](#) en detail beschrieben wurden. Das Auswählen kann sich zudem auf jene Informationen erstrecken, die der Verwertungsabsicht entsprechen, die nicht (allein) auf makropropositionaler Wichtigkeit beruhen, sondern den Relevanzkriterien. In dieser Verwertungsabsicht, die bereits in das Schreiben hineinragt, zeigt sich der fließende Übergang, denn beim *Schreiben* der Diskurssynthese gilt es, Informationen zu selektieren, die den Relevanzkriterien für den Zweck des Schreibens entsprechen. Im Sinne der Schreibprozesse handelt es sich um eine auf dem Lesen basierende Inhaltsgenerierung (Alamargot & Chanquoy, 2001). Wer eine Diskurssynthese verfassen will, braucht dafür die entsprechenden

Inhalte aus den zu verarbeitenden Dokumenten, die im hybriden Lese- und Schreibprozess auszuwählen sind.

- *Informationen organisieren.* Mit diesem Prozess ist bezogen auf das *Lesen* die Rekonstruktion der Makrostruktur von Propositionen gemeint, vor allem jene, die über die einzelnen Dokumente hinausreicht und im Kernbereich des integrierten mentalen Modells liegt. Es geht darum, thematische Bezüge zu erkennen bzw. wo nötig selbst zu schaffen, indem die Leserinnen und Leser intertextuelle Kohärenz stiften. Dies kann zudem eigenen Wissensstrukturen entsprechen, die nicht zwangsläufig mit jener Struktur gleichzusetzen sind, welche der thematischen Entfaltung in den einzelnen Dokumenten folgt. Genauso gut ist es möglich, dass Leserinnen und Leser von vornherein mit einer eigenen Strukturierungsabsicht und -form an die Texte herantreten und passende Informationen über das Auswählen an entsprechende Stellen in ihrer Struktur einfügen, etwa als Details in dimensionalen Vergleichen von Objekten, welche in mehreren Dokumenten beschrieben werden. Das Organisieren benötigen Personen aus der Sicht des Lesens ebenfalls dazu, um für Dokumentenknotten im Intertextmodell des Dokumentenmodells (s. Abbildung 13 in Teilkap. 4.1.2.3) stark aggregierte Hauptaussagen von Dokumenten zum einen zu extrahieren und die darauf aufbauenden Intertextprädikate zum anderen als Strukturgeber für Dokumente als repräsentationale Einheiten zu bilden. Bereits unter dem Punkt der eigenen Makrostruktur beim Lesen ist ein Punkt angesprochen, der für das *Schreiben* von Belang ist. Schreiberinnen und Schreiber müssen Informationen nicht nur selektieren, sondern auch in eine Sequenz und damit in eine eigene Makrostruktur überführen. Eine solche Struktur dient dem angestrebten kommunikativen Zweck und lässt – wo textlinguistisch präformiert – eine textsorten- bzw. genrespezifische Abfolge von Elementen erwarten. Solche Textstrukturen dienen als Gerüst für die konkret ausgewählten Inhalte, die kohärent aufeinander zu beziehen sind. Dies wird in der Terminologie der Schreibprozesse als Inhaltsorganisation als Teil des Planens bezeichnet (Alamargot & Chanquoy, 2001).
- *Informationen verbinden.* Der Prozess des Kohärenzaufbaus, sei es beim Lesen, sei es beim Schreiben, kommt ohne vorwissensbasiertes Inferieren und Integrieren nicht aus. Darum bilden die in Teilkapitel 4.1.2.1 ausführlicher behandelten Prozesse des schlussfolgernden Denkens beim Lesen und gespiegelt dazu das Anreichern gelesener Informationen mit Gedächtnisinhalten beim Schreiben weitere essenzielle Prozesse des Verfassens von Diskurssynthesen. In beiden Fällen geht es darum, die Propositionen mittels eigener Wissensbestände auf logisch nachvollziehbare Weise miteinander zu kombinieren.

Was bei der soeben dargestellten Trias von Prozessen auffällt, ist, dass typische Bestandteile des Lesens multipler Dokumente – darunter das *Sourcing* und generell das *Evaluiieren von Aussagen* (Cho & Afflerbach, 2017) – darin nicht

vorkommen. Selbst in den einschlägigen Veröffentlichungen zur Projekttrias (Nelson & King, 2023; Spivey, 1990) bildet – mutmaßlich über den impliziten Fokus auf das Integrieren erklärbar – das Sourcing einen blinden Fleck. Das erscheint aus zwei Gründen korrekturbedürftig. Erstens ist über das Dokumentenmodell (aus Teilkap. [4.1.2.3](#)) aus lesetheoretischer Sicht das Sourcing bzw. die sich aus ihm speisenden Repräsentationen auf der Intertextmodell-Ebene anzuführen. Nach diesem Modell, welches seinerseits eine idealtypische Variante kognitiver Repräsentation darstellt, sind Metadaten konstitutiv. Das gilt umso mehr, als neuere Lesetheorien betonen, dass insbesondere bei intertextuellen inhaltlichen Konflikten dem Sourcing eine kohärenzstiftende Funktion zukommt, indem das Zustandekommen der Konflikte über Metadaten erklärt und über die Evaluation von Metadaten teils auch aufgelöst werden kann (Braasch & Kessler, 2021; Braasch & Scharrer, 2020; Rouet & Britt, 2022). Zweitens, und dies speist sich aus der Forschung als empirische Perspektive, gibt es gute Gründe, das Sourcing auch in Diskurssynthesen zu fokussieren. Die Grundlagenforschung zeigt, dass Sourcing im Rahmen von Aufgaben mit Diskurssynthesen sowohl prozessual als auch produktbezogen erforscht wird (Mateos & Rijlaarsdam, in Druck) und dass es Zusammenhänge von Sourcing-Produktmaßen und weiteren Leseprozessen einerseits sowie Sourcing-Prozessindikatoren und weiteren Produktmaßen andererseits gibt (Philipp, 2022b). Zudem ist der in der Theorie postulierte Ausgangspunkt für das Sourcing, nämlich das Vorhandensein intertextueller Konflikte, in einer Großzahl der Studien in Grundlagen- (Primor & Katzir, 2018) und Interventionsforschung (Barzilai et al., 2018; Philipp, 2019) oftmals gegeben: Sich inhaltlich widersprechende Dokumente innerhalb vorgegebener Sets dominieren deutlich die bisherige Forschung. Sie implizieren damit das Sourcing, aber Sourcingmaße werden bei Weitem nicht in allen Studien ausgewertet. Es ergibt sich damit ein gesamthaft blinder Fleck, der sich aus den noch zu unverbundenen Forschungsfeldern der Lese- und Schreibforschung erklären lässt (Bråten et al., in Druck; St. Graham, 2020). Dies spricht insgesamt dafür, Sourcing nicht nur beim Lesen multipler Dokumente, sondern auch beim Schreiben darüber noch stärker zu berücksichtigen, um nicht nur die Integration von Inhalten, sondern auch die Nutzung von Dokumenten als soziale Artefakte bzw. Diskursteilnehmer in einer gesamthaft kontextuell zu verortenden intertextuellen Kommunikation adäquater abzubilden (Britt et al., 2013).

5.1.2.2 Knowledge Transforming – eine kognitiv anspruchsvolle Vorgehensweise von Schreiberinnen und Schreibern

Das Verfassen von Diskurssynthesen bedarf umfassender Informationstransformationen. Als Hybrid von Lese- und Schreibprozessen sind aufeinander bezogene Prozesse im rezeptiven und produktiven Umgang mit Informationen notwendig. Die Perspektive des Schreibens ist bereits angeklungen, aber sie bedarf einer

intensiveren Zuwendung. Darum soll zunächst einmal etwas Grundlegendes zu Schreibprozessen angemerkt werden, ehe es um eine als „Knowledge Transforming“ bezeichnete Vorgehensweise gehen wird.

Schreiben ist ein Prozessverbund, der dazu dient, aktiv und schriftlich Bedeutung zu fixieren und zu kommunizieren, also ein Schreibprodukt herzustellen, welches seine kommunikative Funktion zu erfüllen imstande ist (Alamargot & Chanquoy, 2001). Ähnlich wie beim Lesen sind dafür Prozesse in ihrem Verbund nötig, wobei sich drei Hauptprozesse beim Schreiben als Konsens in der Schreibforschung herauskristallisiert haben:

- Der erste Prozess ist das *Planen* als adressaten- und zielorientierte Klärung von Kommunikationskontext, angestrebten Leserinnen und Lesern sowie möglichen Inhalten und deren Sequenz im eigenen Text.
- Der zweite Prozess wird als *Verschriften* oder auch als *Transkription* bezeichnet und beinhaltet das sprachliche Anreichern und Linearisieren geplanter Inhalte unter Berücksichtigung sprachlicher (z. B. grammatischer) und schriftsprachlicher (wie orthografischen) Regeln in graphomotorischen Aktivitäten wie dem Schreiben mit der Hand oder per Tastatur.
- Der dritte Prozess ist das *Revidieren* und fungiert als Reparaturmechanismus für das Schreiben. Er beinhaltet das Diagnostizieren von Problemen an Texten, die Entscheidung über die Modifikationsnotwendigkeit und schließlich die Ausführung der Veränderung am Produkt (Alamargot & Chanquoy, 2001).

Diese Prozesse sind auf diverse Wissensbestände – einige schreibunspezifisch wie das Welt- und Vorwissen, andere schreibspezifisch wie das Textsortenwissen oder das Wissen über Schreibstrategien – angewiesen und müssen, weil sie als Gesamtprozess des Schreibens untereinander interagieren und prinzipiell jederzeit im Schreibprozess auftreten können, metakognitiv gesteuert werden (Hacker, 2018). Anders gesagt: Kognitive Prozesse wie das Planen, Formulieren und Revidieren, welche den Text als Produkt beeinflussen, bedürfen der metakognitiven Selbstregulation (Zimmerman & Risemberg, 1997), was eine bedeutende Parallele zum digitalen Lesen ist (s. dazu Kap. 6). Das Geschehen wird als dynamisches Mehrebenen-Konstrukt angenommen (Hacker, 2018): Kontrollierende, das Planen, Formulieren und Revidieren steuernde Prozesse werden von einer Metaebene der Metakognition aus mit einem Zielbezug initiiert und auf der Objektebene von der Schreiberin bzw. dem Schreiber ausgeführt, um dadurch entweder textbezogene Kognitionen als internes Zwischenprodukt oder Text als externes Aufgabenprodukt selbst zu kreieren. Metakognitive Prozesse steuern demnach die kognitiven Prozesse. Über Rückkopplungen ist die Gegenrichtung ebenfalls angedacht: Indem Personen die Ergebnisse der Prozesse und die entstehenden Produkte wahrnehmen und anhand eines angestrebten Ziels beurteilen, fungieren Überwachungsprozesse als kontinuierlicher Sicherungsmechanismus, der in

Verbindung mit erkannten Veränderungsnotwendigkeiten auf der Objektebene rückkoppelnd zu metakognitiven Kontrollprozessen führt. Insofern spielen beim Schreiben metakognitive und kognitive Prozesse dynamisch zusammen, um sich einem definierten Zielzustand – dem Textprodukt – selbstreguliert zu nähern.

Dieser Zielbezug ist in Modellen des Schreibens mehr oder minder prominenter Bestandteil der Modelle, die sich in der Tradition des Schreibens als Problemlöseprozess sehen (Alamargot & Chanquoy, 2001; Chen, 2019), darunter im Herstellen einer Diskurssynthese (Nelson & King, 2023), welche bereits durch die mehr oder minder spezifizierte Schreibaufgabe eine Zielvorstellung nahelegt, die das schreibende Individuum im Aufgabenmodell repräsentiert und dieses als metakognitive Steuereinheit nutzt. Das gilt nicht ausschließlich, aber doch ebenfalls für Vorgaben zur Zieltextsorte. Welchen textlinguistischen Anforderungen und Konventionen der zu schreibende Text zu genügen hat und wie damit der jeweils konkrete Inhalt aus multiplen Dokumenten kohärent entfaltet wird, ist Gegenstand empirischer Forschung (s. dazu den Kasten „Illustrierendes Beispiel 7“ in Teilkap. 2.2.2.4). Eines der zentralen Ergebnisse aus Studien mit dem Vergleich von Personen, die unterschiedliche Textsorten schrieben, bestand in dem übergreifenden Muster, dass sich bei Maßen, die auf Textprodukten (in expressiven Aufgaben) basierten, Unterschiede ergaben: Bei Argumentationen war es beispielsweise auffällig, dass sich in den Produktdaten relative Überlegenheiten im Vergleich zu vor allem informierenden Zieltextsorten ergaben, wenn a) die zu verarbeitenden Dokumente ein komplementäres intertextuelles Verhältnis aufwiesen und b) die kognitive Leistung besonders anspruchsvoll war, indem es beispielsweise nötig war, Inferenzen zu bilden und das Verhältnis von Aussagen untereinander darzulegen (Philipp, 2022a). Starke vorwissensbasierte Anreicherungen (Elaborationen) und die Restrukturierung von Propositionen für die Eignung des zu verfassenden Zieltexts scheinen also eine Überlegenheit beim Argumentieren nach sich zu ziehen. Damit stellt sich die Ursachenfrage.

Eine mögliche Ursache für dieses Muster besteht darin, dass das Schreiben von Diskurssynthesen deutliche Merkmale eines Vorgehens beim Schreiben aufweist, welches als „*Knowledge Transforming*“ bezeichnet wird (Flower, 1990; Scardamalia & Bereiter, 1987). Das Transformieren des Wissens, welches namensgebend ist, vollzieht sich dabei als ein bidirektionaler Prozess zwischen zwei Problemlöseräumen, einem inhaltlichen und einem rhetorischen. Problemlöseräume haben die zentrale Eigenschaft, dass sie dazu dienen, einen angestrebten Zustand sukzessive zu erreichen. Aus Sicht des Schreibens ist entscheidend, dass das Erreichen des Ziels über interne und iterative Rückkopplungen des Überwachens und Kontrollierens vorangetrieben wird, also selbstreguliert erfolgt (s. Teilkap. 6.1.2.2). Beides lässt sich zunächst analytisch voneinander trennen, und beides trägt die drei Prozesse eines selbstregulierten Planens, Verschriftens und Revidierens in sich:

- Der *inhaltliche Problemlöseraum* bezieht sich auf das Vorwissen und im Fall der Diskurssynthese auf die Inhalte der Dokumente. Die im inhaltlichen Problemlöseraum zu klärende Frage lautet: Was ist das, was die Person in ihrem Text kommunizieren will? Dies erfolgt im Sinne der Problemlösung als Annäherung, indem Personen prüfen, ob sie beim Schreiben ihre geplanten Inhalte im Text realisiert haben.
- Der *rhetorische Problemlöseraum* betrifft die kognitive Repräsentation der Kommunikationsabsicht. In ihm wird die Frage fortlaufend behandelt: Ist der aktuelle Zustand des Textes dazu geeignet, das kommunikative Ziel zu erreichen – und falls nicht: müssen dann Text und/oder das Ziel verändert werden? Es geht hierbei darum, verschiedene Ziele des Schreibens zu überwachen und ggf. zu modifizieren.

Diese beiden Problemlöseräume müssen interagieren, damit es zu einer Wissens-transformation kommt. Das bedeutet, dass inhaltliche Probleme auch aus der Sicht der Rhetorik und rhetorische Probleme ebenfalls aus der Warte des Inhalts zu klären sind. Wer beispielsweise schriftlich argumentierende Diskurssynthesen verfasst und inhaltlich zutreffende, aber die Argumentation an der aktuellen Stelle nicht voranbringende Inhalte aufschreibt, ist gefragt, dieses Problem aus persuasiver Sicht zu lösen, indem er oder sie die Inhalte an anderer, besserer Stelle platziert. Solche Teilziele, die sich im Prozess im Konkreten stellen, führen zu Modifikationen an Zielen, Inhalten und Beziehungen zwischen beidem – dies macht das Vorgehen gemäß dem „Knowledge Transforming“ aus, und dies macht es so anspruchsvoll. Es handelt sich um komplexe interagierende Prozesse, welche metakognitiv reguliert werden müssen. Darum haben Bereiter und Scardamalia (1987) das „Knowledge Transforming“ als Vorgehensweise von geübten Schreiberinnen und Schreibern bezeichnet, die zur Regulation solcher Prozesse aufgrund ihrer umfassenderen Wissensbestände und Schreiberfahrung erst in der Lage sind. Flower (1990) ergänzt dies um eine Reperspektivierung: Sie postuliert, dass „Knowledge Telling“ und „Knowledge Transforming“ als Kontinuumsendpunkte aufzufassen seien und damit Abstufungsgrade zwischen beiden Punkten vorliegen.

Diskurssynthesen stellen die lesenden und schreibenden Personen als hybride Prozesse einforderndes Aufgabenprodukt damit vor Voraussetzungen, welche sich als theoretisch kompatibel erweisen zum „Knowledge Transforming“. Da die Diskurssynthese auf der einen Seite einen Sammelbegriff für unterschiedliche Texte und Textsorten darstellt und darum ohnehin Klärungen im Sinne dessen, was im rhetorischen Problemlöseraum eingefordert wird, nötig sind und weil auf der anderen Seite das Nachdenken über und mit Dokumenten inhaltliche Klärungen bedingt, kann das Verfassen einer Diskurssynthese als umfassender Problemlöseprozess mit vielen Freiheitsgraden modelliert werden. Deshalb bedarf

es diverser strategischer Klärungen, um das eigene Wissen und die Inhalte aus Dokumenten angemessen zu transformieren (Primor & Katzir, 2018).

5.2 Prozesse und Produkte

Diskurssynthesen setzen als Produkte voraus, dass Personen Lese- und Schreibprozesse angemessen nutzen. Dieses Teilkapitel widmet sich Produkten und Prozessen der Diskurssynthese-Herstellung gleichermaßen und stellt die Zusammenhänge von Prozessen und Produkten ins Zentrum. Dazu werden acht Produktmaße vorgängig erläutert und bereits die acht Lese- und sechs Schreibprozessgruppen überblicksartig präsentiert (5.2.1). Es folgen anschließend für Leseprozesse (5.2.2) und Schreibprozesse (5.2.3) ähnliche Entfaltungen der Thematik: Ausgehend von einer Explikation der Prozessmaße werden diese dann von der empirischen Basis ausgehend mit Produktmaßen in Bezug gesetzt. Der folgende Darstellungsschritt geht auf musterhafte Zusammenhänge ein, indem das gemeinsame Auftauchen von Verhaltensweisen im Lesen und Schreiben fokussiert wird, sei es in Bezug auf Personengruppen, das allgemeine Vorgehen oder aber die Veränderung von Lese- und Schreibaktivitäten im Zeitverlauf (5.2.4). Aus den vielen Einzelbefunden wird im finalen Schritt ein prototypisch günstiges Vorgehen abstrahiert (5.2.5).

5.2.1 Darstellerische Vorbemerkungen

Die empirische Erforschung von Prozessen des Lesens und Schreibens bei dem Herstellen von Aufgabenprodukten im Allgemeinen und bei Diskurssynthesen im Besonderen verfolgt zunehmend die Frage, welche kognitiven Prozesse im Umgang mit Texten mit Produktmaßen der Aufgabenprodukte korrespondieren. Aus diesem Forschungszweig liegen inzwischen diverse Studien vor, sodass sich deren Ergebnisse zu mehr oder minder deutlichen Mustern verdichten lassen. Dies erfolgt in diesem Teilkapitel selektiv für die Diskursanalysen. Ebenfalls selektiv ist das Teilkapitel insofern, als es sich allein auf Studien stützt, in denen der Zusammenhang von Prozessen und Produkten bei Diskurssynthesen erforscht wurde. Hierfür gibt es einerseits eine ausreichend breite empirische Basis, andererseits hat gerade die produktbezogene Forschung eine derart divergente Zahl von Maßen und Indikatoren hervorgebracht (Primor & Katzir, 2018), dass allein deshalb eine Komplexitätsreduktion durch Eingrenzungen nötig wirkt.

5.2.1.1 Ein grafischer Überblick über die zentralen Konstrukte in der empirischen Forschung zum Zusammenhang von Prozessen und Produkten

Zur Strukturierung der Inhalte dieses Teilkapitels dient die Sammlung an Konstrukten, die in Abbildung 18 überblicksartig zusammengefasst sind. Es handelt sich um eine induktiv aus den Primärstudien gebildete Kategorisierung, welche auf eigenen Vorarbeiten in Bezug auf Leseprozesse und Produktmaße der Diskurssynthesen beruht (Philipp, 2022b). Die Darstellung der Konstrukte aus Abbildung 18 enthält bei den Produkten unten in der Mitte diverse Kategorien bislang empirisch untersuchter Maße.

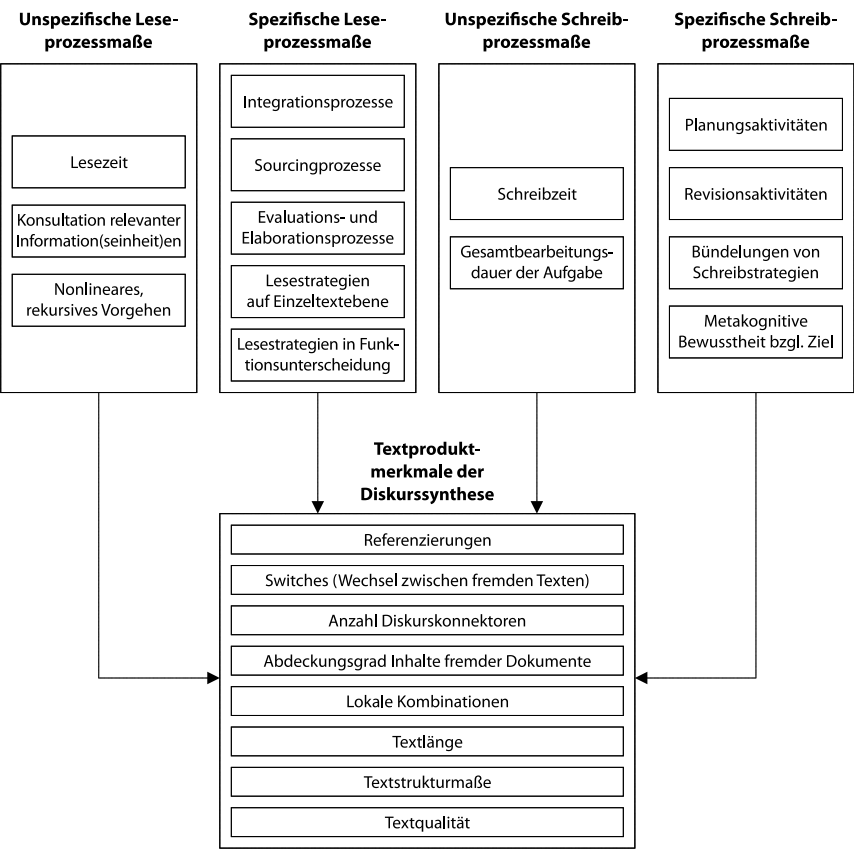


Abbildung 18: Überblick über aktuell empirisch erforschte lese- und schreibbezogene Prozessmaße und Produktmaße beim Verfassen von Diskurssynthesen

Bei den Prozessdaten, links oben dargestellt für das Lesen, rechts oben für das Schreiben, differenziert die Darstellung unspezifische und spezifische Prozessmaße. Unspezifische Maße wie Bearbeitungszeiten geben relativ ungenau an, dass

Personen sich mit dem Lesen bzw. Schreiben befasst haben. Spezifische Maße, deren Gewinnung verschiedene Methodiken wie lautes Denken, genauere Trackingindikatoren oder retrospektive Auskünfte voraussetzt, sind demgegenüber präziser und teilweise aufwändiger in der Auswertung. Während unspezifische Prozessindikatoren demgemäß verdeutlichen, *dass* Prozesse stattfinden, erteilen spezifische Prozessmaße Auskunft darüber, *welche* Prozesse sich ereignet haben. Beide Arten von Daten ergänzen sich und ergeben ein Gesamtbild.

5.2.1.2 Überblick: Welche Produktmaße wurden in Studien erfasst?

Die jeweiligen Prozessdaten für das Lesen und Schreiben werden im Folgenden mit den Produktmaßen ins Verhältnis gesetzt, wobei die Prozesse den Ausgangspunkt bilden. Aus Gründen der besseren Darstellbarkeit und Verständlichkeit werden die Lese- und Prozessmaße in den Teilkapiteln [5.2.2](#) (für das Lesen) sowie [5.2.3](#) (für das Schreiben) beschrieben. Die acht Produktmaße, auf die beide Teilkapitel Bezug nehmen, werden demgegenüber bereits jetzt betrachtet. Die Anzahl von Gruppen der Produktmaße, die ihrerseits übergeordnete Kategorien bilden, stellen nur einen Ausschnitt dar. Die Forschung zu den Diskurssynthesen ist aktuell von einer großen Heterogenität der Maße bei den Produkten geprägt (Nelson & King, 2023; Philipp, 2022a, 2022b; Primor & Katzir, 2018), sodass aus Gründen der Komplexitätsreduktion auch nicht jedes Detail ausgeleuchtet werden kann. Stattdessen werden in den Teilkapiteln [5.2.2–5.2.4](#) die großen Linien und übergeordneten Muster bei den Zusammenhängen herausgearbeitet.

1) Referenzierungen (inkl. Zitationen)

Die Referenzierungen sind die Benennungen der Herkunft von Informationen bzw. Verwendung von Informationen, die auf einzelne Dokumente zurückgeführt werden können; insofern dienen sie dazu, Quellen und Inhalte erkennbar miteinander zu verknüpfen und auszuweisen.

Am häufigsten untersucht wurden die Anzahl expliziter Hinweise auf bzw. die Zitationen schriftlich verarbeiteter Dokumente in der eigenen Diskurssynthese (Anmarkrud et al., 2014; Barzilai et al., 2023; Bråten, Brante & Strømsø, 2018; Du & List, 2020; Lee et al., in Druck; Leroy et al., 2023; List & Alexander, 2018b, 2020; List et al., 2017; List & Du, 2021; List, Du, Wang & Lee, 2019; List, Oaxaca et al., 2021; List, Stephens & Alexander, 2019; Salmerón, Gil & Bråten, 2018b; Strømsø & Bråten, 2014; Strømsø et al., 2013). Teils wurde nur ausgewertet, ob es überhaupt mindestens eine Referenzierung gegeben hat (Schoor et al., 2023; Griffin et al., in Druck). Daneben wurden die Anzahl impliziter Hinweise auf die Anzahl schriftlich verarbeiteter Texte im eigenen Text (Anmarkrud et al., 2014) sowie die Anzahl von im- oder expliziten Benennungen von verarbeiteten Dokumenten mit Verknüpfungen zum Inhalt dieser Dokumente analysiert (Anmarkrud et al.,

2014). Nur wenige Studien (Kiili et al., 2021; Mason et al., 2017; Mason, Scrimin, Zaccoletti et al., 2018) bildeten hingegen Gesamtindizes des Sourcings, in denen Referenzierungen ein Bestandteil unter mehreren war.

2) *Switches*

Die Switches sind Wechsel zwischen den verwendeten Informationsquellen im eigenen Text und fungieren als Indikator für die intertextuelle Vernetzung der Inhalte. Die Anzahl der Wechsel von Informationen, die nicht aus dem gleichen Dokument stammen, gilt als Hinweis auf die Nutzung von intertextuellen Inhalten aus verschiedenen Dokumenten, und wird entsprechend ausgezählt (Bråten, Brante & Strømsø, 2018). Je mehr Wechsel vorgenommen wurden, als desto besser integriert gilt das Textprodukt hinsichtlich der Nutzung der multiplen Dokumente.

3) *Diskurskonnektoren*

Die Diskurskonnektoren bezeichnen die Konjunktionen bzw. Konnektoren, die das semantische Verhältnis von Informationen grammatisch indizieren. Operationalisiert wurde dies, indem die Anzahl der Diskurskonnektoren bestimmt wurde (Du & List, 2020; List, Du, Wang & Lee, 2019). Dies wurde teils noch auf adversative Konjunktionen enggeführt und auf Ausdrücke, die grammatisch nahelegen, dass auf zwei Dokumente Bezug genommen wurde (Schoor et al., 2023).

4) *Abdeckungsgrad an Inhalten*

Der Abdeckungsgrad an Inhalten kennzeichnet das Ausmaß der übernommenen Propositionen aus den gelesenen Dokumenten in der eigenen Diskurssynthese und bildet einen Indikator für die thematische Breite. Eine Gruppe von Variablen aus den Texten stützt sich auf textbasierte Indikatoren, darunter die Anzahl explizit oder implizit übernommener Informationen aus den gelesenen Texten (Ackerman, 1991; Barzilai et al., 2023; Barzilai et al., 2015; Bråten, Brante & Strømsø, 2018; Cerdán & Vidal-Abarca, 2008; List & Lin, 2023; Peterson & Alexander, 2020; Spivey, 1992), die Anzahl korrekter und inkorrektter Konzepte im eigenen Text (Castells et al., 2023; Wiley et al., 2009) und die Anzahl textbasierter Sätze (List & Du, 2021) bzw. Anzahl der Ideeneinheiten in Texten (Ackerman, 1991), darunter auch als negative Ausprägung in Form von wortwörtlichen Übernahmen von Inhalten aus Texten (Castells et al., 2023; Yang & Plakans, 2012) oder falschen Informationen (Gil et al., 2011); weiterhin wurden die Häufigkeiten von textbasierten Begründungen als Indikator genutzt (Du & List, 2021; List, Stephens & Alexander, 2019). Eine weitere, einen anderen Bezugspunkt wählende Operationalisierung besteht in der Anzahl der verwendeten Dokumente im eigenen Text (Barzilai

et al., 2023; Barzilai et al., 2015). Ein anderes, computerlinguistisch grundiertes Maß ist die thematische Ähnlichkeit der Diskurssynthese zu den Inhalten der zu verarbeitenden Texten (Raković, Fan et al., 2022; Raković et al., 2023).

5) *Lokale Kombinationen*

Lokale Kombinationen beziehen sich auf intra- bzw. intertextuelle Verknüpfungen von Informationen aus den gelesenen Dokumenten und gelten als Indikator für intra- resp. intertextuelles Integrieren. Als Maße hierfür kamen zum Einsatz die Anzahl korrekt gebildeter intertextueller Inferenzen (s. Teilkap. [4.1.2.1](#) und [4.2.1.1](#)) zwischen den Inhalten gelesener Dokumente (Barzilai et al., 2023; Gil et al., 2011; Leroy et al., 2021, 2023; Linderholm et al., 2014, Studie 1; List & Lin, 2023; List, Oaxaca et al., 2021), darunter zu Belegen bei der Einschätzung von kontroversen Themen (Du & List, 2021), die Anzahl von Aussagen mit intertextuellen Integrationen (Latini et al., 2019; Lee et al., in Druck; List & Du, 2021; List, Du, Wang & Lee, 2019, Studie 1; Mason, Tarchi et al., 2022) sowie die korrekte Wiedergabe von intertextuell behandelten Hauptthemen (List & Alexander, 2020). Hinzu kommen intratextuelle und andere, vorwissensbasierte Inferenzen, die in Studien mit Prozessdaten bislang selten untersucht worden sind (Gil et al., 2011).

6) *Textlänge*

Die Textlänge bei den Diskurssynthesen basiert auf quantitativen Auszählungen verschiedener Bestandteile von Texten. Dies können textsortenspezifische Bestandteile sein wie die Anzahl von Argumenten in der eigenen argumentativen Diskurssynthese (List et al., 2017; List, Stephens & Alexander, 2019) und die Anzahl von Begründungen für Argumente im eigenen Text (List et al., 2017). Als weitere Operationalisierung wählte man die Textlänge in Wörtern (Bräten, Brante & Strømsø, 2018; Latini et al., 2019; Leroy et al., 2023; Mason, Tarchi et al., 2022; Raković et al., 2023; Schoor et al., 2023) bzw. Sätzen (List & Du, 2021).

7) *Textstrukturmaße*

Textstrukturmaße widmeten sich der makrostrukturellen Themenentfaltung, und zwar hinsichtlich der Textsortenmerkmale und/oder thematischen Struktur. Dies erfolgte heterogen in den Studien. Einige Studien prüften die Adäquanz der in den Texten beschriebenen Inhalte, insbesondere fokussierten die Studien darauf, ob intertextuelle Bezüge hergestellt und die inhaltlichen Verhältnisse von Inhalten aus multiplen Dokumenten expliziert wurden (Bräten, Anmarkrud et al., 2014; Haverkamp et al., 2024; Kobayashi, 2009; Wiley et al., 2009). Andere Studien gingen auf die Vollständigkeit von Elementen bei Argumentationen (Barzilai et al., 2015), die kohärente Themenentfaltung (Spivey, 1992; Spivey &

King, 1989), eine thematisch strukturierte Themenentfaltung (Castells et al., 2023) sowie die textsortenadäquate Strukturierung des Texts (Binks et al., 2022; Schoor et al., 2023; Spivey, 1991; Spivey & King, 1989) ein. Schließlich wurde in einer Studie in Anbindung an die Lesetheorie die Einschätzung getroffen, ob die Diskurssynthesen ein Dokumentenmodell-Äquivalent darstellten (Mason et al., 2020). Bezogen auf Argumentationen als Zieltextsorte kamen drei weitere Maße in einer Studie von Binks et al. (2022) zum Einsatz, denn in ihr wurden zudem a) die Klarheit von Behauptungen, b) die Überzeugungskraft der Argumente und c) die Zweiseitigkeit von Argumenten überprüft.

8) *Textqualität*

Die aus der Schreibforschung stammende Variable Textqualität ist ein Globalmaß für die kommunikative Gelungenheit von Texten, oftmals wird sie mit Fokus auf Texttiefenmerkmale wie Kohärenz analysiert, meist mit menschlichen Raterinnen und Ratern. Die Studien lassen sich im Vorgehen zwei Gruppen zuordnen. Die erste Gruppe operationalisierte die holistisch eingeschätzte Qualität der Texte hinsichtlich des Erreichens der textsortenspezifischen Kommunikationsabsicht mit einem Globalurteil (Anmarkrud et al., 2013, 2014; Barzilay et al., 2023; Binks et al., 2022; Chau et al., 2022; Du & List, 2020; Holdinga et al., 2021; Lee & List, 2021, 2023b; Leijten et al., 2019; List, in Druck, Studie 2; List & Alexander, 2020; List & Du, 2021; List, Stephens & Alexander, 2019; Spivey, 1992; Spivey & King, 1989). Die zweite Gruppe nutzte mehrere Merkmale, welche dann zu Gesamtindikatoren gebündelt wurden, d. h. die mehrdimensional eingeschätzte Qualität der Texte fußte auf Dimensionen wie Aufbau, Themenentwicklung und Sprachgebrauch (Chang, 2023; Cheong et al., 2019; Golparvar & Khafi, 2021; H. Kim, 2020; Liao & Zhao, 2024; Liao et al., 2021; Neumann et al., 2020; Raković et al., 2023; Risemberg, 1996; Tarchi, Villalón et al., in Druck; van der Graaf et al., 2022; Yang, 2014; Yang & Plakans, 2012; Zhao et al., 2023).

5.2.2 **Perspektive 1: Zusammenhänge von Leseprozessen und Textprodukten bei Diskurssynthesen**

5.2.2.1 **Unspezifische Leseprozessmaße**

Die unspezifischen Lesemaße wurden in den Studien über drei Gruppen von Daten erfasst und ausgewertet: Lesezeiten, Konsultations- bzw. Navigationsverhalten und non-lineares Vorgehen. Unter diese übergeordnete Kategorie lassen sich diverse Operationalisierungen versammeln.

- Ein sehr grobes Maß ist die *Lesezeit*, also die gemessene Zeitspanne, die nur mit dem Lesen verbracht wurde; damit ist keine Gesamtbearbeitungsdauer gemeint, welche umfassender ist. Hinter der Kategorie Lesezeit verbergen sich die Zeiten der Konsultation von ganzen Dokumenten, spezifischen Passagen und von Dokumenten in verschiedenen Intervallen der Lesezeit. Darunter fallen in den Studien konkret diese Maße: Zeit, die mit dem Lesen von eigenständig geöffneten digitalen Texten verbracht wurde (Barzilai et al., 2023; Bråten, Brante & Strømsø, 2018; Cerdán & Vidal-Abarca, 2008; Lee & List, 2021, 2023b; Liao & Zhao, 2024; List, Oaxaca et al., 2021; Schoor et al., 2023), mitunter noch differenziert nach Passagen, die relevante bzw. irrelevante Informationen enthielten (Cerdán & Vidal-Abarca, 2008); Lesezeit (Bråten, Anmarkrud et al., 2014; Du & List, 2020; Holdinga et al., 2021; Latini et al., 2019; List, Stephens & Alexander, 2019; Mason, Tarchi et al., 2022), darunter auch: betrachtet nach Auftreten in dem nach in Drittel unterteilten Gesamtprozess (Vandermeulen, van den Broek et al., 2020); als ein Gesamtindikator die Lesezeit unter Berücksichtigung der Wechsel zwischen den einzelnen zu lesenden Dokumenten (Chau et al., 2022; Leijten et al., 2019) sowie die Lesezeit vertrauenswürdiger Websites im Vergleich zum Lesen nicht-vertrauenswürdiger Websites (Wiley et al., 2009). Zudem fallen kombinierte Maße mit Lesen und Hervorheben und Notizenanfertigen in diese Kategorie (Raković, Fan et al., 2022).
- Die *Konsultation relevanter Information(seinheit)en* bezeichnet die Häufigkeit, in welcher die Testpersonen die für die Aufgabenerreichung nötig erscheinenden Informationseinheiten bzw. ganze Dokumente aufgesucht, betrachtet oder gelesen haben.
In den Studien wurden eher unspezifische Indikatoren dafür gewählt, nämlich die Anzahl insgesamt geöffneter Dokumente (Du & List, 2020; Lee & List, 2023b; List & Du, 2021; List, Oaxaca et al., 2021; List, Stephens & Alexander, 2019). Spezifischer fallen im Vergleich jene Maße aus, welche Rückschlüsse auf die für die Aufgabenbearbeitung nötigen Informationen und Informationsträger erlauben. Dazu zählen die Zeit für das Lesen der für die Aufgabenbearbeitung relevanteren Texte (Lee & List, 2023b), der Anteil der relevanten Texte von allen besuchten Texten (Lee & List, 2023b), die Anzahl an linearen Konsultationen einzelner Informationen sowie erneuten nicht-linearen Konsultationen relevanter Informationen (Cerdán & Vidal-Abarca, 2008) bzw. vertrauenswürdiger Websites (Wiley et al., 2009) oder die Anzahl erneuter Öffnungen relevanter digitaler Dokumente (Leroy et al., 2021, 2023) bzw. erneuter Konsultationen von Dokumenten allgemein (List & Alexander, 2018b).
- Ein *non-lineares, rekursives Vorgehen*, was aus der Kombination von mehreren Daten als Muster extrahiert wurde, gibt alternierende Aktivitäten samt Wiederholungen an. Darunter fallen Wechsel von Lese- und partiell Schreibaktivitäten und ein insgesamt iteratives Handeln, bei dem Prozesse zusammenspielen

und teils mehrere ähnliche Aktivitäten in adaptiven Sequenzen hintereinander ablaufen – hierin liegt eine Differenz zu einem eher linearen Vorgehen. Operationalisiert wurde dies in den Studien über die Anzahl Wechsel beim Lesen zwischen den einzelnen Dokumenten (Chau et al., 2022; Leijten et al., 2019; Liao & Zhao, 2024; Peterson & Alexander, 2020; Schoor et al., 2023; Vandermeulen, van den Broek et al., 2020), Wechseln zwischen zu lesenden Texten und eigenem Text (Liao & Zhao, 2024; Vandermeulen, van den Broek et al., 2020), der Anzahl der Wechsel zwischen Lesen und Notieren sowie Lesen und Schreiben (Barzilai et al., 2023) und ein rekonstruiertes Navigationsmuster mit Tendenz zum wiederholten Konsultieren von vertrauenswürdigen Websites (Wiley et al., 2009).

Die Zusammenhänge von diesen drei Gruppen von Prozessindikatoren und den Produktmaßen gestalten sich nicht immer einheitlich oder konsistent, wie die nun folgende Darstellung es als Muster erbringt.

Unspezifische Leseprozesse 1: Lesezeit

Längeres Lesen, so das erste Muster aus der Empirie, ist von Vorteil. Wer längere Zeit mit dem Lesen der Dokumente verbracht hatte,

- vollzog in den Diskurssynthesen mehr Wechsel der Informationen aus verschiedenen Dokumenten (Switches; Bråten, Brante & Strømsø, 2018),
- übernahm in den Diskurssynthesen mehr Inhalte aus den gelesenen Dokumenten (Bråten, Brante & Strømsø, 2018) – insbesondere korrekte Inhalte, wenn die Personen vergleichsweise viel Zeit auf vertrauenswürdigen Websites verbracht hatten (Wiley et al., 2009),
- nutzte mehr textbasierte Begründungen (List, Stephens & Alexander, 2019),
- gab mehr inferenzbasierte Kombinationen von Informationen in den Texten an (Mason, Tarchi et al., 2022),
- schrieb längere Diskurssynthesen (Schoor et al., 2023),
- band mehr Argumente in argumentative Diskurssynthesen ein (List, Stephens & Alexander, 2019),
- verwendete mehr Diskurskonnektoren (Du & List, 2020; Schoor et al., 2023),
- dessen Diskurssynthese umfasste mehr Referenzierungen (Du & List, 2020; List, Oaxaca et al., 2021; List, Stephens & Alexander, 2019; Schoor et al., 2023),
- bezog mehr Perspektiven zur kontroversen Thematik erfolgreich aufeinander (Bråten, Anmarkrud et al., 2014),
- gab komplexere und dem Gegenstand angemessenere Kausalzusammenhänge schriftlich wieder (Wiley et al., 2009),
- erzielte eine höhere Textqualität (Du & List, 2020; Holdinga et al., 2021; Lee & List, 2021, 2023b; Liao & Zhao, 2024; List, Stephens & Alexander, 2019).

Die obige Listung von positiven Zusammenhängen darf nicht darüber hinwegtäuschen, dass zur Lesezeit überdies weniger eindeutige Ergebnisse existieren. So gibt es Hinweise aus Studien, dass die reine Lesezeit (Barzilai et al., 2023; Cerdán & Vidal-Abarca, 2008; Latini et al., 2019; List, Oaxaca et al., 2021) bzw. Kombinationen von Lesezeit und anderen Prozessindikatoren (Chau et al., 2022; Leijten et al., 2019; Raković, Fan et al., 2022) *keinerlei linearen Zusammenhänge* mit den Merkmalen der finalen Diskurssynthesen aufwies. Sogar *negative Zusammenhänge* mit der Textqualität bzw. Textstrukturmaßen wurden ermittelt, sofern die Lesezeit bei der Erstfixation länger dauerte (Mason et al., 2020) bzw. einen hohen Anteil an der frühen Aufgabenbearbeitung (Holdinga et al., 2021) oder am Ende des Herstellens von Diskurssynthesen (Vandermeulen, van den Broek et al., 2020) ausmachte. Ebenfalls haben sich in einem Fall lange Lesezeiten ausgerechnet von relevanten Informationen als ungünstig für den Abdeckungsgrad eben jener Inhalte in den Diskurssynthesen erwiesen (Cerdán & Vidal-Abarca, 2008). Mitunter gestaltete sich der Zusammenhang von Lesezeit und Textqualität auch nicht als linear, sondern als *kurvilinear*, d. h. es war günstig, weder zu kurz noch zu lang in einzelnen Phasen des Diskurssyntheseschreibens zu lesen (Vandermeulen, van den Broek et al., 2020). Das Gesamt der Befunde verweist auf eine tendenzielle Überlegenheit bei einem längeren Lesen, doch ist dies allein noch kein Indiz dafür, dass später entstehende Textprodukte davon allein schon profitieren würden.

Unspezifische Leseprozesse 2: Konsultation relevanter Informationen und Informationseinheiten

In den Studien zum Herstellen der Diskurssynthesen sind häufig die Informationen, welche für die Vollständigkeit der mentalen Repräsentation eines meist strittigen Sachverhalts nötig sind, über mehrere Dokumente verteilt. Es liegt diesen Experimenten also die Annahme zugrunde, dass Personen mit dem Auftrag, eine Diskurssynthese herzustellen, gut beraten sind, sämtliche Dokumente zu rezipieren. List und Du (2021) sowie Lee und List (2023b) fanden ein positives Verhältnis von Anzahl aufgerufener Dokumente und der Textqualität. List, Stephens und Alexander (2019) ermittelten keinen Zusammenhang mit der Textqualität, aber dafür mit drei anderen Maßen, nämlich der Zahl textbasierter Begründungen, der Zahl von Argumenten und der Zahl von Referenzierungen. Keine direkten Effekte berichten hingegen Du und List (2020) sowie List, Oaxaca et al. (2021). Es wirkt demnach angesichts der heterogenen Befundlage so, als wäre die Breite abgedeckter Dokumente beim Lesen noch nicht ausreichend dafür, um bessere Diskurssynthesen herzustellen.

Andere Studien setzten andere Akzente in der Auswertung, diese sind allerdings insgesamt so heterogen, dass sie sich kaum abstrahierend bündeln lassen. Es gibt sowohl positive als auch Nulleffekte in der Empirie:

- *Positive Effekte:* Das häufige Lesen einzelner Textteile, d.h. für die Aufgabenstellung relevanter Absätze, ging damit einher, dass ein höherer Abdeckungsgrad an Inhalten in den Diskurssynthesen feststellbar war (Cerdán & Vidal-Abarca, 2008). Für intertextuelle Inferenzen war es zudem günstig, wenn Personen relevante Dokumente erneut geöffnet und Informationen hervorgehoben hatten (Leroy et al., 2021). Zudem enthielten Diskurssynthesen dann mehr Referenzierungen, wenn Personen Dokumente häufiger erneut geöffnet hatten (List & Alexander, 2018b).
- *Nulleffekte:* Wiley et al. (2009) prüften, ob die Anzahl der erneuten Konsultationen besonders vertrauenswürdiger Websites mit Produktmaßen korrelierte, und fanden keine statistischen Zusammenhänge. Ebenfalls keine Zusammenhänge mit Produktmaßen ermittelten Leroy et al. (2023) für die Anzahl erneuter Öffnungen von Dokumenten und Cerdán und Vidal-Abarca (2008) für häufigeres nicht-lineares Lesen von mehreren Textteilen mit relevanten Informationen. Ähnliche Nulleffekte berichten zudem Castells et al. (2023) für die Häufigkeit von Relektüren sowohl vor dem Verfassen als auch während des Verfassens von Diskurssynthesen. Mason et al. (2020) ermittelten keinen Effekt der längeren Lesezeit von erneut mit den Augen fixierten Inhalten und der Vollständigkeit von Dokumentenmodellen in der Diskurssynthese. Und schließlich berichteten Lee und List (2023b) ebenfalls davon, dass weder höhere Anteile der konsultierten relevanten Dokumente an der Gesamtzahl der besuchten Dokumente noch die Lesezeit dieser relevanten Dokumente mit einer höheren Textqualität einhergingen.

Wie schon bei der Lesezeit ist damit der positive Effekt des Konsultierens relevanter Informationen bzw. Einheiten nicht per se ein Indikator für bessere Textprodukte.

Unspezifische Leseprozesse 3: non-lineares rekursives Vorgehen

Die dritte Gruppe von unspezifischen Prozessmaßen ist bereits stark aggregiert. Es wurden hierfür einzelne Verhaltensindikatoren in Bezug zueinander gesetzt, um Verhaltensmuster zu extrahieren, in diesem Fall eines der Rekursion. Damit ist gemeint, dass Personen nicht linear vorgegangen sind, sondern Sequenzen wiederholten. Die Forschung hierzu liefert vereinzelte Befunde:

- Wer häufig zwischen verschiedenen Dokumenten wechselte, dessen Textprodukt enthielt mehr Informationseinheiten, die sich gemäß den Operationalisierungen zuvorderst aus übernommenen Inhalten speisten (Peterson & Alexander, 2020).
- Häufige Wechsel zwischen Lesen und grafischen Notieren bzw. dem Schreiben korrespondierten mit mehr Referenzierungen in der Diskurssynthese, zudem

verwendeten die Testpersonen auch mehr Dokumente in der Diskurssynthese, wenn Lesen und Schreiben sich häufiger in der Phase des Schreibens abwechselten (Barzilai et al., 2023). Jedoch sind nur drei von insgesamt zehn Korrelationen in der Studie signifikant gewesen.

- Wer dazu neigte, mehrfach vertrauenswürdige Websites aufzusuchen, schrieb Diskurssynthesen mit mehr korrekten Inhalten zum einen und mit weniger inkorrekten Inhalten zum anderen (Wiley et al., 2009).
- Je mehr Wechsel zwischen den Texten beim Lesen erfolgt waren, desto mehr adversative Konjunktionen und desto mehr Ausdrücke mit Hinweis auf die Berücksichtigung zweier Dokumente enthielten die Diskurssynthesen. Außerdem waren die Texte expliziter in der Benennung der Diskrepanz des Themas, umfangreicher in puncto Wortzahl und wiesen zudem eher die Struktur eines kontinuierlichen Texts im Vergleich zu einer reinen Listung auf (Schoor et al., 2023).
- Häufige Wechsel bei der Lektüre gingen mit besserer Textqualität einher (Chau et al., 2022; Leijten et al., 2019; Liao & Zhao, 2024), wobei dies je nach Aufgabe – in diesem Fall: informierenden Diskurssynthesen – günstig zu Beginn der Bearbeitung war und ungünstig gegen Ende (Vandermeulen, van den Broek et al., 2020).
- Ebenfalls war ein Wechsel zwischen Leseaktivitäten und dem eigenen entstehenden Text bei argumentativen Diskurssynthesen für die Textqualität günstig (Liao & Zhao, 2024), mitunter aber nur dann wenn dieses Verhalten zu Beginn gezeigt worden war (Vandermeulen, van den Broek et al., 2020).

Im Licht dieser Ergebnisse wirkt das non-lineare Vorgehen partiell günstiger als ein lineares, doch es scheinen Spezifika zu bestehen, für welche Variable und für welchen Zeitpunkt Rekursionen positiv sind.

5.2.2.2 Spezifische Leseprozessmaße

Welche fünf Gruppen der Leseprozessmaße wurden in Studien erfasst?

Spezifische Lesemaße kann man zu fünf Gruppen bündeln, wobei sich die ersten drei in die Logik bisheriger Taxonomien gut einfügen lassen (Cho, Afflerbach & Han, 2018; List, 2020; Philipp, 2020a): Integrations-, Sourcing- sowie Evaluations- und Elaborationsprozesse. Hinzu kommen zwei weitere Gruppen, die sich zum einen mehr auf Einzeltexte beziehen und zum anderen Spezifika in den Funktionen betreffen, die in anderen Klassifikationen noch nicht abgedeckt worden sind. Diese fünf Gruppen werden zunächst kurz skizziert, ehe danach jeweils die Zusammenhänge von Produktmaßen und den Prozessindikatoren aus den einzelnen Gruppen vertieft behandelt werden.

1) Integrationsprozesse

Großes Interesse haben in der Forschung die Integrationsprozesse evoziert, womit inferenzbasierte kognitive Kombinationen von Informationen aus verschiedenen Dokumenten gemeint sind (s. Teilkap. 4.2.1); mehrheitlich handelte es sich um textuelle Informationen.

Hier lassen sich anhand verschiedener Datenquellen die Zugänge unterteilen. Es dominieren *Selbstauskünfte in Fragebögen* zum intertextuellen Integrieren beim Lesen (Bråten, Anmarkrud et al., 2014; Cheong et al., 2019; Golparvar & Khafi, 2021; Haverkamp et al., 2024; Liao et al., 2021; List, Du, Wang & Lee, 2019; Yang, 2014; Yang & Plakans, 2012) vs. Fokussierung auf Informationsakkumulation beim Lesen (List, Du, Wang & Lee, 2019). Prozessnah erfolgten Analysen auf der Basis von ausgezählten *Lautdenk-Protokoll-Passagen*, in denen zwischen Dokumenten Verknüpfungen hergestellt werden (Anmarkrud et al., 2014; Barzilai et al., 2015; List & Alexander, 2020). Ein letzter Indikator für intertextuelles Integrieren waren Verhaltensindikatoren bei digitalen Dokumenten, die ein Gruppieren von thematisch zusammengehörigen Dokumenten betreffen, wofür intertextuelles Verstehen nötig ist (Leroy et al., 2021, 2023).

2) Sourcingprozesse

Die Sourcingprozesse in Studien wurden geprüft, um die Verwendung von Metadaten für die Auswahl, Einschätzung und Nutzung von Dokumenten als soziale Artefakte zu rekonstruieren.

Dies erfolgte multimethodisch. Mittels *Tracking von Verhalten* wurde beispielsweise das gezielte Anwählen von allgemeinen Metadaten (Du & List, 2020; List & Alexander, 2018b; List et al., 2017; List & Du, 2021), eingebetteten Metadaten in Dokumenten (List & Lin, 2023) bzw. von spezifischen Metadaten zum Autor/zur Autorin erfasst (List & Alexander, 2018b). In Studien mit *Eye-Tracking* ließen sich Anzahl und Dauer von Erst- und Zweitfixationen von Metadaten sowie Gesamtfixationsdauer bei Metadaten bestimmen (Salmerón, Gil & Bråten, 2018b). Studien mit *Lautdenk-Protokollen* untersuchten die Häufigkeit eines aufmerksamen Bemerkens von Metadaten (Barzilai et al., 2015; Strømsø & Bråten, 2014; Strømsø et al., 2013) sowie die Anzahl der Verbindungen von inhaltlichen Aussagen und Metadaten (Barzilai et al., 2015).

3) Evaluations- und Elaborationsprozesse

Die dritte hier interessierende Gruppe setzt sich aus Evaluations- und Elaborationsprozessen zusammen, welche die Verarbeitung der gelesenen Texte mit Informationen jenseits der konkret fokussierten Textstelle, vornehmlich vorwissensbasiert, eint. Die Operationalisierungen sind heterogen. Sie reichen von

der Anzahl vorwissensbasierter Einschätzungen und/oder Anreicherungen der Informationen, die man gelesen hat (Ackerman, 1991), über die Anzahl von Beurteilungen von Textinhalten (Anmarkrud et al., 2014) bzw. die positiven oder negativen Einschätzungen der Relevanz einzelner Textsegmente für das Aufgabenziel (Anmarkrud et al., 2013) hin zu kritisch-analytischen Prozessen auf Einzeldokumentebene bzw. auf Ebene mehrerer Dokumente in Lautdenk-Protokollen (List & Alexander, 2020). Weiter wurden elaborative Fragen, Elaborationen und Evaluationen (Du & List, 2021) untersucht und lassen sich aufgrund der vorwissensbasierten Verknüpfungen dieser Gruppe von strategischen Prozessen zuordnen. Ein anderer Ansatz besteht darin, retrospektiv Probandinnen und Probanden zu fragen, welche Strategien des kritischen Denkens sie eingesetzt haben (List, in Druck, Studie 2).

4) *Lesestrategien auf Einzeltextebene*

Überprüft wurde in den Studien zudem, auf welche *Lesestrategien auf Einzeltextebene* Probandinnen und Probanden zurückgreifen, d. h. welche zielbezogenen Vorgehensweisen sie anwenden, die sich auf intratextuelle Informationseinheiten beziehen. Die bislang erkennbaren Zugangsweisen in Studien betreffen erstens die Paraphrase gelesener Inhalte auf Einzeltextebene (Du & List, 2021; Golparvar & Khafi, 2021; Yang, 2014), zweitens interessierten die Unterstreichungen/Her hervorhebungen beim Lesen (Griffin et al., in Druck; Kobayashi, 2009; Leroy et al., 2021; List & Alexander, 2020) sowie – drittens – geäußerte Verstehensprozesse auf Einzeldokumentebene in Lautdenk-Protokollen (List & Alexander, 2020).

5) *Lesestrategien nach Funktionsunterscheidung*

Mit den Lesestrategien nach Funktionsunterscheidung ist eine weitere Klassifikation angesprochen, die eher selten in Studien vorgekommen ist, aber klassische Trennungen von Strategiefamilien aufgreift und vor allem mit Lautdenk-Protokollen realisiert wurde. Dies betrifft zum einen den Fokus auf metakognitive Monitoringstrategien beim Lesen (Anmarkrud et al., 2014; Du & List, 2020; Linderholm et al., 2014, Studie 1) bzw. Selbstregulationsstrategien (Du & List, 2021; List & Alexander, 2020) und zum anderen die Unterscheidung von oberflächen- und tiefenorientierten Strategien (Du & List, 2020). Die erstgenannte Gruppe von Unterscheidungen geht auf die Metakognition und Selbstregulation ein (s. dazu Teilkap. [6.1.2](#)), die andere Differenzierung betrifft die Tiefe der kognitiven Verstehensprozesse (Dinsmore & Hattan, 2020).

Die Integrationsprozesse wurden mit Fragebögen, Lautdenk-Protokollen und – in deutlich geringem Maß – mit Verhaltensindikatoren erfasst. In dieser Reihenfolge der Operationalisierungen werden die Ergebnisse im Folgenden entfaltet.

Mittels eines ersten *Fragebogens*, der das Vorgehen beim Lesen multipler Dokumente retrospektiv erfasst (Bråten & Strømsø, 2011), ließen sich zwei Arten vom Umgang mit Informationen extrahieren: einerseits ein Ansammeln von Informationen, welches bestenfalls als Auswahlprozess im Sinne des Teilkapitels [5.1.2.1](#) wertbar ist, andererseits ein Bemühen um das Herstellen von intertextueller Kohärenz. Ein zweiter Fragebogen aus dem L2-Kontext bildete ein verwandtes Konstrukt ab, das als „Diskurssynthesestrategien“ bezeichnet wurde (Cheong et al., 2019; Yang & Plakans, 2012) und primär Auswahl- und Organisationsprozesse von Informationen beinhaltet.

Die Skala zum Sammeln von Informationen aus dem ersten Fragebogen, die bisher in nur einer Studie mit Diskurssynthesen verwendet wurde (List, Du, Wang & Lee, 2019, Studie 1), war statistisch nahezu vollständig unkorreliert mit Produktmaßen. Demgegenüber war die Skala zum Herstellen intertextueller Kohärenz aus dem ersten Fragebogen und die Skala „Diskurssynthesestrategien“ aus dem zweiten Fragebogen bedeutsamer. Wer nämlich angab, mehr um intertextuelle Kohärenz beim Lesen bemüht gewesen zu sein,

- verwendete mehr Diskurskonnektoren in der Diskurssynthese (List, Du, Wang & Lee, 2019, Studie 1);
- hatte mehr Aussagen auf der Basis intertextueller Inferenzen im Text eingebunden (List, Du, Wang & Lee, 2019, Studie 1);
- hatte mehr Perspektiven aus den Texten rund um ein strittiges Thema aufeinander bezogen (Bråten, Anmarkrud et al., 2014; Haverkamp et al., 2024);
- nutzte mehr Referenzierungen (List, Du, Wang & Lee, 2019, Studie 1) und
- erzielte eine höhere Textqualität (Cheong et al., 2019; Golparvar & Khafi, 2021; Yang, 2014; Yang & Plakans, 2012).

In einem dritten Instrument (Liao et al., 2021), in welchem intertextuelles Integrieren in zwei von vier Faktoren abgedeckt und diese Faktoren zu einem latenten Konstrukt zusammengefasst wurden, ergaben sich ebenfalls positive Zusammenhänge mit der Textqualität.

Insbesondere die Textqualität scheint demnach davon zu profitieren, wenn laut Selbstauskünften Integrationsprozesse beim Herstellen einer Diskurssynthese stattgefunden haben. Studien mit *Lautdenk-Protokollen* gelangen zu ähnlichen Befunden. Wenn Personen mehr intertextuelle Verknüpfungen in Lautdenk-Protokollen äußerten, dann zitierten sie mehr Dokumente explizit (Anmarkrud et al., 2014; List & Alexander, 2020), stellten mehr Verbindungen von Quellen

und Aussagen her (Anmarkrud et al., 2014), hatten besser strukturierte Argumentationen (Barzilai et al., 2015) und erzielten zudem in ihren argumentativen Diskurssynthesen eine höhere Qualität, was eine ausgewogenere Darstellung der Kontroverse impliziert (Anmarkrud et al., 2014).

In zwei Experimenten wurde geprüft, ob das Gruppieren von thematisch zusammengehörigen Dokumenten – ein *behavioraler Indikator für intertextuelles Integrieren* auf der Basis ganzer Dokumente und ein Indiz für Intertextprädikate – mit Merkmalen der Diskurssynthesen zusammenhing. Das war tatsächlich der Fall in Form längerer Diskurssynthesen (Leroy et al., 2023) und mehr intertextueller Aussagen in den Synthesen (Leroy et al., 2021).

Die in den Studien auf verschiedene Arten ermittelten Integrationsprozesse illustrieren, dass es konvergente und konsistente Ergebnisse gibt: Wer intertextuelle Integrationsprozesse beim Lesen absolviert, verfasst qualitativ höherwertige Diskurssynthesen und zitiert mehr. Auffällig ist außerdem, dass sehr viele unterschiedliche Produktmaße mit dem Integrieren positiv korrelierten, was ebenfalls die Bedeutsamkeit des Integrierens neuerlich unterstreicht.

Spezifische Leseprozesse 2: Sourcingprozesse

Wenngleich das Sourcing im Kontext der Multiple-Documents-Literacy-Forschung zunehmend und starke Forschungsaktivitäten hervorgerufen hat (Braasch & Scharrer, 2020; Bråten et al., 2017), sind Studien mit einem Fokus auf Diskurssynthesen spärlich vorgelegt worden und wenn, dann vorrangig im Kontext von Studien mit konfligierenden Dokumentensets. In ihnen wurde als ein erstes und freilich inhaltlich wenig präzises Maß untersucht, ob in spezifischen Apps die Metadaten zu Dokumenten angewählt wurden. Dies erwies sich als vorteilhaft für die Zahl von Zitationen in der Diskurssynthese (Du & List, 2020; List & Alexander, 2018b) bzw. die Anzahl zitierter Dokumente (List et al., 2017), die Anzahl von Argumenten in Diskurssynthesen sowie die Anzahl von Belegen für diese Argumente (List et al., 2017). Keine Zusammenhänge mit Diskurskonnektorenzahl und Qualität der Diskurssynthesen fanden sich bei Du und List (2020). Bei eingebetteten Quellen, bei denen überprüft wurde, wie viele Testpersonen Metadaten mittels Link-Aktivierung konsultieren, ergaben sich bei List und Lin (2023) positive Zusammenhänge mit der Anzahl übernommener Argumente, aber nicht bei der Anzahl von intertextuellen Inferenzen in der Diskurssynthese. Eye-Tracking-Daten wiederum illustrieren zusätzlich, dass vor allem längere und wiederholte Konsultationen von Metadaten dazu geführt hatten, dass später im Text mehr Zitationen auftauchten (Salmerón, Gil & Bråten, 2018b). Die Forschungsergebnisse sind insgesamt noch sehr lückenhaft und wirken mit Blick auf die Erträge des Sourcings uneinheitlich. Gleichwohl deuten sich positive Effekte ganz unterschiedlicher Sourcingaktivitäten für Diskurssynthesen an.

Studien mit Lautdenk-Protokollen zeigen, dass ein allgemeines Bemerkens von Metadaten mit mehr Zitationen in den Diskurssynthesen korrespondiert und dass insbesondere nicht nur die Zitationen allgemein von mehr Sourcingaktivitäten profitieren, sondern dass sich die Wahrscheinlichkeit erhöht, dass die aktuell beim Lesen in ihren Metadaten aufmerksam fokussierte Quelle tatsächlich später explizit im Text zitiert wird (Strømsø & Bråten, 2014). Mitunter betraf dies jedoch nur einzelne Dokumente (Strømsø et al., 2013). Sourcingprozesse, seien es die Repräsentationen von Metadaten, seien es die Verknüpfungen von Metadaten und Inhalt, gingen zudem mit mehr Argumenten und Begründungen und besser strukturierten, vollständigeren Argumentationen einher (Barzilai et al., 2015). Sourcingaktivitäten haben im Licht dieser Prozesse vor allem eine Nähe zu Zitationen im späteren eigenen Text, der Diskurssynthese. Andere Textmaße sind seltener untersucht, doch auch bei ihnen offenbarten sich hin und wieder Zusammenhänge, welche durchaus Texttiefenmerkmale wie die Struktur betreffen. Jedoch bilden Sourcingprozesse im Kontext der prozess- und produktbezogenen Forschung zu Diskurssynthesen immer noch ein Desiderat.

Spezifische Leseprozesse 3: Evaluations- und Elaborationsprozesse

Prozesse, die deutliche Züge des angewandten Vorwissens in sich bergen, sind in den Studien ebenfalls untersucht worden, dabei überwiegen Lautdenkdaten. Die Studien hierzu sind bislang selten durchgeführt worden. Sie kommen bei den Elaborationen zu weniger eindeutigen Ergebnissen als bei den evaluativen Kommentierungen.

Bei den *Elaborationen*, die kaum untersucht sind, sind die Befunde nicht eindeutig. Ackerman (1991) ermittelte beispielsweise kaum Zusammenhänge von elaborativen Aussagen und Produktmaßen. Auffällig war jedoch, dass elaborative Prozesse damit einhergingen, dass die Testpersonen umso mehr Inhalte jenseits der Inhalte aus den gelesenen Dokumenten in ihre Diskurssynthesen aufgenommen hatten, wenn sie aktiv mehr Elaborationen vorgenommen hatten. Das heißt: Je mehr sie vorwissensbasierte Kommentare geäußert hatten, desto mehr eigene Ideen fügten sie ihrer Diskurssynthese hinzu und desto weniger Inhalte übernahmen sie explizit. Elaborationen waren jedoch in der Studie von Du und List (2021) durchaus in Regressionsanalysen Prädiktor für ein Produktmaß, nämlich die angemessene Darstellung kontroverser Gründe für kontrovers berichtete Sachverhalte. Für elaborative Fragen galt dieser Zusammenhang in derselben Studie nicht.

Evaluative Kommentare wirken hingegen deutlicher positiv mit Produktmaßen verknüpft. In einer norwegischen Studie (Anmarkrud et al., 2014) ging die Anzahl evaluativer Kommentierungen beim Lesen positiv einher mit der Textqualität. Eine vertiefende Analyse derselben Daten aus der Veröffentlichung von Anmarkrud et al. (2014) verdeutlicht die Bedeutung der Relevanz von

Informationen bei den Integrationsprozessen: So gab es nur dann einen positiven Zusammenhang von Textqualität und Relevanzurteilen, wenn zum einen die Leserinnen und Leser relevante Textpassagen beim Lesen als relevant erkannten und zum anderen irrelevante Textteile als solche beurteilten (Anmarkrud et al., 2013). Außerdem waren kritisch-analytische Prozesse in der Studie von List und Alexander (2020) mit der Zahl von Zitationen, nicht aber der Textqualität korreliert. Demgegenüber fand List (in Druck, Studie 2) positive Zusammenhänge von der Zahl retrospektiv geschilderter kritisch-evaluativer Strategien und der Qualität von Argumentationen. Evaluationen waren außerdem nicht nur korrelativ, sondern auch regressionsanalytisch nachweislich mit der Ausgewogenheit verschiedener Belege für kontroverse Themen verbunden (Du & List, 2021). In letztgenannter Studie leisteten damit sowohl elaborative als auch evaluative Aussagen eigenständige Varianzaufklärungen.

Spezifische Leseprozesse 4: Lesestrategien auf Einzeltextebene

Intertextuelles Verstehen für das Verfassen von Diskurssynthesen setzt auf dem Verstehen einzelner Dokumente auf. Darum sind Studien von Interesse, die dieses intratextuelle Verstehen genauer fokussieren. Die Ergebnisse unterstreichen, dass strategische intratextuelle Verstehensprozesse nicht immer zu besseren Diskurssynthesen führen:

- *Günstige Effekte des Paraphrasierens:* Wer (wichtige) gelesene Inhalte in eigenen Worten wiedergibt und damit im Arbeitsgedächtnis hält und bereits erste Transformationsleistungen erbracht hat, der paraphrasiert. Solche Paraphrasen sind günstig für die Textqualität, wie es Fragebogenstudien (Golparvar & Khafi, 2021; Yang, 2014 – kombiniert mit dem Zitieren während des Synthetisierens) illustrierten. Für eine ausgewogene Darstellung variierender Gründe in kontroversen Themen war das Wiederholen von Aussagen laut Regressionsanalysen ebenfalls günstig (Du & List, 2021). Diese drei Studien kommen damit einhellig zum Schluss, dass sinngemäße Wiederholungen von Inhalten einen positiven Beitrag für Texttiefenmerkmale zu leisten imstande sind.
- *Mehrheitlich keine Zusammenhänge mit dem Unterstreichen/Hervorheben:* Das Unterstreichen bzw. Hervorheben wurde in fünf Studien (Griffin et al., in Druck; Kobayashi, 2009; Leroy et al., 2021; List & Alexander, 2020; List & Lin, 2023) untersucht und dort auf das Vorhandensein hin überprüft bzw. die Anzahl der Hervorhebungen gezählt. In vier Studien gab es keine direkten Zusammenhänge mit Produktmaßen (Griffin et al., in Druck; Kobayashi, 2009; Leroy et al., 2021; List & Alexander, 2020). Die fünfte Studie von List und Lin (2023) ermittelte hingegen differenzielle Effekte bei den sechs Indikatoren der Hervorhebungen bezogen auf zwei Diskurssynthesemaße. Für den Abdeckungsgrad von Inhalten erwiesen sich drei Hervorhebungsmaße

als positive Prädiktoren: Je mehr Schlüsselbegriffe und je mehr Metadaten hervorgehoben wurden und je mehr Hervorhebungen als Summenwert feststellbar waren, desto mehr Argumente aus den Texten wurden aufgegriffen. Für die lokalen Kombinationen als Anzahl intertextueller Inferenzen waren zwei Hervorhebungsmaße prädiktiv. Zum einen betraf dies Metadaten und zum anderen unterstützende Details. Dies verweist auf komplexere Zusammenhänge, die allerdings noch zu wenig erforscht wirken. In der Studie von List und Lin (2023) ist jedoch ein wichtiges Ergebnis gewesen, dass die Anzahl von metadatenbezogenen Hervorhebungen mit anderen Hervorhebungen statistisch unverbunden war, was auf eine Spezifik beim Hervorheben im Sinne einer Sensibilität für Informationsarten verweist.

- *Keine Zusammenhänge mit groben Kategorien des intratextuellen Verstehens:* Der Vollständigkeit halber ist noch die Studie von List und Alexander (2020) abschließend anzuführen. Diese prüfte, ob Lautdenk-Protokollaussagen, die als intratextuelles Verstehen bezeichnet werden können, mit den Diskurssynthesemaßen korrelierten. Dem war nicht so.

Damit deuten die Ergebnisse an, dass es bei intratextuellen Leseverstehensprozessen Differenzen gibt, welcher Art sie sind und ob sich Zusammenhänge mit Produktmaßen haben feststellen lassen. Lediglich das Paraphrasieren erwies sich mehrheitlich mit Leistungsdaten verquickt.

Spezifische Leseprozesse 5: Strategien nach Funktionsunterscheidung

In diese letzte Kategorie von Strategien gehören Maße, die von den anderen Kategorien noch nicht abgedeckt wurden. Darunter fallen *metakognitive Prozesse bzw. selbstregulatorische Strategien* aus sechs Studien mit Lautdenk-Protokollen. Hier gab es keine Zusammenhänge von metakognitiven Strategien und Textqualität und Zitationen (List & Alexander, 2020) sowie der Ausgewogenheit der Belege bei kontrovers dargestellten Sachverhalten (Du & List, 2021). Hingegen fanden Linderholm et al. (2014, Studie 1) positive Zusammenhänge von Verstehensüberwachung zum einen und negative Zusammenhänge von Verstehensproblemen zum anderen mit der Anzahl lokaler Kombinationen. Anmarkrud et al. (2014) ermittelten durchaus und substanzielle Zusammenhänge mit der Textqualität. Du und List (2020) konnten keine Zusammenhänge zwischen metakognitiven Strategien und Textqualität eruieren, aber zwischen diesen Strategien und Anzahl der Referenzierungen sowie Diskurskonnektoren. Auch van der Graaf et al. (2022) konnten keine direkten Zusammenhänge zwischen metakognitiven Strategien und Textqualität feststellen. Die Befundlage ist damit gemischt und uneindeutig.

Bei der *Unterscheidung in oberflächen- und tiefenorientierte Strategien*, die Du und List (2020) und mit anderer Bezeichnung, aber ähnlichem Zugang ebenfalls van der Graaf et al. (2022) getroffen haben, fanden sie für beide Gruppen

positive Zusammenhänge mit der Textqualität, was insofern bemerkenswert war, als dass das Auftreten bzw. die Häufigkeit beider Arten von Strategien teils völlig unabhängig voneinander war (Du & List, 2020) bzw. sogar ein stark negativer Zusammenhang bestand (van der Graaf et al., 2022). Häufigere Oberflächenstrategien korrelierten bei Du und List (2020) ferner mit mehr Diskurskonnektoren; mehr tiefenorientierte Strategien hingen mit mehr Zitationen zusammen. Damit ist – zumindest in diesen beiden Studien – für beide Gruppen von Strategien ein überwiegend positiver Zusammenhang mit zwei von drei untersuchten Produktmaßen nachweisbar gewesen.

5.2.3 Perspektive 2: Schreibprozesse im Rahmen lesebezogener Aufgaben am Beispiel der Diskurssynthesen

In diesem Teilkapitel werden zwei Bereiche der Prozesse behandelt. Dabei handelt es sich um die unspezifischen Prozessmaße (5.2.3.1) und die spezifischen Prozessmaße (5.2.3.2). Die letztgenannte Gruppe stellt differenziertere und auch mehr Befunde zur Verfügung und ist darum umfangreicher in der Darstellung.

5.2.3.1 Unspezifische Schreibprozessmaße

Unspezifische Schreibprozesse wurden auf zweierlei Arten fokussiert, von denen die erste die präzisere ist:

- *Schreibzeit*, also die erfasste Zeitspanne, in der Personen schreibbezogene Aktivitäten ausgeführt haben bzw. sich in den dafür vorgesehenen Bereichen von Apps aufgehalten haben. Das kann die erfasste Zeit in der Maske für das Schreiben in digitalen Apps sein (Bråten, Brante & Strømsø, 2018; Du & List, 2020; Schoor et al., 2023) bzw. die erfasste Dauer des Schreibens bei Paper-and-Pencil-Aufgaben, teils noch untergliedert in Häufigkeit und ausgewertet in Bezug auf die Anteile des Schreibens an der Gesamtbearbeitungsdauer (Holdinga et al., 2021). Mitunter wurde die Gesamtzeit der Schreibphasen nochmals untergliedert in Lese-, Mapping- und Schreibzeit (Barzilai et al., 2023).
- Ein im Vergleich zum Spiegelstrich zuvor ungenaueres, in der Empirie gleichwohl häufig antreffbares Maß ist die *Gesamtdauer der Aufgabenbearbeitung* beim Herstellen von Diskurssynthesen, das nicht mehr Lesen und Schreiben trennt. Einige Studien lieferten nur diese Gesamtdauer als Summenwert aus verschiedenen Lese- und Schreibaktivitäten (Ackerman, 1991; Spivey, 1992; Spivey & King, 1989) oder summierten Werte der einzelnen Prozesse nochmals, um sie auszuwerten (Barzilai et al., 2023).

Unspezifische Schreibprozesse 1: Dauer der reinen Schreibaktivitäten

Längere Bearbeitungszeiten rein bezogen auf die Schreibaktivitäten korrespondieren mit

- längeren Textprodukten (Bråten, Brante & Strømsø, 2018),
- mehr übernommenen Inhalten aus den gelesenen Dokumenten (Barzilai et al., 2023; Bråten, Brante & Strømsø, 2018),
- mehr in den Diskurssynthesen verwendeten Dokumenten (Barzilai et al., 2023),
- mehr Wechseln zwischen den Dokumenten (Switches; Bråten, Brante & Strømsø, 2018),
- mehr Diskurskonnektoren in der Diskurssynthese (Du & List, 2020),
- mehr Referenzierungen (Barzilai et al., 2023; Bråten, Brante & Strømsø, 2018; Du & List, 2020) sowie
- höherer Textqualität (Du & List, 2020; Holdinga et al., 2021).

Daneben erweisen sich auch die Häufigkeiten der Schreibaktivitäten und ein hoher Anteil des Schreibens an der Gesamtbearbeitungsdauer als günstig für die Textqualität (Holdinga et al., 2021). Die Schreibdauer ist also mit einer Vielzahl von Produktmaßen positiv assoziiert und verdeutlicht, dass eine längere Dauer des Schreibprozesses sich günstig in den Diskurssynthesen niederschlägt.

Unspezifische Schreibprozesse 2: Dauer der Gesamtbearbeitung bei Aufgaben zum Herstellen von Diskurssynthesen

Neben empirischen Zugängen zur spezifischen Schreibzeit gibt es Studien, die die Gesamtbearbeitungszeiten inklusive der Schreibzeit erfassen. Je länger Gesamtzeiten ausfielen,

- desto mehr Inhalte enthielten die Texte, darunter jene aus den zu verarbeitenden Dokumenten (Barzilai et al., 2023; Spivey, 1992; Spivey & King, 1989),
- desto länger wurden die Diskurssynthesen (Schoor et al., 2023),
- desto mehr Dokumente wurden für die Diskurssynthese genutzt (Barzilai et al., 2023),
- desto mehr Referenzierungen waren in den Diskurssynthesen enthalten (Barzilai et al., 2023; Schoor et al., 2023),
- desto eher waren grammatische Bestandteile enthalten, welche Dissens bzw. den Einbezug mehrerer Dokumente markierten (Schoor et al., 2023),
- desto kohärenter wurden die Themen entfaltet (Spivey & King, 1989),
- desto mehr entsprach das Textprodukt einem kontinuierlichen Fließtext (Schoor et al., 2023) und

- als desto besser wurde die Textqualität beurteilt (Holdinga et al., 2021; Spivey, 1992; Spivey & King, 1989).

Einige Studien fanden jedoch auch keine Zusammenhänge von Gesamtbearbeitungszeit und Produktmaßen (Ackerman, 1991), insbesondere Texttiefeigenschaften (Barzilai et al., 2023). Gleichwohl zeigt sich hierin ein prinzipiell ähnliches Muster wie bei der reinen Schreibzeit: Je mehr Zeit Personen in den Gesamtprozess ihrer Diskurssyntheseproduktion insgesamt investierten, um so hochwertiger wurde in der Regel auch ihr Text in verschiedenen Indikatoren.

5.2.3.2 Spezifische Schreibprozessmaße

Überblick: Welche vier Gruppen der Schreibprozessmaße wurden in Studien erfasst?

Die Studien nutzten unterschiedliche Prozessindikatoren, um spezifische Prozesse des Schreibens zu erfassen. Vier Gruppen kann man extrahieren: Planungs- sowie Revisionsaktivitäten als zwei Hauptgruppen der Schreibstrategien (Alamargot & Chanquoy, 2001), Bündelungen verschiedener Strategien zu Gesamtindizes und schließlich noch Vorgehensweisen mit hoher metakognitiver Bewusstheit für das Schreibziel, die von den vorherigen drei Kategorien nicht abgedeckt waren.

1) Planungsaktivitäten

Planungsaktivitäten dienen beim Verfassen von Diskurssynthesen dazu, Inhalte aus gelesenen Dokumenten und dem eigenen Vorwissen bezogen auf den kommunikativen Anlass angemessen auszuwählen und zu strukturieren.

Die erste Gruppe von Maßen betrifft die Länge bzw. eingeschätzte Häufigkeit von planerischen Aktivitäten, darunter die Länge aller Planungsaktivitäten in Lautdenk-Protokollen (Holdinga et al., 2021), die Verweildauer in der Planungsmaske in einem digitalen Tool zum Herstellen von Textplänen (Raković, Fan et al., 2022) oder die retrospektiv eingeschätzten Planungsstrategien (Chang, 2023; Golparvar & Khafi, 2021; Yang, 2014) und – als eine eher weniger wünschenswerte Vorgehensweise – die oberflächlichen Strategien zum Finden und Übernehmen von isolierten Inhalten aus gelesenen Texten, um diese später in die eigene Diskurssynthese als Versatzstücke unverändert einzupassen („Test Wiseness Strategy Use“; Yang & Plakans, 2012).

Genauer in ihren Analysen und an transitorischen, von den Testpersonen hergestellten *Zwischenprodukten* orientiert sind Maße einer zweiten Gruppe von Analysezugängen, die Zwischenprodukte als Indikator des Planens prüften. Diese Studien nutzten diverse Maße aus den Notizen. Es lassen sich zwei übergeordnete Gruppen von Analysegegenständen aus den Studien abstrahieren. Die

erste, zahlenmäßig größere und in sich heterogenere Gruppe vereint in sich Maße, die Notizen nach verschiedenen Bezugspunkten auf Informationsauswahl und mitunter auf die Verbindungen von Inhalten, Metadaten, ganzen Dokumenten bzw. mit dem Vorwissen analysierten. Damit nimmt diese Gruppe grosso modo das auf, was in den Transformationsprozessen in Teilkapitel [5.1.2.1](#) als Auswählen und Verknüpfen beschrieben wurde. Die zweite Gruppe widmete sich restrukturierenden Vorgängen der Informationsorganisation.

Innerhalb der *ersten Gruppe mit Auswahl- und Verknüpfungsindikatoren* erfassten einige Studien die inhaltlichen Aussagen, die aus den gelesenen Texten in Notizen übernommen wurden (Du & List, 2020; List & Du, 2021), hierunter fallen ebenfalls zusammenfassende Notizen (Kobayashi, 2009). In der Studie von Kobayashi (2009) wurden Notizen ferner danach klassifiziert, ob sie zum einen evaluativen Charakter hatten und ob es zum anderen explanatorische Notizen gab. Es handelte sich also um vorwissensbasierte Anreicherungen bzw. Reaktionen auf die intratextuellen Inhalte. Eine besonders detaillierte Klassifikation von Notizen kam in der Studie von List und Lin (2023) vor, in der Notizen nach verschiedenen Transformationen (Paraphrasen, Elaborationen, Kategorisierungen und intertextuellen Inferenzen) und zudem noch hinsichtlich der Evaluationen kodiert wurden.

Die *Verbindung von Metadaten und Inhalt* wurden über Häufigkeitszählungen über berücksichtigte Dokumente erfasst. Dies wurde über die Anzahl von Dokumenten mit Metadaten in Notizen (Du & List, 2020; List & Du, 2021) bzw. die Häufigkeit, mit der in Notizen Inhalte und Metadaten kombiniert wurden (Du & List, 2020), operationalisiert.

Einige der Studien gingen dezidiert auf *intertextuelle Maße* ein. So wurde etwa erfasst, ob die Testpersonen intertextuelle Notizen angefertigt hatten, welche z. T. inhaltliche Transformationen beinhalteten (Du & List, 2020; Kobayashi, 2009) bzw. anhand derer sich rekonstruieren ließ, wie viele unterschiedliche Dokumente in den Notizen verarbeitet worden waren (Du & List, 2020; List & Du, 2021).

Unter die *zweite Gruppe von Studien mit Organisationsprozessen bzw. deren Folgen* in den Notizen fallen Untersuchungen, die das binäre Unterscheiden des Vorhandenseins angefertigter planerischer Notizen (Spivey, 1991) oder eine auf Skalenniveau erfolgte Einschätzung von strukturierten Planungsnotizen mit Hinweisen auf Transformation der Inhalte (Risemberg, 1996) vorgenommen haben. Teils geschah dies in Verbindung mit Unterstreichungen und Annotationen (Spivey, 1992; Spivey & King, 1989). Zudem wurden Notizen beim Planen auch mehrdimensional in ihrer Qualität eingeschätzt (H. Kim, 2020). Eine die Prozesse des Notizenanfertigers genauer fokussierende Analyse besteht darin, dass in einer Kombination von Zwischenprodukt- und Prozessdaten überprüft wurde, ob Notizen linear im Sinne eines Notierens pro Text oder eher non-linear durch Modifikationen und/oder thematische Gruppierungen von Notizen angefertigt wurden (Gil et al., 2011). Erfasst wurde zudem, ob die Reihenfolge der

Inhalte aus den Notizen eine andere war als in den finalen Textprodukten, also ob es zu weiteren Restrukturierungen im Prozess des Schreibens gekommen ist (Du & List, 2021).

Zudem wurden elektronisch erstellte *grafische Notizen* („Mapping“ genannt), detailliert mit verschiedenen Indikatoren wie der Anzahl der Ideen, der Anzahl von Dokumenten, intra- und intertextuellen Inferenzen, dem Grad an Vollständigkeit von Inhalten für die verlangte Zieltextsorte Argumentation und schließlich die Vollständigkeit als grafisches Dokumentenmodell ausgewertet (Barzilai et al., 2023). Eine besonders detaillierte Analyse von digital erstellten Maps haben schließlich Binks et al. (2022) vorgenommen. Sie extrahierten 33 Indikatoren aus den Maps, die sie vor allem mit Blick auf Transformationen von Informationen im Rahmen der planerischen Aktivitäten untersuchten.

2) Revisionsaktivitäten

Revisionsaktivitäten dienen dazu, be- und entstehende Diskurssynthesen zu optimieren, indem inhaltliche Veränderungen vorgenommen werden. Erfasst wurden handlungsnah am Text selbst, welche *Zahl von inhaltlichen Revisionen* am Text beim Schreiben des Textes realisiert wurden (Du & List, 2020; Spivey, 1992; Spivey & King, 1989), bzw. die zeitliche Dauer von Veränderungen am Text (Holdinga et al., 2021; H. Kim, 2020; Raković, Fan et al., 2022). Daneben wurden immer wieder auch retrospektiv eingeschätzte *Revisionsstrategien in Fragebögen* (Chang, 2023; Golparvar & Khafi, 2021; Yang, 2014), teils in Verbindung mit metakognitivem Überwachen (Selbstregulationsstrategien) (Yang & Plakans, 2012) erhoben, die eine andere, handlungsfernere Facette des Revidierens fokussierten.

3) Bündelungen/Aggregationen von verschiedenen Schreibstrategien

Die Trennung von funktionsbezogenen Aktivitäten ist nicht durchgängig in Studien vorgenommen wurden; einige von ihnen nutzten als dritte Variablengruppe die *Bündelungen von verschiedenen Schreibstrategien* – in diesem Fall sind verschiedene Prozesse zu Gesamtindizes zusammengefasst worden, d.h. verschiedene Strategien und Prozesse mit variierenden einzelnen Funktionen wurden aggregiert. Ein Beispiel hierfür ist die „Informationsverwendung“ als Summenwert primär von genutzten Notizen und Übernahmen von Informationen aus den gelesenen Dokumenten in eigene Diskurssynthesen (Du & List, 2020). Erfasst wurde zudem unter der Bezeichnung „mentales Modell“, wie Personen versuchen, argumentativ kohärente Texte mit divergierenden Aussagen herzustellen (Chang, 2023). Ein wiederum anderes Exempel stellt die „Herstellung des Textes“ dar, die basierend auf Summen von Planungs-, Strukturierungs- und

Revisionsaktivitäten in Lautdenk-Protokollen aggregiert ausgewertet wurde (Du & List, 2020). Dies wurde ähnlich auch mit Fragebogenauskünften in Strukturgleichungsanalysen getan (Liao et al., 2021; Zhao et al., 2023).

4) *Metakognitive Bewusstheit zur Zielerreichung*

Die vierte Gruppe ist die *metakognitive Bewusstheit zur Zielerreichung*, was hier ein Sammelbegriff für Aktivitäten darstellt, die darauf hinweisen, dass Personen Informationsquellen gezielt aufsuchen, die bei der Lösung der Aufgabenstellung jenseits des eigentlichen Inhalts helfen können, bzw. Aktivitäten, die auf einen klaren Zielbezug beim schreibbezogenen Vorgehen Rückschlüsse erlauben – beiden Varianten ist gemein, dass sie metakognitive Steuerungen des Vorgehens implizieren.

Die Konstrukte, die zum Einsatz gelangten, sind vielfältig und reichen über die „rhetorische Bewusstheit“, verstanden als Lautdenk-Protokoll-Aussagen mit Fokus auf Konventionen des Schreibens (Ackerman, 1991), die Häufigkeit metakognitiver Strategien während des Schreibens gemäß Lautdenk-Protokollen (Du & List, 2020), die Suche nach Informationen zur Aufgabenlösung, erfasst über die Dauer der Konsultation dieser Informationsquellen (Risemberg, 1996), bis hin zu Überwachungsaktivitäten beim Schreiben durch wiederholtes Konsultieren von aufgabenrelevanten Informationen (Raković, Fan et al., 2022).

Diese vier Gruppen von Variablen werden im Folgenden sukzessive betrachtet und mit den Mustern bei den Produktmaßen präsentiert.

Spezifische Schreibprozesse 1: Planungsaktivitäten

Die Planungsaktivitäten beim Verfassen der Diskurssynthesen werden in diesem Teilkapitel zweigeteilt betrachtet. Im ersten und wegen der empirischen Datenlage erheblich kürzeren Teil sind Prozessdaten bzw. retrospektive Selbstauskünfte von Interesse. Hierbei ist es auffällig, dass die Produktmaße allesamt die Textqualität betreffen oder Gesamtindizes aus mehreren Teildimensionen darstellen. Der zweite Teil geht auf die beim Planen entstehenden Zwischenprodukte in Form von Notizen bzw. „Maps“ ein. Dort ist ein konsistentes Muster, dass sich grob zwei Arten von Befundlagen herauskristallisieren: Auf der einen Seite ist die Aufnahme von Informationen in Notizen ein Indikator, aber kein hinreichender Garant dafür, dass die eigenen Texte eher oberflächenorientiert besser glücken. Auf der anderen Seite haben für Texttiefenmerkmale der Diskurssynthese primär die Struktur und Organisation betreffende Prozesse einen positiven Effekt: Je stärker die Zwischenprodukte Spuren solcher Transformationen beinhalteten, desto günstiger war dies für die Textprodukte.

Befunde vor allem aus retrospektiven Selbsteinschätzungen zum Planen und partiell aus Lautdenk-Protokollen

Die reine Verweildauer in für das Planen vorgesehenen Bereichen von Apps in Experimentalstudien hängt nicht mit der Textqualität zusammen (Raković, Fan et al., 2022). Demgegenüber zeichnen Studien mit einer genaueren nachträglichen Kodierung der Prozesse bzw. der gezielten Befragung zum Vorgehen mittels Selbstauskunftsskalen ein anderes Bild, welches die Bedeutsamkeit des Planens klar herausstellt.

Spezifische Ergebnisse zu den Zusammenhängen von in Fragebögen retrospektiv eingeschätzten Planungs- sowie weiteren kognitiven Prozessen ergaben erstens die Studien von Chang (2023), Golparvar und Khafi (2021) sowie Yang (2014). Diesen Untersuchungen zufolge war in Strukturgleichungsanalysen das allgemeine Planen mit diversen Merkmalen der Diskurssynthesen lediglich indirekt verbunden: Generische Planungsprozesse wurden von spezifischen Prozessen mediiert. Anders gesagt: Planen war nicht per se günstig, sondern bedurfte einer spezifischen strategischen Informationstransformation, damit es sich in aggregierten Gesamtindizes der Diskurssynthesen positiv niederschlug. In einer weiteren Studie (Yang & Plakans, 2012) stellte es sich als ungünstig für den Gesamtindex der Diskurssynthesen heraus, wenn die Testpersonen eine oberflächliche Strategie des Patchwritings genutzt hatten. Ein wenig planvoll wirkendes Vorgehen war demnach negativ mit Leistungen assoziiert.

Nicht ex post eingeschätzte, sondern parallel zur Aufgabenbearbeitung entstandene Prozessdaten gehen – zweitens – bei den Befunden in eine ähnliche Richtung: Häufigere und insgesamt längere Planungsepisoden über den gesamten Prozess der Aufgabenbearbeitung gemäß Lautdenk-Protokollen korrelieren substanziell und positiv mit der Textqualität, zumindest in einzelnen Aufgaben (Holdinga et al., 2021).

Ein genauerer Blick auf hybride Zwischenprodukte als Planungshilfsmittel: Notizen und Mapping

Planungsprozesse wurden auch über Notizen erfasst, also über Zwischenprodukte, die sowohl Lese- als auch Schreibaktivitäten erfordern und damit einen dezidiert hybriden Charakter aufweisen (Peveryly & Wolf, 2019). Bezogen auf das Herstellen von Diskurssynthesen erfüllen Notizen nicht nur die Funktion, Informationen aus den Texten bzw. elaborative Kommentare der lesenden Personen aufzunehmen. Vielmehr spielen sie insofern beim Schreiben eine Rolle, als sie bereits erste Vorarbeiten im Sinne einer strukturierten Planung darstellen können (Arias-Gundín & Robledo, 2023). Ähnliches gilt für das Verwenden des „Mappings“ als zweidimensionale, grafische Elemente nutzende Variante, Inhalte

grafisch zu strukturieren. Gerade der Mapping-Ansatz offeriert reichhaltige Daten, darum ist ihm auch das Beispiel im Kasten „Illustrierendes Beispiel 20“ gebührend Platz gewidmet.

Das Anfertigen von Notizen wurde vergleichsweise häufig in Studien untersucht. Hierbei sind verschiedenste Maße zum Einsatz gekommen. Eine erste Gruppe von Studien prüfte, *wie viele Inhalte aus den Dokumenten übernommen* wurden (Du & List, 2020; List & Du, 2021), teils sogar mit Fokus auf wortwörtliche und damit oberflächliche Übernahmen (Castells et al., 2023; Neumann et al., 2020). Diese Studien kommen zu teils unterschiedlichen Befunden:

- *Null-Befunde:* In den Studien von Du und List (2020), Castells et al. (2023), List und Lin (2023) sowie Payant et al. (2019) bestanden keine Zusammenhänge mit Textmaßen.
- *Negative Zusammenhänge bei wortwörtlichen Übernahmen ohne Transformationen:* Andere Studien berichten negative Zusammenhänge. Castells et al. (2023) fanden einen negativen Zusammenhang von vielen wortwörtlichen Übernahmen von Inhalten in Notizen und einer kohärenten Themenentfaltung vor. Dieses Muster berichteten hinsichtlich der Textqualität ebenfalls Neumann et al. (2020). In den Studien von List und Du (2021) sowie Castells et al. (2023) gingen mehr übernommene Inhalte in den Notizen generell mit Diskurssynthesen einher, deren Inhalte verstärkt wortwörtlich aus den gelesenen Dokumenten stammten. Die Notizen erfüllten hier also die reine Funktion eines Zwischenspeichers.
- *Positive Zusammenhänge bei transformierten Notizeninhalten:* Kobayashi (2009) konnte zeigen, dass eine Variante von Notizen, nämlich jene mit Zusammenfassungen der einzelnen Texte, mit besseren schriftlichen Leistungen der intertextuellen Kohärenzbildung korrespondierte. Hierbei war entscheidend, dass lediglich unterschieden worden war, dass es Notizen gab, die Zusammenfassungen enthielten, nicht aber, wie viele und in welcher Qualität. Neumann et al. (2020) ermittelten ganz ähnlich, dass die Textqualität dann besser war, wenn die Notizen bereits Transformationsleistungen in Form von paraphrasierten Inhalten enthielten.

Alles in allem wirkt es so, dass das Abdecken von gelesenen Inhalten, teils in Bezug auf bereits forcierte Unterscheidung von wichtigeren und weniger wichtigen Propositionen, nicht zwingend ein Garant dafür ist, dass bessere Textprodukte entstehen. Es gibt vereinzelte Hinweise aus den Studien zu positiven Effekten, doch es fällt gesamthaft auf, dass sie keine Mehrheit darstellten und sich zudem, zumal bei einem kleinen Pool an aktuell verfügbaren Studien, kein klares Gesamtmuster extrahieren lässt.

Bei den *rekonstruierbaren Metadaten aus einzelnen Dokumenten* als zweitem Schwerpunkt ergaben sich in zwei Studien, in denen sie überprüft wurden (Du &

List, 2020; List & Du, 2021), konvergente Ergebnisse: Enthielten die Notizen mehr Metadaten, dann beinhalten die Diskurssynthesen später auch substantiell mehr Referenzierungen. In beiden Studien gab es keine weiteren Zusammenhänge mit anderen Produktmaßen. In beiden Ergebnissen, sowohl den vorhandenen als auch den nicht vorhandenen Korrelationen, spiegelt sich wider, dass die Notizen die Enkodierungsfunktion von Informationen, in diesem Fall: Metadaten, erfüllen. Kurzum: Notizen dienen der Aufnahme von Informationen, um sie später wieder zu verwenden.

Vereinzelte prüften in einem dritten Zugang Studien, in welchem Grad *intertextuelle Inferenzen in den Notizen* vorkamen, wobei sich die Genauigkeit der Maße stark unterschied. So berichtete Kobayashi (2009) grob als binäre Kategorie, ob es solche intertextuellen Notizen gab, während Du und List (2020) auf einer Skala einschätzten, in welchem Maß intertextuelle Verbindungen hergestellt worden waren. In beiden Studien waren intertextuelle Notizen mit besseren Leistungen verbunden, nämlich einer genaueren Repräsentation der intertextuellen Bezüge in den Texten (Kobayashi, 2009) bzw. einer höheren Textqualität (Du & List, 2020). Intertextuelle Inferenzen in den Notizen hingen bei List und Lin (2023) hingegen nicht mit zwei Produktmaßen (Anzahl übernommener Argumente und Anzahl der Diskurskonnektoren) zusammen.

Wurde hingegen aus den Notizen extrahiert, *wie viele Dokumente sie abdeckten*, dann ergab sich ein eher unklares Bild. Du und List (2020) fanden keine Zusammenhänge mit Produktmaßen, List und Du (2021) ermittelten positive Zusammenhänge. Je mehr Dokumente in den Notizen vorkamen, um so länger waren die Diskurssynthesen zum einen und desto mehr Inhalte aus den Dokumenten flossen zum anderen in sie ein.

Daneben wurden – viertens – auch die Inhalte von Notizen vereinzelt auf *Ähnlichkeiten von Notizinhalt und Diskurssynthesen* abgeglichen (Castells et al., 2023; Du & List, 2021; Gil et al., 2011). Gemäß den Studien von Castells et al. (2023) und Gil et al. (2011) ist ein erstes Ergebnis, dass sich die Arten von Informationen in Notizen teils eng gekoppelt mit den Inhalten der Diskurssynthesen erwiesen haben. Sowohl a) falsche als auch b) wörtliche und c) inhaltliche transformierte (in diesem Fall: inferenzbasiert verknüpfte) Inhalte aus den Notizen gehen mit dem Vorhandensein derselben Arten von Informationen in den Texten einher, was dafür spricht, dass Inhalte aus den Notizen recht häufig unverändert in Texte einfließen. Die Studie von Gil et al. (2011) prüfte zudem noch den *Prozess der Notizenveränderung*. Konkret wurde mittels Trackingdaten in der digitalen Lese- und Schreibumgebung ermittelt, ob die Testpersonen Notizen eher linear oder non-linear erstellt hatten. Linear bedeutet, dass die Notizen in der Sequenz der gelesenen Texte angefertigt worden waren. Non-linear waren die Notizen dann, wenn sie im Laufe der Zeit verändert wurden bzw. wenn Notizeninhalte thematisch sortiert worden waren. War dies der Fall, enthielten die Diskurssynthesen zum einen mehr intratextuelle Inferenzen und

zum anderen mehr transformierte Inhalte. Eine zweite Studie zeigte – wenn auch in anderen Indikatoren der Produkte – eine Überlegenheit der Transformation der Notizeninhalte. Wenn die Reihenfolge der notierten Inhalte von derjenigen in der späteren eigenen Diskurssynthese abwich, erwies sich das in der Studie von List und Du (2021) in zwei Dritteln der sechs Textmaße als günstig. In dem Fall waren die Texte insgesamt länger, enthielten mehr Aussagen aus den gelesenen Dokumenten, mehr Zitationen und vor allem mehr intertextuelle Verbindungen. Damit wirkt es so, als ob Notizen dann hilfreich sind, wenn sie zunächst als Speicher intratextueller Verstehensleistungen angefertigt werden, dann aber im Prozess der Aufgabenbearbeitung transformierend für intertextuelle Verstehensleistungen flexibel genutzt werden.

Studien, in denen als fünfter Zugang schriftliche Notizen hinsichtlich *inhaltlicher Restrukturierungen* der gelesenen Inhalte ausgewertet wurden, kommen mit zwei Ausnahmen (H. Kim, 2020; Spivey, 1992) zum Befund, dass planerische Notizen positive Zusammenhänge mit späteren Texten aufweisen. Dabei ist weniger entscheidend, dass Notizen angefertigt werden, sondern dass diese Notizen die teils transformierten Inhalte rund um intertextuelle Makropropositionen strukturieren und/oder die geforderte Textstruktur der Diskurssynthese antizipieren und dass in dieses Gerüst passende Informationen eingefüllt worden waren. Wenn Notizen dies leisteten, dann

- erzielten die Testpersonen höhere Werte in der Textqualität (Risemberg, 1996; Spivey & King, 1989),
- hatten besser strukturierte Diskurssynthesen verfasst (Spivey, 1991; Spivey & King, 1989),
- entfalteten die Themen kohärenter (Spivey & King, 1989) und
- inkludierten mehr Inhalte in ihre Diskurssynthesen (Spivey & King, 1989).

Illustrierendes Beispiel 20

Mit digitalen Graphic Organizers das Schreiben von Argumentationen vorbereiten und Visualisierungen zur Aufbereitung diverser digitaler Dokumente nutzen

A) Was war das Erkenntnisinteresse?

Welche unterstützende Leistung kann es haben, wenn Personen Graphic Organizer (GO) herstellen, wenn sie einen kontroversen, in mehr als einem halben Dutzend digitalen Dokumenten dargestellten Sachverhalt erkennen und in einem eigenen Text darstellen sollen? Dieser Frage verschrieb sich eine Studie aus Israel (Barzilai et al., 2023).

B) Wer wurde untersucht?

Das Sample waren 40 israelische Jugendliche aus der Sekundarstufe I, mangels Altersangaben kann nur auf die Klassenstufen 7 bis 9 geschlossen werden. Das Geschlechterverhältnis war paritätisch. Die Jugendlichen hatten das digitale Tool zum Erstellen von GO

bereits einmal verwendet, die Daten beziehen sich auf einen zweiten Einsatzzeitpunkt. Es gab keine eigene Einführung, wie das Tool zu nutzen ist, aber die Jugendlichen wurden vorab mit dessen Funktionsweise vertraut gemacht.

C) Welche Instrumente kamen zum Einsatz?

Die Jugendlichen hatten ein Set von sieben kurzen inhaltlich konfligierenden Dokumenten zur Verfügung, über deren Inhalte sie einen argumentativen Text verfassen sollten. Diese Dokumente waren Bestandteil eines digitalen Tools zum Anfertigen der GO mit mehreren Funktionen. Unter diesen Funktionen waren das Herstellen von GO und eine Schreibumgebung, zwischen denen die Jugendlichen wechseln konnten. Das digitale Tool zum Herstellen von GO erlaubte es, die Relevanz und Vertrauenswürdigkeit von Dokumenten zu beurteilen, was grafisch visualisiert wurde. Zusätzlich konnten die Jugendlichen frei GO herstellen und dabei inhaltliche Aussagen frei ergänzen und – besonders wichtig – diese Aussagen mit den Dokumenten per Linien verknüpfen. Auch die Dokumente selbst konnten mittels Linien verbunden werden, wobei die Linien mittels Farben die semantische Relation angaben (Bsp. „Dokument A widerspricht Dokument B“ mit roter Linie; „Aussage A wird von den Dokumenten C und G“ unterstützt mittels grüner Linie etc.).

Das Erstellen dieser GO und die Schreibphase wurden zum einen per Screen-Capturing aufgenommen und zum anderen in einem retrospektiven Einzelinterview mit den Jugendlichen betrachtet, sodass es möglich war, die Dauer von verschiedenen Lese-, GO-Konstruktions- und Schreibaktivitäten zu erfassen und die Ziele zu rekonstruieren. Neben der Dauer dieser drei Hauptaktivitäten wurden auch Wechsel zwischen ihnen kodiert. All dies diente dazu, *Prozessdaten* zu erheben, wobei die Phase der GO-Konstruktion und eine Schreibphase unterschieden wurden.

Die *entstandenen GO* wurden ebenfalls umfassend ausgewertet. So interessierte man sich für a) die Zahl inhaltlicher Aussagen, seien es intra-, seien es intertextuelle Inhalte. Auch die Verbindungen vom Ursprung dieser Aussagen wurden kodiert, nämlich ob sie mit b) einzelnen (intratextuellen) oder c) mit mehreren (intertextuellen) Quelle-Inhalt-Verbindungen visualisiert worden waren. Eine weitere Kodierung betraf d) den Grad an systematischen Verknüpfungen von Inhalten einerseits und Dokumentenrelationen andererseits. Von besonderem Interesse war, ob die GO den Idealzustand eines umfassenden Dokumentenmodells (s. Teilkap. [4.1.2.3](#)) widerspiegeln. Schließlich wurden die GO noch dahingehend ausgewertet, e) ob ihre Inhalte ausreichend ausgewogen für eine Argumentation Positionen und Beispiele inkludierten.

Die *Textprodukte* – Argumentationen zum strittigen Sachverhalt – bildeten schließlich einen dritten Analysefokus. Die Texte wurden danach kodiert, a) wie viele Ideeneinheiten sie beinhalteten, b) wie viele Dokumente zweifelsfrei verwendet worden waren, c) wie viele explizite Angaben/Zitationen zur Herkunft von Aussagen vorhanden waren, d) in welcher Zahl Informationen aus mehreren Dokumenten intertextuell integriert worden waren und e) wie argumentativ ausgewogen die Argumentation entfaltet worden war.

Die reichhaltigen Prozess-, Zwischenprodukt- und Produktmerkmale wurden deskriptiv ausgewertet und in Korrelationsanalysen zueinander in Bezug gesetzt. Dies erfolgte in der Studie bei den Prozessdauern teils sehr differenziert für einzelne Subprozesse, auf die hier aus Platzgründen nicht eingegangen wird.

D) Was waren die Hauptbefunde?

Die *Prozessdaten* illustrierten, dass in der Phase der GO-Konstruktion rund zwei Drittel auf das Lesen entfielen, wobei auch während der Konstruktionsphase noch hohe Leseanteile zu beobachten waren, was sich darin manifestierte, dass es zu vielen Wechseln zwischen dem Lesen der Dokumente und der GO-Konstruktion kam, nämlich zu durchschnittlich fast 17. Die Schreibphase selbst wurde vom Verfassen des Textes dominiert, welches vier Fünftel ausmachte. Dennoch lasen die Jugendlichen noch in den Texten nach und betrachteten ihre eigenen GO. Jeweils sechs Wechsel zwischen Schreiben und Lesen bzw. GO-Konsultation waren im Mittel beobachtbar.

Die GO als *Zwischenprodukte* enthielten durchgängig viele Inhalte, dabei überwogen intratextuelle, also aus einzelnen Dokumenten stammende Aussagen. Bei den Verknüpfungen dominierten hingegen intertextuelle Verbindungen, die visualisierten, dass Aussagen von mehreren Dokumenten unterstützt wurden. In Bezug auf die Dokumentenmodelle ist auffällig gewesen, dass etwas mehr als die Hälfte umfassende Dokumentenmodelle darstellte, also die Jugendlichen inhaltliche Verknüpfungen und Quelle-Inhalt-Verbindungen genutzt hatten.

Die *Textprodukte* waren überwiegend geglückte Argumentationen, die sich auf beide Positionen bezogen und diese abwogen. Dabei wurde mehr als die Hälfte der Dokumente im Text genutzt, aber wenig explizit zitiert und intertextuell integriert. Im Durchschnitt enthielten die Texte ein Dutzend Ideeneinheiten.

Die zentralen Zusammenhänge zwischen den drei fokussierten Hauptkategorien sind in Abbildung 19 dargestellt, nämlich als mit Linien visualisierte Korrelationen. In der Grafik sind die vorhandenen Interkorrelationen bei den (Zwischen-)Produktmaßen aus Gründen der Komplexitätsreduktion ausgespart. Die Darstellung der zentralen Ergebnisse zeigt komplexe Ergebnismuster. Für den Gesamtscore der Argumentationen im Text war eine längere Lesedauer günstig, die sich vermittelt über die Aufnahme vieler Inhalte und vollständiger Positionen in den GO nachvollziehen lässt. Daneben wurde die Ausgewogenheit dadurch begünstigt, wenn die Konstruktionsphase länger gedauert hatte. In den GO selbst nahm die Zahl der Inhalte offenbar eine zentrale Position ein, da diese mit mehr Inhalten in der späteren Argumentation, mehr intertextuellen Integrationen und wie schon erwähnt ausgewogeneren Argumentationen korrespondierte. Je inhaltlich vollständiger der GO war, desto inhaltlich umfassender, vernetzter und textsortenkonformer geriet später der Text. Demgegenüber waren andere Merkmale der GO für den eigenen Text eher unbedeutend. Daneben gab es direkte Zusammenhänge von Prozessindikatoren und Text: Häufige Wechsel zwischen den Hauptprozessen schlugen sich – wie eine längere Schreibphase – in gleich mehreren Textmerkmalen positiv nieder. Wer seinen Schreibprozess besser orchestrierte und länger aufrechterhielt, zitierte mehr und nutzte mehr Dokumente. Wer

länger schrieb, band mehr Informationen ein. Damit scheint das Herstellen eines GO im Rahmen eines Schreibauftrags vor allem der Vollständigkeit von Informationen zuträglich, die Personen in ihren GO und Texten verwenden und aufeinander beziehen.

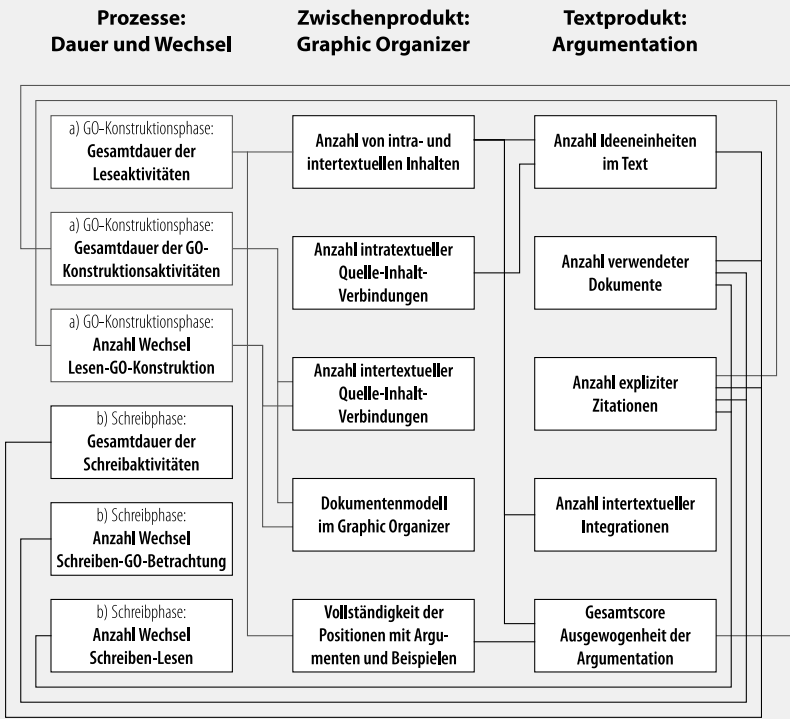


Abbildung 19: Zusammenhänge von Prozess-, Zwischenprodukt und Produktdaten beim argumentativen Schreiben von Jugendlichen (eigene Darstellung, basierend auf Barzilai et al., 2023, S. 830 und 834, Verbindungslinien geben signifikante Korrelationen an)

E) Diskussion

Die israelische Studie ist eine der ersten, die sich genuin damit auseinandersetzen, wie (ohne weitere didaktische Intervention) der Nutzen von GO für das Aufbereiten gelesener Dokumente konturiert werden kann. Die reichhaltigen Befunde illustrieren spezifische Muster, welche Prozesse in Dauer und Wechsel mit Eigenschaften grafisch organisierter Zwischenprodukte und späteren Textprodukten korrespondieren. Auffällig ist, dass sich die untersuchten Personen vor allem darauf konzentrierten, die einzelnen Texte zu erschließen, und dennoch waren ihre GO mehrheitlich einer sehr elaborierten mentalen Repräsentation zuordenbar, nämlich dem Dokumentenmodell. Diese elaborierte Visualisierung war jedoch nicht mit späteren Textmaßen korreliert – möglicherweise ist dies über nicht im Erkenntnisinteresse der Studie stehende weitere Transformationsprozesse erklärbar. Dennoch ist es den Jugendlichen auffällig gut gelungen, ausgewogene Argumentationen auf der Basis der gelesenen Dokumente zu verfassen.

Ein weiterer besonders wichtiger Befund liegt darin, dass nicht nur Prozessdauern mit (Zwischen-)Produkten positiv assoziiert sind, es scheinen vor allem die (selbstregulierten) Wechsel zwischen Hauptprozessen zu sein, die mit besseren Leistungen zusammenhängen. Dieses Resultat ist deshalb so wichtig, weil es auf eine Überlegenheit eines rekursiven, nicht-linearen Vorgehens deutlich hinweist, das in der Forschung als grundsätzlich nötig erachtet wird, um komplexe Probleme wie das Herstellen eines Aufgabenprodukts auf der Basis multipler (digitaler) Dokumente erfolgreich zu bewältigen.

Das *Mapping* schließlich wurde in zwei Studien untersucht, in denen es darum ging, argumentative Diskurssynthesen zu verfassen (Barzilai et al., 2023; Binks et al., 2022). Die digital entstandenen Planungsnotizen mit grafischer Anordnung, also die Maps, wurden in beiden Studien unterschiedlich intensiv ausgewertet und in einer der beiden (Barzilai et al., 2023) mit weiteren Prozessdaten in Beziehung gesetzt (s. Kasten „Illustrierendes Beispiel 20“). Beide Studien kommen trotz unterschiedlicher Methodik zu teils übereinstimmenden Befunden. Ein erstes Muster betrifft die Selektion von Inhalten und deren spätere Verwendung: Je mehr Inhalte aus den gelesenen Texten die Maps enthielten,

- desto mehr Ideeneinheiten fanden sich in den Diskurssynthesen wieder (Barzilai et al., 2023),
- desto besser waren auch die intertextuellen Integrationsleistungen (Barzilai et al., 2023),
- desto ausgewogener wurde der kontroverse Sachverhalt dargestellt (Barzilai et al., 2023; Binks et al., 2022) und desto effektiver waren die Texte strukturiert (Binks et al., 2022),
- desto überzeugender wurde die gesamte Argumentation (Binks et al., 2022).

Ein zweites Muster betrifft die Restrukturierung von Inhalten mittels und infolge der Map-Nutzung. Die Studie von Binks et al. (2022) erbrachte mehrere, in sich konsistente Befunde. In vier von fünf Textproduktmaßen – Überzeugungskraft der Argumentation, Textstruktur, Ausgewogenheit der Positionen und schließlich der Textqualität – erwies es sich als günstig, wenn die Map-Inhalte Informationen enthielten, die aber später im Text weiter ausgebaut worden waren, indem dort aus den einzelnen Inhalten der Maps mehrere Aussagen und Bestandteile extrahiert wurden. Waren hingegen die Maps reine Listen von Inhalten, die ohne Veränderung übernommen worden waren, dann hing dies negativ mit der Textstruktur zusammen. Für die Mehrperspektivität war es zudem ungünstig, wenn die Testpersonen ihre Maps als rein lineare Abfolge erstellt hatten. Diese negativen Korrelationen sprechen dafür, dass ohne Restrukturierungen teils Texte sogar schlechter glücken.

Ein letzter Befund, der der Vollständigkeit halber erwähnt wird: Hatten die Maps bereits deutlich mehr Verbindungen von Dokumenten und deren Inhalten,

fürte dies später zu inhaltlich vollständigeren Diskurssynthesen (Barzilai et al., 2023).

Zusammenfassend lässt sich für Notizen konstatieren: Die Übernahme von verschiedenen Informationen und teils ihrer vorwissensbasierten Anreicherung – also Ergebnisse vorderhand von Auswahl- und Verknüpfungsprozessen – hat sich teils als vorteilhaft erwiesen. Hier scheinen sich Notizen vor allem bei eher lokal zu nennenden Produktindikatoren positiv auszuwirken. Im weitesten Sinn dienen die Notizen als Vorarbeiten, als Reservoir für später nutzbare Informationen. Für Produktmaße, die eine höhere Verarbeitungstiefe erfordern, genügen Maße nicht mehr, die sich auf einzelne Informationen bzw. singuläre Notizeninhalte erstrecken. Solche die intertextuelle Kohärenz bzw. die Texttiefenmerkmale adressierenden Prozesse erfordern das Erkennen und (Re-)Organisieren, und hier unterstreichen die ausgewerteten Studien hinsichtlich textstruktureller bzw. -struktureller Prozesse und Produkte eine Kongruenz.

Die Ergebnisse aus Studien mit Planungsnotizen und Mapping als Planungshilfe ergeben komplementäre Befunde: Der Abdeckungsgrad zum einen und Transformationstendenzen zum anderen scheinen damit bei den schriftlichen Planungen entscheidend dafür zu sein, dass Diskurssynthesen in verschiedenen Indikatoren als positiver beurteilt worden sind. Entscheidend ist also, dass die Planungsaktivitäten indizieren, dass Personen bereits beim Anfertigen der hybriden Zwischenprodukte Verarbeitungsprozesse durchgeführt haben, die gleichsam die Basis für weitere, sich im Schreibprozess vollziehende Inhaltstransformationen bilden. Dieses Befundmuster ist nicht auf das Schreiben von Diskurssynthesen limitiert, sondern wurde auch in weiteren Bereichen der Bildungsforschung für das eigene Herstellen von strukturierten Notizen für Lese- und Schreibleistungen ermittelt (Ponce et al., 2023).

Spezifische Schreibprozesse 2: Revisionsaktivitäten

Revidieren dient beim Schreiben allgemein dazu, Probleme in entstehenden Texten zu erkennen und sie zu beheben (Fitzgerald, 1987). Dieser Prozess des Schreibens ist in den Studien nicht nur seltener untersucht worden als das Planen, es ergibt sich zudem auch kein kohärentes Gesamtbild über die Studien hinweg.

- *Positive Zusammenhänge von Prozessen und Produkten:* Je länger die Revisionsaktivitäten ausfallen, desto mehr Inhalte werden in Diskurssynthesen abgedeckt (Raković, Fan et al., 2022), und teils bestehen diese Zusammenhänge in einzelnen Zeitabschnitten und bei einzelnen Aufgaben bei der Textqualität (Holdinga et al., 2021). Daneben erwiesen sich mehr inhaltlich substanzielle Revisionen – im Vergleich zu sprachformalen Revisionen – als günstig für Texttiefenmerkmale der Diskurssynthesen (H. Kim, 2020).

- *Indirekte Effekte bei Fragebogenstudien und mehrdimensional gebildeten Produktmaßen:* Studien mit Strukturgleichungsmodellen auf der Basis von Fragebogenselbstauskünften zum Vorgehen fanden das Muster, dass Revidieren (als „Evaluieren“ bezeichnet) über andere Prozesse des Schreibens mediiert wird (Chang, 2023; Golparvar & Khafi, 2021; Yang, 2014; Yang & Plakans, 2012). Das bedeutet, dass Revidieren erst in Zusammenhang mit Integrationsprozessen zu positiven Effekten auf das Textprodukt indirekt beiträgt.
- *Nulleffekte:* Studien mit genaueren Analysen der Anzahl inhaltlicher Revisionen kommen hingegen zu konträren Befunden: Hier gab es keine Zusammenhänge mit den in den Studien analysierten Produkteigenschaften und der Zahl inhaltlicher Modifikationen (Du & List, 2020; Spivey, 1992; Spivey & King, 1989).

Angesichts dieser Befundlage kann nur vorsichtig extrapoliert werden, dass das Revidieren partiell zu besseren Diskurssynthesen führt.

Spezifische Schreibprozesse 3: Bündelung von Schreibstrategien zu übergeordneten Indizes

In vereinzelt Fällen in der Empirie wurden Strategien des Schreibens nicht mehr in ihren Funktionen unterschieden, sondern aggregiert. Chang (2023) wies mit einem Gesamtscore zum Herstellen intertextueller Kohärenz bei Argumentationen nach, dass diese retrospektiv eingeschätzten Prozesse die Textqualität in Strukturgleichungsanalysen positiv prädizierten. Du und List (2020) summierten verschiedene Schreibaktivitäten zu zwei Indizes, nämlich Informationsverwendung und Produktherstellung. Keines dieser Maße hing mit den drei fokussierten Textmerkmalen in ihrer Studie in Korrelationsanalysen zusammen, weder mit der Textqualität noch mit der Zitationsanzahl noch mit der Anzahl der Diskurskonnektoren. In Regressionsanalysen war hingegen die Summe aus Planungs- und Revisionsstrategien als Teil des Herstellens des Textprodukts günstig für die Textqualität. Die (inhaltliche) Verwendung der Dokumente und Informationen als zweiter Index erwies sich allerdings als nachteilig für die Zahl der Zitationen in den Diskurssynthesen. Bei den Diskurskonnektorenanzahlen ergaben sich positive Beiträge beider Prozessgesamtindizes, jedoch nur unter Berücksichtigung ihres Auftretens in Dritteln des Gesamtprozesses.

In zwei anderen Studien (Liao et al., 2021; Zhao et al., 2023) wurden vier verschiedene Schreibstrategien als Konstrukt „integrierte Schreibstrategien“ in einem Messmodell gebündelt und dann mit Schreibproduktmaßen, ebenfalls als ein latentes Konstrukt modelliert, statistisch in Beziehung gesetzt. Hier ließen sich direkte und positive Effekte auf die Schreibprodukte nachweisen.

Aus Logfiledaten bildeten schließlich Raković et al. (2023) einen Index, den sie als „Elaboration/Organisation“ als Teil selbstregulatorischen Vorgehens

bezeichneten. Dieser Index umfasste das Anfertigen von Notizen und das allmähliche Ergänzen und Modifizieren von textuellen Informationen, vor allem aber Schreibaktivitäten im Verbund von zuvor stattfindenden Zielklärungen. Dieser Indikator hing stark positiv mit der Textqualität und mit der Textlänge zusammen, nicht aber mit der Abdeckung der Themen aus den gelesenen Dokumenten.

Trotz aller Differenz in der Operationalisierung und Auswertung lässt sich festhalten: Ein Vorgehen aus mehreren Strategien erweist sich als günstig für die Qualität der Texte.

Spezifische Schreibprozesse 4: metakognitive Bewusstheit zur Zielerreichung

Schreiben wird als zielgerichtete Aktivität konzeptualisiert, bei der das mit dem Schreiben verbundene rhetorische Ziel konstitutiv ist und die Grundlage bildet für interne Feedbackschlaufen mit Blick auf den angestrebten Soll-Zustand des Texts und den gegenwärtigen Ist-Zustand bei entstehenden Texten sowie das aktuelle prozessuale Vorgehen (Hacker, 2018). Der Bezug zu solchen Zielen – sei es als Evaluation des Vorgehens, sei es als erkannte Abweichung, welche andere Vorgehensweise aus Sicht der schreibenden Person erfordert – lässt sich als metakognitive Bewusstheit werten, also als steuernde metakognitive Instanz. In den Studien ist dies untersucht worden, indem qualitative (Ackerman, 1991; Du & List, 2020) bzw. quantitative (Raković, Fan et al., 2022; Risemberg, 1996) Indikatoren des metakognitiv gesteuerten Verhaltens über mehrere Bezugspunkte erfasst und aggregiert wurden.

- *Höhere Ausmaße an Bewusstheit der rhetorischen Zielerreichung korrespondieren mit Übernahmen wichtiger Aussagen und der sprachlichen Transformation übernommener Inhalte.* In einer Studie mit Studierenden wurden die Lautdenk-Protokolle kodiert und mit Produktmaßen korreliert (Ackerman, 1991). Dabei erwies sich die sogenannte „rhetorische Bewusstheit“ beim Schreiben als Verbundkategorie von Prozessen als korreliert mit zwei Produktmaßen, nämlich zum einen der Anzahl implizit übernommener Inhalte aus den gelesenen Texten und zum anderen der Wichtigkeit dieser Inhalte in der Makrostruktur der Aussagen aus den Texten. Je stärker im Prozess Aussagen getätigt wurden, die sowohl planerische als auch revisorische Bezugspunkte hatten, die dazu dienten, den rhetorischen Vorgaben des angestrebten Textprodukts zu entsprechen, desto mehr sprachlich modifizierte inhaltliche Rekurse (= implizit übernommene Inhalte) wiesen die Texte auf und desto mehr von als zentral erachteten Aussagen kamen vor. Anders gesagt: Eine höhere rhetorische Sensibilität im Prozess des Schreibens führte zu einer stärkeren Abdeckung von Inhalten aus den gelesenen Texten, welche bereits sprachliche Transformationen aufwiesen. Andere Produktmaße blieben davon unberührt.

- *Je mehr metakognitive Strategien Schreibende anwenden, desto mehr Referenzierungen enthalten ihre Texte.* Wer in der Studie von Du und List (2020) im Laufe des Verfassens der Diskursynthese mehr metakognitive Überwachungen äußerte, hatte später im Text mehr Zitationen vorgenommen. Textqualität und Anzahl von Diskurskonnektoren hingen nicht signifikant mit metakognitiven Strategien zusammen, auch wenn die Koeffizienten auf schwach positive Korrelationen verweisen.
- *Längere Zeitdauern beim optionalen Prüfen von aufgabenbezogenen Informationen stehen in Verbindung mit qualitativ höherwertigen Diskurssynthesen.* Dies ist das Ergebnis der Studie von Risemberg (1996). Er erfasste, wie lange Personen vor und beim Schreiben auf drei Informationsquellen zugriffen, die Informationen zur Lösung der Aufgabe enthielten, nämlich zwei Mustertexte mit ähnlichem Textaufbau wie jener, welcher für die Aufgabe verlangt war, und einer nicht näher beschriebenen Liste mit Hinweisen zum Verfassen der verlangten Zieltextsorte. Die Ergebnisse ergaben, dass Personen dann bessere Texte verfassten, wenn sie vor allem in der Phase vor dem Schreiben die zur Verfügung gestellten Informationen zur Aufgabenlösung länger betrachtet hatten. Taten sie dies beim Schreiben, hatte dies keinen positiven Effekt mehr in Bezug auf die Textqualität.
- *Je länger Personen mit Überwachungsaktivitäten befasst waren, desto mehr Inhalte integrierten sie in ihre Texte.* In ihrer Studie bildeten Raković, Fan et al. (2022) aus Eye-Tracking- und Navigationsverhaltensindikatoren einen Gesamtwert für das Monitoring. Darunter fielen sämtliche Aktivitäten erneuter Lektüren von Aufgabenstellungen, Zielvorgaben, bereits gelesenen Textteilen, eigenen Notizen und Unterstreichungen sowie das Überwachen der verstrichenen Gesamtarbeitszeit. Dieser Gesamtwert korrelierte stark mit einem Index des verfassten Textes, welcher vor allem auf der thematischen Abdeckung sowie der Ähnlichkeit zu den Bezugstexten basierte. In diesem Zusammenhang ist bemerkenswert, dass es für die Integrationsleistungen in den Texten *ungünstig* war, wenn Personen vor der eigentlichen Aufgabenbearbeitung längere Zeit die Aufgabenstellung und die Zielkriterien lasen. Erst im Verbund mit anderen Prozessen ergab sich damit ein positiver Effekt von metakognitiv gesteuerten Aktivitäten.

Die einzelnen Studien haben somit jeweils eigene Operationalisierungen gewählt, um das, was arbeitsdefinitorisch in Anlehnung an Tarricone (2011) als „metakognitive Bewusstheit“ bezeichnet wurde, zu operationalisieren. Alles in allem zeigt sich jedoch, dass eine höhere metakognitive Überwachung mit einer stärkeren Abdeckung von Inhalten und Dokumenten einhergeht.

Schreibprozesse wurden im Vergleich zur Domäne Lesen seltener untersucht, und innerhalb der Schreibprozesse sind wiederum ebenfalls Ungleichgewichte zu attestieren. Die Essenz der Befunde lautet:

- Günstig sind allgemein *metakognitive Steuerungen bzw. Bewusstheit* zum einen und *Planungsprozesse* zum anderen, denn hier ergibt sich bei aller Unterschiedlichkeit ein Muster: Je mehr planerischer Aufwand, primär in der Strukturierung der Inhalte, und Zielbezug rekonstruierbar waren, desto besser wurden in der Regel die Diskurssynthesen.
- Beim *Revidieren* sind derartige Zusammenhänge weniger deutlich erkennbar, es scheint zudem indirekte Wirkungen zu geben, welche über die zielbezogene Nutzung der Dokumente vermittelt werden.
- *Gesamtindizes*, welche für eine Vielzahl von verschiedenen Prozessen im Herstellen der Diskurssynthesen gebildet werden, sind insgesamt kaum untersucht. Sie hängen jedoch, wo sie dafür genutzt wurden, in varianzanalytischen Auswertungen interindividuelle Unterschiede bei Produktmaßen aufzuklären, positiv mit den Diskurssynthesen zusammen.

Damit unterstreichen die hier vorgestellten Resultate, dass spezifisch in ihrer Funktion zugeordnete Schreibprozesse empirisch bedeutend sind, vor allem jene in Relation zu metakognitiven Zielen und dem Suchen nach übergeordneten Sinnstrukturen der zu verarbeitenden Dokumenteninhalte.

5.2.4 Perspektive 3: Zur Rekursivität und zum Zusammenhang von (glückenden) Prozessen der Diskurssynthese

Die in den Teilkapiteln [5.2.2](#) und [5.2.3](#) eingenommenen Perspektiven waren rein lese- bzw. schreibspezifisch. Für das Verfassen von Diskurssynthesen müssen jedoch sowohl die rezeptive als auch die produktive Komponente zusammenwirken. Teils spiegelt sich das bereits in den Studien nieder, wenn beispielsweise ausgewertet wird, welche Lese- und Schreibaktivitäten miteinander korrelieren (Ackerman, 1991; Bråten, Brante & Strømsø, 2018; Du & List, 2020; Schoor et al., 2023) oder ob sie zusammengehörige Faktoren von Konstrukten abbilden (Liao et al., 2021). Andere Ansätze bestehen darin, mittels aggregativer Datenverfahren prototypische Muster zu extrahieren, welches kombinierte Auftreten und teils auch welche Sequenz von Aktivitäten, Prozessen und Strategien als latente Struktur in den Daten zugrunde liegt.

5.2.4.1 Gruppen von Personen mit ähnlichem Vorgehen – Ergebnisse aus Cluster- und Latente-Klassen-Analysen

Ein Ansatzpunkt in sogenannten personenzentrierten Ansätzen bei der Auswertung von Prozessdaten besteht darin, Gruppen von Individuen zu finden, die sich in ihren Vorgehensweisen, etwa in den Prozessen beim Verfassen von Diskurssynthesen, untereinander ähneln, sich aber als Gruppe von anderen Personen(gruppen) unterscheiden (Fryer & Shum, 2020). Mit solchen statistischen Verfahren lassen sich also Anzahl von Gruppen und Gruppengrößen bestimmen, zudem kann man die Gruppen in ihren Prozess- und Produktmerkmalen miteinander vergleichen. Solche Studien sind bislang noch recht selten vorgelegt worden (Escorcia et al., 2017; H. Kim, 2020; van Steendam et al., 2022). Bezogen auf Texttiefenmerkmale und die Zugehörigkeit zu Gruppen mit spezifischen Profilen im Vorgehen ergaben sich in der Mehrheit der Vergleiche allerdings kaum Auffälligkeiten (mit der Ausnahme zugunsten einer Gruppe stark inhaltlich revidierender Personen bei H. Kim, 2020). Gleichwohl lassen sich aus den bis dato wenigen Studien bei aller Unterschiedlichkeit der berücksichtigten Prozessdaten studienübergreifend Gruppen profilieren, die sich eher in Lese- und eher in Schreibprozessen manifestierten.

Personengruppen mit Auffälligkeiten im Lesen

Personengruppen lassen sich bezüglich des Lesens in drei Gruppen unterteilen, die in den Studien unterschiedlich häufig vorfindbar waren:

- *Personen, die extensiv lesen:* Hierunter fallen in der Untersuchung von H. Kim (2020) jene Personen, die bei einem offenen Arbeitsauftrag viele Rechercheaktivitäten absolvierten und dafür wenig planten und revidierten. Escorcia et al. (2017) fanden eine vergleichbare Gruppe, die viel Zeit mit Lesen und dem Anfertigen von Notizen verbrachte und sich dabei darauf konzentrierte, die Inhalte aus den gelesenen Dokumenten zu übernehmen. Van Steendam et al. (2022) konnten zudem eine Gruppe von Personen identifizieren, die über den gesamten Bearbeitungsprozess ausgiebig lasen und immer wieder die Dokumente aufsuchten.
- *Personen, die häufig zwischen Dokumenten wechseln und sie früh im Gesamtprozess konsultieren:* Diese Gruppe von Personen wechselte in der Studie van Steendams et al. (2022) im ersten Drittel häufig zwischen den Dokumenten, tat dies in der Mitte ebenfalls, wobei auch die Wechsel zwischen Diskurssynthese und Dokumenten in dieser Periode sehr häufig erfolgte. Dafür schrieben sie gegen Ende langsamer.

- *Personen, die erst spät lesen:* Van Steendam et al. (2022) nennen dieses Vorgehen „Konsultieren der Dokumente in letzter Minute“. Die Textproduktion dieser Personen startete früh, aber ohne starke Leseaktivitäten. Erst im letzten Drittel suchten diese Personen die Dokumente verstärkt auf.

Personengruppen mit Auffälligkeiten im Schreiben

Beim Schreiben gibt es wie beim Lesen mehrere Gruppen, nämlich drei, welche sich in verschiedenen konkreten Ausprägungen in den Studien extrahieren ließen:

- *Personen, die stark planen:* Personen, die ein solches Verhalten in der Studie von H. Kim (2020) zeigten, widmeten sich stärker dem Planen, welches einen erkennbaren Schwerpunkt in ihrem Gesamtbearbeitungsprozess ausmachte. Sie revidierten dafür weniger.
- *Personen, die intensiv revidieren:* Bei H. Kim (2020) waren dies Personen, die weniger planten, dafür aber ihre entstehenden Diskurssynthesen in einem besonders starken Maß inhaltlich modifizierten. Auch Escorcia et al. (2017) fanden eine Gruppe von Personen, die viel inhaltlich überarbeiteten und dabei häufig eigenen Text und zu verarbeitende Dokumente konsultierend lasen. In der Studie von H. Kim (2020) gab es noch eine weitere Gruppe von stark revidierenden Personen, doch hier bezogen sich die dominierenden Revisionen eher auf oberflächliche Modifikationen am Text.
- *Personen, die zügig schreiben:* Das Vorgehen dieser Personen, welche in der Studie von Escorcia et al. (2017) als „spontane Schreiberinnen und Schreiber“ bezeichnet wurden, war von wenig planerischen und revisorischen Aktivitäten gekennzeichnet. Van Steendam et al. (2022) konnten ein solches Profil ebenfalls feststellen: Hier verbrachten die Personen früh und relativ viel Zeit durchgängig mit dem Schreiben und lasen relativ wenig.

Diese jeweils drei Muster sowohl beim Lesen als auch beim Schreiben, deren markanteste profilgebenden Merkmale aufgezählt wurden, wobei selbstredend weitere Prozessdaten Bestandteile der Profile bilden, unterstreichen unterschiedliche Zugänge. Beim Lesen ist die Konstanz bzw. Massierung von Leseprozessen (bzw. mit Blick auf das Schreiben bei den zügig schreibenden Personen: der Mangel an Lesen) erkennbar. Beim Schreiben wirkt es so, als gebe es eher planende und eher revidierende Personen. Die Absenz deutlich über- oder unterlegener Produkte, die von diesen einzelnen Gruppen hergestellt wurden, verdeutlicht, dass es mehrere zielführende Vorgehensweisen gibt. Dies gilt freilich nur für ältere Jugendliche ab der zehnten Klassenstufe und Studierende, die in den Experimenten untersucht wurden.

5.2.4.2 Zusammenspiel von Lesen und Schreiben: musterhaftes Vorgehen und teils differenzielle Zusammenhänge

In einigen Studien wurden die Vorgehensweisen beim Herstellen von Diskurssynthesen als Muster bestimmt. Die zugrundeliegenden Fallzahlen waren in diesen qualitativen Zugängen aufgrund des hohen Kodier-, Analyse- und Auswertungsaufwands in der Regel klein. Aus den bislang noch wenigen Studien lassen sich zwei Muster abstrahieren, von denen das zweite Muster breiter abgestützt wirkt.

- *Muster 1: Vorgehen mit oder ohne Erstellen von planerischen Zwischenprodukten.* Kirkpatrick und Klein (2016) konnten mit den Daten von neun Jugendlichen zwei Vorgehensweisen extrahieren. In der einen Variante erstellten die Jugendlichen beim Schreiben über Online-Dokumente Zwischenprodukte wie Notizen und Gliederungsentwürfe, die sie im Rahmen ihrer Rechercheaktivitäten anfertigten und veränderten, um diese dann später für das Schreiben zu verwenden. Sie nutzten hierfür eigene Zwischenprodukte, die sowohl die inhaltlichen Elemente enthielten als auch solche, welche die Textstruktur bedienten. Die zweite Variante, ein Schreiben mit den just gelesenen Dokumenten, trennte Recherche/Planungsaktivitäten weniger stark von der Textproduktion. In dieser Studie wurden keine Zusammenhänge mit Textprodukten berichtet.
- *Muster 2: lineare und rekursive Vorgehensweisen.* Gleich mehrere Studien haben Prozessdaten aggregiert und entlang eines Kontinuums von rein linearen und diskreten Lese- und Schreibprozessen über teils rekursive bis hin zu hochdynamisch rekursiven Prozessen des stark interagierenden Lesens und Schreibens in zwei (Tarchi, Villalón et al., in Druck) bzw. drei Gruppen gereiht (Lenski & Johns, 1997; Mateos et al., 2008; Mateos & Solé, 2009; Solé et al., 2013). Diesen Studien zufolge lassen sich Personen danach unterscheiden, wie dynamisch sich die Prozesse bei ihnen gestalten. Diese Studien fanden als deskriptives Ergebnis vor, dass komplexere, rekursive Vorgehensweisen in stärker integrierte Darstellungen münden (Lenski & Johns, 1997; Mateos et al., 2008; Solé et al., 2013), teils kombiniert mit Indikatoren des besseren Verstehens des Sachverhalts (Mateos & Solé, 2009). Die einzige Studie, die gezielt Personen mit hoher bzw. geringer Rekursivität im Prozess statistisch verglich (Tarchi, Villalón et al., in Druck), ergab zugunsten der Personen mit höherer Rekursivität, dass ihre argumentativen Diskurssynthesen in zwei von drei Textqualitätsmaßen qualitativ höherwertig waren: Sie waren besser darin, das übergeordnete Thema zu benennen und Lösungen für die Thematik vorzuschlagen. Keine Differenzen gab es hingegen in der Integration von Argumenten und Gegenargumenten. Zusammengefasst: Je linearer die Lese- und Schreibprozesse verliefen, desto weniger integriert waren die Inhalte aus den gelesenen Dokumenten; je

komplexer und damit rekursiver die Lese- und Schreibprozesse abliefen, desto thematisch besser strukturiert und besser aufeinander bezogen waren die Inhalte in den Diskurssynthesen.

Prozessgruppen können also unterschieden werden, je nachdem welches Verhältnis separater Lese- und Schreibanteile mit planerischer Funktion vorliegen (Muster 1) bzw. in welcher strikten Trennbarkeit sie erfolgen (Muster 1 und 2). Dynamische und damit aus selbstregulatorischer Sicht (s. Kap. 6) aufwändigere Prozessverbünde scheinen sich in verschiedenen Produkten positiv niederzuschlagen (Muster 2), was allerdings aus statistischer Sicht noch nicht ausreichend gut abgesichert ist.

5.2.4.3 Zeitliche Veränderungen im Gesamtprozess

Das Zusammenspiel von Lese- und Schreibprozessen beim Verfassen von Diskurssynthesen ist nicht nur als Verbund verschiedener Prozesse in ihrer Interaktion beschreibbar, sondern – mit der Prozessperspektive direkt einhergehend – auch chronologisch veränderlich. Die zeitliche Perspektive auf die Abfolge bzw. Häufigkeit von Prozessen in Intervallen des Gesamtprozesses sind bislang noch randständig, aber Untersuchungen dazu kommen durchaus vor. Einige Studien teilten den Gesamtprozess in Viertel (McGinley, 1992) bzw. Drittel (Buyuktas Kara et al., 2018; Vandermeulen, van den Broek et al., 2020), andere betrachteten eine experimentell vom Schreiben abgetrennte Lese-Phase (Escorcía et al., 2017) und/oder Drittel der Schreibphase (Du & List, 2020; Escorcía et al., 2017). Einige konturieren sogar mithilfe der Kodierungen auf Einzelfallebene, wie individuell die Vorgehensweisen waren (Barzilai et al., 2023; Du & List, 2020), bzw. extrahierten typologisierend prototypische Vorgehensweisen (Kirkpatrick & Klein, 2016; Mateos & Solé, 2009; Solé et al., 2013, s. Teilkap. 5.2.4.2). Aus den Studien lassen sich mehrere Befundmuster abstrahieren:

- *Lese- und Schreibprozesse weisen zwar Affinitäten zu Zeitpunkten auf, sie kommen jedoch im gesamten Bearbeitungsprozess vor.* Die Studien gelangen zwar zum Befund, dass es Massierungen einzelner dem Lesen bzw. Schreiben zuordenbaren Aktivitäten gibt, gleichwohl finden Schreibprozesse bereits sehr früh und Leseprozesse noch sehr spät im Gesamtprozess statt (Buyuktas Kara et al., 2018; Du & List, 2020; Escorcía et al., 2017; McGinley, 1992; Vandermeulen, van den Broek et al., 2020). In der Regel, das zeigt sich in der Dauer von Prozessen, sind Leseprozesse deutlich zu Beginn der Aufgaben beobachtbar, gegen Ende überwiegen eher die Schreibprozesse (Barzilai et al., 2023; Escorcía et al., 2017). Es kommt also zu Häufungen bzw. im Falle des Lesens zu Funktionsverschiebungen, wie es der nächste Spiegelstrich verdeutlicht.

- *Gleichwohl sind Schwerpunkte erkennbar: Geht es zunächst um die lesende Aneignung von Informationen aus den Dokumenten, steht gegen Ende die Produktion mit einem stärkeren Fokus auf die Produktion und Optimierung des eigenen Texts im Vordergrund.* Besonders deutlich hat sich dies in den Studien von McGinley (1992), Escorcia et al. (2017), Buyuktas Kara et al. (2018), Vandermeulen, van den Broek et al. (2020) sowie Du und List (2020) gezeigt. In McGinleys (1992) Untersuchung mit dem Fokus auf den gesamten Prozess der Diskurssynthese-Herstellung überwogen das Lesen und Notieren in den ersten beiden Vierteln, während das Produzieren des Texts später dominierte. Zudem kam es zu stärkeren Fixierungen des Sinngehaltes im Laufe des Gesamtprozesses, d. h. der zunächst offene Prozess der Sinnkonstruktion wurde tendenziell enggeführt und zielte auf eine Konkretisierung ab. Ähnliches zeigte sich bei Escorcia et al. (2017): Das Lesen der fremden Dokumente nahm hier deutlich im Lauf der Aufgabenbearbeitung ab, wurde von der Textproduktion und danach von der Textrevision allmählich abgelöst, und das Vorgehen wurde zunehmend von metakognitiver Überwachung geprägt. Buyuktas Kara et al. (2018) nutzten Lautdenk-Protokolle und fanden Leseaktivitäten in den ersten beiden Dritteln und Schreibprozesse im mittleren und letzten Drittel gehäuft vor. Mittels Logfile-Daten ermittelten Vandermeulen, van den Broek et al. (2020) längere Lesezeiten und Wechsel zwischen verschiedenen Dokumenten im ersten Drittel der Aufgabenbearbeitung und stärkere Wechsel zwischen fremden Dokumenten und eigener Diskurssynthese in der Mitte des Prozesses. In der Studie von Du und List (2020) schließlich mit anderem Fokus rein auf Schreibstrategien ergab sich im ersten Drittel vor allem ein Organisieren der Struktur, und es folgte dann eher ein Auffüllen der Struktur mit gelesenen Inhalten bzw. eine Form des Auswählens von Inhalten aus den einzelnen Dokumenten. Zusammengefasst: Die Prozesse des lesenden Aneignens von Informationen sind rückläufig, je weiter der Gesamtprozess voranschreitet. Stattdessen werden – je nach geschaffener Strukturierung des eigenen Textes – dann Inhalte genutzt, um die Diskurssynthese zu erstellen und gegen Ende zu prüfen.
- *Es gibt Aufgabeneffekte, die sich teils in den Frequenzen einzelner Prozesse, noch deutlicher aber in den Zusammenhängen mit den Produkten niederschlagen.* Aufgabeneffekte beim Verfassen von Diskurssynthesen sind bei den Produktmaßen relativ stark erforscht worden (s. dazu Teilkap. [2.2.2.4](#)), bei den Prozessmaßen steht derlei noch aus. Momentan ist insbesondere die Experimentalstudie von Vandermeulen, van den Broek et al. (2020) anzuführen, in der mit Logfile-Daten das Verhalten beim Schreiben aufgezeichnet wurde. Die getesteten Schülerinnen und Schüler zehnter bis zwölfter Klassenstufen verfassten mehrere Diskurssynthesen, die sich zwei Zieltextsorten zuordnen lassen: argumentativen und informierenden Diskurssynthesen. In einer separaten Veröffentlichung (Vandermeulen, Maeyer et al., 2020) wurden elf

Prozessdaten aus fünf Gruppen – a) Zeitnutzung für Lese- und Schreibaktivitäten, b) Textproduktion, c) Pausen während der Textproduktion, d) Revisionsaktivitäten und e) Wechsel zwischen Dokumenten bzw. Dokumenten und eigenem Text – auf etwaige Aufgabeneffekte überprüft. In 12 der 27 Vergleiche gab es solche Differenzen. Sämtliche Differenzen traten in den Maßen zutage, welche einzelne Intervalle fokussierten. Gebündelt lässt sich sagen, dass argumentierende Personen früher mit dem Schreiben anfangen und dabei längere Zeit beim Schreiben pausierten. Bestand das Ziel darin, eine informierende Diskurssynthese zu verfassen, lasen die Personen länger und wechselten häufiger zwischen den gelesenen Dokumenten und der eigenen Diskurssynthese. Es wirkt demnach so, als sei das Lesen beim argumentativen Text eher selektiv und bereits deutlicher vom Schreiben geprägt, während es beim Informieren eher suchend-orientierend wirkt. Auch die Vorhersage der Textqualität war je nach Zieltextsorte abhängig von anderen Variablen (Vandermeulen, van den Broek et al., 2020): Informierende Diskurssynthesen wurden dann als besser beurteilt, wenn die Testpersonen beim Lesen anfänglich häufig zwischen den Dokumenten wechselten und gegen Ende weniger. Außerdem war es günstig, wenn im letzten Drittel die mit Lesen verbrachte Zeit in den Dokumenten eher moderat ausfiel. Bei den argumentativen Texten hingen hingegen anfänglich mittlere Aufenthaltszeiten in den Dokumenten und wenig mit den Dokumenten verbrachte Zeit im letzten Drittel positiv mit der Textqualität zusammen. Zusätzlich war ein häufigerer Wechsel zwischen Diskurssynthese und den zu verarbeitenden Dokumenten vorteilhaft. Diese differenziellen Zusammenhänge in der bisher von der Samplegröße umfangreichsten Studie weisen darauf hin, dass Aufgabenspezifika bestehen.

- *Nicht jeder einzelne Prozess in einem Gefüge von Prozessen ist günstig für die Textqualität – es gibt Effekte des Zeitpunkts seines Auftretens.* In der Schreibforschung gibt es Hinweise, dass es für eine höhere Textqualität von Belang ist, wann Schreibprozesse im Gesamtprozess aufgetaucht sind. Dies nehmen neuere Studien (Du & List, 2020; Vandermeulen, van den Broek et al., 2020) zum Teil auf und illustrieren, dass dieses generische Befundmuster sich auch auf Diskurssynthesen applizieren lässt. Die spezifischen Zusammenhänge divergieren jedoch je nach Studie, der Operationalisierung der Prozessmaße und den Textproduktmaßen. Dies zeigt das nachstehende Beispiel einer aktuellen Studie im Kasten „Illustrierendes Beispiel 21“.

Illustrierendes Beispiel 21

Die komplexen Zusammenhänge von Lese- und Schreibprozessen und Produkten beim Herstellen von Diskurssynthesen

A) Was war das Erkenntnisinteresse?

Die Studie hat sich der Frage verschrieben, welche Prozesse in separaten Lese- und Schreibphasen beim Herstellen von Diskurssynthesen junge Erwachsene bemühen, wie solche Prozesse untereinander zusammenhängen und ob sie für verschiedene Merkmale des Textprodukts prädiktiv sind (Du & List, 2020).

B) Wer wurde untersucht?

Die Teilnehmenden waren 32 US-amerikanische Studierende, mehrheitlich europäischer Herkunft (drei Viertel) und weiblich (73 Prozent). Diese Studierenden stammten aus verschiedenen Etappen des Studiums, die Fachrichtungen sind nicht berichtet worden. Ausgewertet werden konnten 28 Datensätze, die restlichen vier Datensätze waren unvollständig.

C) Welche Instrumente kamen zum Einsatz?

Die Studie erfasste außergewöhnlich viele Prozess- und Produktdaten. Unterschieden werden können bei den Prozessdaten jene, die aus der Lese- und der Schreibphase stammten, beide Phasen wurden gezielt experimentell voneinander getrennt. In der im Mittel 21-minütigen *Lese*phase erhielten die Testpersonen einen von zwei Aufträgen (Forschungsbericht oder Argumentation zum Thema Überbevölkerung verfassen), konnten in einer digitalen Bibliothek sechs Texte mit zugehörigen Metadaten konsultieren, wobei dies mit Logfile-Daten erfasst wurde, und Notizen anfertigen; parallel zum Lesen verbalisierten die Probandinnen und Probanden ihre Gedanken. Hieraus wurden zehn Prozessindikatoren extrahiert. Aus den Logdaten der digitalen Leseumgebung wurden a) Lesezeiten, b) Zahl aufgerufener Dokumente sowie c) die Häufigkeit des Zugriffs auf Metadaten extrahiert. Die Notizen wurden danach ausgewertet, d) wie viele Dokumente in den Notizen vorkamen, e) wie viele Dokumente mit Metadaten aufgeführt wurden, f) wie viele inhaltliche Aussagen die Notizen enthielten sowie g) wie stark Informationen bereits transformiert worden waren. Die Lautdenk-Protokolle wurden nach drei Arten von Lesestrategien analysiert: h) oberflächenorientierte Strategien der Informationsaufnahme, i) tiefenorientierte Strategien der Informationsbeurteilung und -elaboration sowie j) metakognitive Strategien. In der im Durchschnitt 18-minütigen *Schreib*phase, in der die Studierenden keinen Zugang mehr zu den Texten hatten, wurden ebenfalls Lautdenk-Protokolle und per Screen-Capturing Schreibprozessmaße erhoben, diese ergaben insgesamt fünf Maße. Mit dem Screen-Capturing wurden a) die Schreibzeit und b) die beim Schreiben durchgeführten Revisionen am entstehenden Text aufgezeichnet. Aus den Lautdenk-Protokollen generierten die Forscherinnen drei Kategorien von Strategien, nämlich c) die Herstellung eines Aufgabenprodukts primär mittels Planen und Revidieren, d) Verwendung von Informationen aus den und über die gelesenen Dokumente sowie e) metakognitive Strategien.

Die am Computer erstellten *Textprodukte*, also die Diskurssynthesen, wurden in Bezug auf drei Maße analysiert: a) eine holistisch eingeschätzte Textqualität, b) die Anzahl von Zitationen und c) die Anzahl von Diskurskonnektoren.

Diese Daten wurden mittels Korrelations- und Regressionsanalysen auf bi- und multivariate Art ausgewertet. Die Korrelationen fungierten vor allem zur Prüfung von etwaigen Zusammenhängen der 15 Prozess- und drei Produktmaße, die Regressionen dienten dazu, Unterschiede in den Produktmaßen auf Prozess- und weitere erhobene Variablen zurückzuführen.

D) Was waren die Hauptbefunde?

Die Studie lieferte reichhaltige Befunde, von denen an dieser Stelle nur ein kleiner Ausschnitt gebündelt dargestellt werden kann, und zwar verdichtet zu drei Punkten:

- *Lese- und Schreibprozesse waren untereinander mit wenigen Ausnahmen statistisch unverbunden.* Die Ausnahmen betrafen als unspezifisches Schreibprozessmaß die Schreibzeit und als spezifische Maße die metakognitiven Strategien beim Schreiben sowie die Sammelkategorie namens „Informationsverwendung“. Die Schreibzeit korrelierte mit allen Indikatoren des Lesens positiv, aber nur die Hälfte der Korrelationen war statistisch überzufällig, also signifikant. Die metakognitiven Strategien als spezifische Maße der Schreibprozesse hingen stark positiv mit Tiefenverarbeitungsstrategien beim Lesen (also Indikatoren der Informationstransformationen) und metakognitiven Strategien des Lesens zusammen. Eine letzte positive Korrelation bestand zwischen der ausgiebigeren Informationsverwendung von Informationen aus den gelesenen Texten beim Schreiben und der Anzahl von zuvor notierten Inhalten aus den Texten. Je mehr Informationen jemand in der Lesephase notiert hatte, desto stärker neigte die Person zur Verwendung von textuellen Informationen beim Schreiben.
- *Die Zusammenhänge von Produkt- und Prozessmaßen fallen in Korrelationen unterschiedlich aus.* Der auffälligste Befund bestand darin, dass von 45 potenziellen Korrelationen nur ein Drittel signifikant war – und dies waren mehrheitlich nicht jene Variablen, die sich in den Regressionsanalysen positive (oder negative) Prädiktoren erwiesen. Bei den Korrelationen waren die Lese- und Schreibzeiten als unspezifische Maße mit allen drei Produktmaßen positiv assoziiert: Je länger jemand las bzw. schrieb, desto hochwertiger wurden die Texte und enthielten überdies mehr Zitationen und Diskurskonnektoren. Ebenfalls auffällig war, dass vor allem die Leseprozesse, die mit lautem Denken erfasst wurden (oberflächen- bzw. tiefenorientierte Strategien und metakognitive Strategien), mit den drei Produktmaßen positiv zusammenhängen – wenn auch jede Kategorie bei Prozess bzw. Produkt auf unterschiedliche Weise.
- *Lese- und Schreibprozesse waren unterschiedlich stark prädiktiv für Textmaße in Regressionsanalysen.* Für die Anzahl der Zitationen gab es sowohl positive als auch negative Beiträge von Prozessmaßen. Mehr Zitationen entstanden bei längeren Lesezeiten und mehr planerischen und revisorischen Aktivitäten sowie metakognitiven Strategien beim Schreiben. Weniger Zitationen fanden sich hingegen, wenn die Personen stärker zu transformierenden Notizen beim Lesen und zur Inhaltsgeneration

beim Schreiben geneigt hatten. Für die *Diskurskonnektorenzahl* ergaben sich positive Beiträge nur bei spezifischen Schreibprozessen, nämlich bei verstärkter Informationsverwendung im mittleren Drittel der Schreibphase zum einen und mehr Planungs- und Revisionsprozessen im finalen Drittel des Schreibens zum anderen – bedeutend ist hier die Interaktion von Zeitpunkt und Prozessart. Für die *Qualität* der Diskussynthesen schließlich war nur eine einzige Schreibvariable rechnerisch bedeutend: Die Zahl an Planungs- und Revisionsstrategien ging positiv einher mit besseren Texten.

Alles in allem bestehen also komplexe bivariate Zusammenhänge zwischen Prozess- und Produktmaßen, und zwar sowohl in Bezug auf die Korrelationen zwischen Lese- und Schreibprozessen untereinander als auch in Bezug auf Prozess-Produkt-Verbindungen. Dies schlägt sich prinzipiell in den multivariaten Analysen nieder, für die nicht nur Leseprozesse teils rechnerisch unbedeutend waren (Diskurskonnektoren, Textqualität), sondern auch verschiedene Schreibprozesse (und Leseprozesse) zum einen teils gegenläufige Effekte hatten (Textqualität) und zum anderen der Zeitpunkt des Auftretens von Belang war (Diskurskonnektoren).

E) Diskussion

Die mit hohem methodischem Aufwand betriebene Studie erlaubt differenzierte Blicke auf Zusammenhänge von Produkt- und zuvorderst Prozessmaßen. Das Lesen zeichnete sich dadurch aus, dass die Studierenden vor allem Informationen oberflächenorientiert sammelten und diese später für das Schreiben nutzten. Diese Prozesse – zusammen mit den groben Kategorien der Gesamtlese- und Schreibzeit – hingen substantiell mit den drei Produktmaßen zusammen, welche sehr unterschiedliche Facetten der Multiple Documents Literacy abbildeten. Diese bivariaten Ergebnisse wurden von den multivariaten Analysen nur zum Teil bestätigt. In ihnen erwiesen sich nur die Schreibprozessindikatoren als vorhersagestark. Dies verdeutlicht die Bedeutsamkeit von verschiedenen Prozessen (im Verbund) mit spezifischen Erträgen für das Textprodukt, was für die weitere Forschung ein deutliches Signal und Gelegenheit für Vertiefungen darstellt.

Resümierend lässt sich anhand der noch wenigen Studien sagen, dass Lese- und Schreibprozesse dynamisch sind, aber es gewisse Musterhaftigkeiten gibt, die auf Aufgabenerfordernisse wie das Aneignen und Verarbeiten der Inhalte verweisen. Außerdem wirkt es so, als gebe es prototypisch günstigere und weniger günstige Vorgehensweisen bei Aufgaben und überdies spezifische Beiträge einzelner Prozesse für einzelne Produktmerkmale. All dies unterstreicht neben aller Dynamik des Verfassens der Diskurssynthesen emergierende überindividuelle Muster.

5.2.5 Zusammenfassung: ein abstrahiertes idealtypisches Vorgehen beim Schreiben von Diskurssynthesen

Die bislang zu Analyse- und Darstellungszwecken vorgenommene Separation von einzelnen Perspektiven auf einzelne Prozessgruppen hat dazu gedient, das Forschungsfeld zu systematisieren, was es mit sich gebracht hat, dass einzelne Prozesse aus gesamthaft erfassten Prozessverbünden isoliert dargestellt wurden. In dem Zugang, Prozesskategorien voneinander getrennt voneinander zu betrachten, wird der Blick auf größere Zusammenhänge erschwert. Darum wird abschließend an dieser Stelle empiriebasiert ein Merkmalskatalog günstiger Vorgehensweisen extrahiert, der aufgrund des aktuellen Kenntnisstands legitimierbar ist. Mit dieser Bündelung der Befunde geht einher, dass nicht sämtliche Studienresultate berücksichtigt werden, zumal jene, deren Grundlage fragmentarisch wirkt und/oder die aus wenigen Untersuchungen stammen.

Diskurssynthesen weisen dann eine über diverse Konstrukte und Merkmale beschreibbare qualitativ höherwertige Ausprägung auf, wenn

- *die Lese- und vor allem die Schreibzeit als Teil einer Gesamtbearbeitungszeit länger ausfällt.* Mit diesem unspezifischen Maß der reinen Dauer lassen sich zwar keine Aussagen über die Art und die Qualität der Prozesse treffen, doch speziell eine längere Schreibzeit ist relativ konsistent und vor allem positiv mit Produktmaßen assoziiert. Insbesondere die unspezifischen Leseprozessindikatoren wirken momentan als notwendige, aber noch nicht hinreichende Bedingung für erfolgreiche Diskurssynthesen.
- *mehr Inhalte schon beim Lesen ausreichend abgedeckt und teils bereits in Zwischenprodukten wie (grafischen) Notizen verarbeitet wurden.* Diskurssynthesen setzen Inhalte voraus, die zwar nicht alleinig, doch aber aus der informierten Lektüre der zur Verfügung gestellten Dokumente stammen. Dies nehmen Prozessindikatoren auf, welche das Konsultieren angemessen vieler Dokumente auszählen, also eine ausreichende Abdeckung des zu verarbeitenden Inhalts oder die Anzahl von verschiedenen inhaltlichen und Metadateneinheiten in Zwischenprodukten des lesebasierten Schreibens. Solche Maße lassen sich dem Auswählen zuordnen. Die empirischen Befunde hierzu gestalten sich disparat – analog zur Lesezeit als Prozessindikator: Nicht jede Studie fand (lineare) Zusammenhänge. Möglicherweise ist die quantitative Reichweite der Leseaktivitäten kein Garant dafür, dass umfassendere kognitive Leistungen tatsächlich erbracht werden, das zeigen besonders eindrücklich die sogar negativen Zusammenhänge mit der Textqualität, wenn Personen rein oberflächenorientiert vorgegangen waren.
- *die Personen erkennbare Integrationsleistungen erbringen, welche bereits für den Schreibanlass Restrukturierungs- und Transformationsleistungen*

beinhalten. Im Fall des Lesens ist damit gemeint, dass die Leserinnen und Leser intertextuelle Inferenzen herstellen und damit Makropropositionen und integrierte mentale Modelle konstruieren. Hier ergibt sich ein konsistentes studienübergreifendes Ergebnismuster. Für das Schreiben lässt sich Ähnliches festhalten. Vor allem Planungsaktivitäten, die gesamthaft erkennen lassen, dass Personen die gelesenen Inhalte organisierend strukturieren bzw. restrukturieren, führen zu besseren Textprodukten. Sowohl beim Lesen als auch beim Schreiben sind hierzu im Vergleich zu anderen Prozessen besonders viele und besonders in ihren Befunden übereinstimmende Studien durchgeführt worden, welche dokumentieren, dass sich rezeptive und produktive Prozesse mit Zielpunkt Makrostruktur als sowohl zentral als auch günstig für die Diskurssynthesen erweisen.

- *das Vorgehen rekursiver Natur ist*. Das bedeutet, dass das Vorgehen aus mehreren Teilprozessen besteht, die nicht linear absolviert werden, sondern sich insgesamt dynamisch gestalten und sich durch Wiederholungen, Rekombinationen und nicht durch distinkte und abgeschlossene Teile in strikter Sequenzialität auszeichnen.
- *Personen bezogen auf das Schreiben als einem Bestandteil insgesamt eher planen als revidieren*. Es besteht ein Ungleichgewicht von Studien in Hinblick auf zwei Hauptprozesse des Schreibens, denn planerische Aktivitäten, die sich unter anderem in Zwischenprodukten wie Notizen und grafisch organisierten Inhalten wie Maps niederschlagen, bilden den ungleich besser untersuchten Forschungsgegenstand. Bei den Revisionstätigkeiten sind direkte, indirekte, teils auch keine Zusammenhänge mit Textmaßen erkennbar. Momentan wirkt es demnach ungeklärt, inwieweit Revisionsprozesse in den Studien leistungssteigernde Wirkungen haben. Diese Uneindeutigkeit beim Revidieren gestaltet sich leicht anders, wenn man den engen Fokus auf Veränderungen des Textprodukts zugunsten einer breiter gefassten Betrachtung von metakognitiv gesteuerten Prozessen des Lesens und Schreibens verlässt. Denn nimmt man Studien hinzu und in den Blick, welche metakognitive Strategien und metakognitive Bewusstheit beinhalten, darunter solche zur Zielerreichung und dem Überwachen des Fortschritts – beides finden sich auch in der Detektion von textbezogenen Problemen und dem Optimierungsbedarf von Schreibprozessen als Teil des Revidierens –, dann zeigen sich durchaus positive Zusammenhänge. Jedoch wirkt die Rolle metakognitiver Prozesse gegenwärtig noch unterbelichtet.

Es lassen sich also fünf prozessuale Merkmale aus der derzeit vorliegenden Empirie ableiten, welche sich mehr oder minder deutlich als günstig für die Textprodukte (Diskurssynthesen) erwiesen haben. Diese fünf Merkmale sind nicht gänzlich trennscharf, aber sie helfen dabei, ein insgesamt eher produktives Vorgehen einzugrenzen.

5.3 Determinanten und interindividuelle Unterschiede

Komplexe Aufgaben wie das Herstellen einer Diskurssynthese als hybride, diverse Fähigkeiten erfordernde Lese- und Schreibanlässe lassen es wahrscheinlich wirken, dass Personengruppen sich darin unterscheiden bzw. dass individuelle Merkmale die Ausgestaltung der Prozesse und deren Ergebnisse, die Produkte, beeinflussen. Insgesamt fünf Bereiche beleuchtet dieses Teilkapitel. Es beginnt mit einem Fokus auf kognitive Leistungsgruppen (Teilkap. 5.3.1), die auf sehr verschiedene Weise gebildet wurden. Es folgen das inhaltliche Vorwissen (5.3.2) und die thematischen Überzeugungen zum Gegenstand, die hinsichtlich des empirisch breit nachweisbaren Phänomens „Text-Überzeugungskonsistenz-Effekt“ problematisiert werden (5.3.3). Im Anschluss werden die rezeptiven (5.3.4) und produktiven Leistungsfähigkeiten (5.3.5) fokussiert. Nicht überall in diesen fünf Bereichen sind Prozess- und Produktmaße gleichermaßen untersucht worden, die thematischen Überzeugungen (5.3.3), die Lese- (5.3.4) und die Schreibfähigkeiten (5.3.5) werden darum nur aus der Produktperspektive betrachtet.

Da Diskurssynthesen in der Forschung häufig als Maß für verschiedene, kognitiv anspruchsvollere Verstehensleistungen im Sinne „expressiver Tests“ eingesetzt werden (Primor & Katzir, 2018), besteht eine gewisse Schnittmenge zu den Darstellungen im Kapitel 3 und 4. In den dort behandelten Inhalten ging es um die auf Verstehensprozesse abzielenden Einschätzungs- bzw. Integrationsleistungen, die nicht notwendigerweise in schriftliche Texte, eben die Diskurssynthesen, münden müssen. Gleichzeitig sind bestimmte individuelle Merkmale mit Verstehensprozessen verbunden, auf die hybride Lese- und Schreibleistungen beim Erstellen von Diskurssynthesen angewiesen sind. Diese Hybridität bringt die schon erwähnte Schnittmengenproblematik aus darstellerischer Sicht mit sich. Um die Redundanz mit den Teilkapiteln 3.3 und 4.3 gering zu halten, werden beispielsweise die epistemischen Überzeugungen gezielt ausgeklammert.

5.3.1 Leistungsunterschiede

Leistungsunterschiede in verschiedenen Leistungsgruppen sind in den Studien bisher auf unterschiedliche Weise gebildet und in Bezug auf diverse Diskurssynthesemaße analysiert worden. Folgende zwei grundsätzliche Ansätze für die Gruppenbildung sind gewählt worden:

- *Differenzierung nach erbrachter Leistung*: Ausschlaggebend ist für die Gruppenbildung die Nutzung einer Ausgangsleistung, sei es im *Vorwissenstests* (Gil et al., 2010a; Le Bigot & Rouet, 2007), sei es in standardisierten *Lesetests* (Kennedy, 1985; Primor et al., 2021, Experimente 1 und 2; Spivey, 1992; Spivey & King, 1989). Zudem wurden ex post Gruppen anhand der Produktleistungen

in der *Textqualität* gebildet und die Prozessindikatoren miteinander deskriptiv verglichen (Du & List, 2020; Plakans, 2009; Yang, 2014).

- *Differenzierung nach erwartbarer Kompetenz*: Teils wurde auch über die Nähe des Textthemas zum Studiengang der Testpersonen angenommen, diese hätten bei Kongruenz von Thema und Studienfach ein hohes oder geringes Wissen, wodurch *Gruppenkontraste von Studierenden mit unterschiedlichen Fächern* legitimiert wurden (z. B. Ackerman, 1991; Rouet et al., 1997). Teils wurden Grenzen gezogen zwischen *Bildungsetappen bei Studierenden* über die Zugehörigkeit zu Erststudiumsangehörigen (Undergraduates) und Personen, die einen weiteren Bildungsabschluss anstreben (Graduates) (List & Oaxaca, 2024). Ein anderer, einer ähnlichen Logik folgender Personengruppenvergleich besteht in der Gegenüberstellung von *Altersgruppen* (McCarthy et al., 2022; Spivey & King, 1989).

5.3.1.1 Prozesse

Die verschiedenen Gruppen wurden sowohl in Prozessindikatoren als auch in Produktmaßen miteinander verglichen. Die Prozesse der Personen unterscheiden sich zwar nicht durchgängig, doch gibt es Auffälligkeiten zugunsten der leistungstärkeren Personengruppen. Leistungstärkere Personen, die eine Diskurssynthese verfassen,

- gehen eher *selektiv und zielorientiert* vor und demonstrieren ein eher auf globales Verstehen fokussierendes Vorgehen (Plakans, 2009),
- verbringen mehr *Zeit mit dem Lesen* der zu verarbeitenden Dokumente (Du & List, 2020) bzw. insgesamt mit der Aufgabe (Spivey & King, 1989) – allerdings berichten mit einem Fokus auf Vorwissen Le Bigot und Rouet (2007) das Gegenteil, weil Personen mit geringerem Wissen länger lesen –,
- nutzen mehr *Informationsverwendungsstrategien*, setzen also beispielsweise Notizen strategischer für ihren späteren Text ein (Du & List, 2020; Kennedy, 1985) bzw. erstellten elaboriertere Pläne für ihre Diskurssynthese (Spivey & King, 1989),
- *beurteilen* im Prozess der Lektüre von Dokumenten Informationen häufiger (Ackerman, 1991; Kennedy, 1985),
- verwenden mehr *planerische und Organisationsstrategien* im Schreibprozess (Du & List, 2020; Kennedy, 1985),
- geben mehr *Formulierungsaktivitäten* zu Protokoll (Du & List, 2020) und
- verwenden gemäß Selbstauskünften sämtliche *Lese- und Schreibstrategien* häufiger als leistungsschwächere Personen (Yang, 2014).

5.3.1.2 Produkte

Die Produktmaße erweisen sich ebenfalls teils als sensitiv für Leistungsgruppen. Leistungsstärkere Personen

- *reichern* Diskurssynthesen mit *mehr domänenspezifischem Wissen an* (Rouet et al., 1997),
- *übernehmen mehr korrekte Inhalte* aus den gelesenen Texten (Perin et al., 2003; Spivey, 1992; Spivey & King, 1989),
- verfassen *besser strukturierte Diskurssynthesen* (Spivey, 1992; Spivey & King, 1989),
- *transformieren Inhalte* aus Texten inhaltlich stärker (Gil et al., 2010a, Experimente 1 und 2),
- wechseln stärker zwischen den gelesenen Dokumenten in ihrer Diskurssynthese, ihre Texte weisen also mehr *Switches* auf (Gil et al., 2010a, Experiment 2),
- konnten mehr *Generalisierungen* auf der Basis des Gelesenen bilden (Primor et al., 2021, Experiment 1),
- enthalten sich bei kontroversen Themen einer explizit *einseitigen Positionierung* (Rouet et al., 1997) und
- erzielen eine höhere holistisch eingeschätzte *Textqualität* (McCarthy et al., 2022; Spivey, 1992; Spivey & King, 1989).

5.3.1.3 Abschlussbemerkung

Neben den eher eindeutig wirkenden Befundmustern in Prozessen und Produkten sind solche Ergebnisse Teil der empirischen Wahrheit, die auf Ähnlichkeit in den Gruppen verweisen. Denn nicht in allen angestellten Gruppenvergleichen waren statistisch überzufällige Merkmalsausprägungen in Produktindikatoren vorfindbar (Ackerman, 1991; Le Bigot & Rouet, 2007; List & Oaxaca, 2024; Primor et al., 2021, Experiment 1, v. a. aber Experiment 2; Rouet et al., 1997). Ähnliches gilt für die nur sehr spärlich untersuchten Prozessmaße (Le Bigot & Rouet, 2007).

Diese Divergenz der Ergebnisse mag methodischer Natur sein. Denn die in diesem Teilkapitel vorgestellten Studien unterstreichen allein bei der Gruppenbildung von Leistungsgruppen, dass eine Einheitlichkeit in den unabhängigen Variablen nicht gegeben ist. Dies allein erschwert schon die Interpretation von Ergebnissen. Erschwerend kommt hinzu, dass die Heterogenität sich bei den abhängigen Variablen fortsetzt. Gleichwohl sind nicht nur relativierend die Einschränkungen zu konstatieren. Selbst wenn der Kenntnisstand noch fragmentarisch ist, so gibt es in den bisher publizierten Studien wichtige Hinweise für die weitere Forschungsarbeit:

- Auffällig ist beispielsweise eine *Prozess-Produkt-Divergenz*: So fand Spivey (1992) nur Vorsprünge in den Produkt-, aber nicht in den Prozessindikatoren; sie replizierte damit bei Studierenden einen Befund, den sie bei Jugendlichen bereits tendenziell vorgefunden hatte (Spivey & King, 1989). Damit stellt sich die Frage, ob Leistungsdifferenzen sich tatsächlich konsistent über diverse Indikatoren der Diskurssynthese und des Prozesses dorthin manifestieren.
- Die Frage nach der *Eigenständigkeit von mehreren Faktoren der Differenz* stellt sich, ist aber noch unbeantwortet: Hervorzuheben ist die Studie von Spivey und King (1989), denn in ihr wurden nicht nur Altersgruppen, sondern auch Leseverstehensleistungen als Indikatoren für Leistungsgruppen erfasst und – was in diesem Zusammenhang besonders bedeutend ist – ihre einzelnen bzw. kombinierten Effekte statistisch bestimmt. Dabei stellte sich heraus, dass die Leseleistungen die wichtigeren und Unterschiede zwischen den Gruppen besser vorhersagenden Merkmale waren. Gil et al. (2010a, Experiment 2) kamen zusätzlich zum Ergebnis, dass auch die Aufgabe als solches mit Vorwissensgruppenzugehörigkeit interagiert: So war ein hohes Vorwissen teils nur bei argumentativen Diskurssynthesen von Vorteil. Solche Ergebnisse sind deshalb von hoher Bedeutung, weil sie dabei helfen, die relative Wichtigkeit von verschiedenen Einflussfaktoren realistisch abzuschätzen.
- Die *Gleichartigkeit oder Besonderheit des Zusammenwirkens von diversen Prozessen innerhalb von Leistungsgruppen* verdient Beachtung: Mindestens ebenso wichtig sind Studien, die prüfen, ob gleiche Zusammenhänge zwischen Strategien und Leistungsindikatoren für die Leistungsgruppen gelten. In Strukturgleichungsanalysen ermittelte Yang (2014), dass ein oberflächenorientiertes Übernehmen von Inhalten aus den Dokumenten nur bei leistungsschwächeren Personen die Textqualität positiv vorhersagt, aber nicht unter leistungstärkeren Peers. Differenzielle Effekte und Moderatoranalysen sind jedoch noch ein Desiderat.

Alles in allem sind konsistente Muster durch die Vielzahl der unabhängigen und abhängigen Variablen in den Studien und in einem eher unterbelichteten Forschungsfeld schwer auszumachen. Gleichwohl wirkt es aktuell so, als würden leistungstärkere Personen insgesamt strategischer vorgehen und in ihren Produkten mehr Transformationen des Gelesenen erreichen.

5.3.2 Thematisches Vorwissen

Das thematische Vorwissen wird aus Sicht der Theorie als interne Ressource für das Anfertigen von Diskurssynthesen betrachtet (Britt et al., 2018; List & Alexander, 2019; Rouet & Britt, 2011). Dem liegt die Annahme zugrunde, dass das Vorwissen dabei hilft, elaborative Inferenzen zwischen verschiedenen Aussagen,

vor allem den intertextuell vorliegenden Informationen, zu bilden. Diese Fähigkeit ist bedeutend, weil das Verhältnis von Aussagen durch die Wissensbestände geklärt werden kann und dadurch intertextuelle Kohärenz hergestellt wird. Solche Inferenzen benötigen Personen dazu, die verschiedenen und verschiedenartigen Informationen aus Texten und über Texte hinweg kohärent zu organisieren, z. B. in Form der komplexen „Dokumentenmodelle“ (s. Teilkap. [4.1.2.3](#), Rouet & Britt, 2022). In ihrem Rahmenmodell, das seinerseits eine Integration diverser Modellierungen darstellt, beschreiben List und Alexander (2019), dass – je nach Leseziel und damit einem internalen individuellen Benchmark sowie diverse interne und externe Ressourcen nutzend – Inferenzen variierender Abstraktionsleistung nötig sind, um ein Aufgabenprodukt wie die Diskurssynthese herzustellen.

Das Vorwissen wird ausgehend von seiner veranschlagten hohen Bedeutung in einer Vielzahl von Studien wahlweise als Kovariate oder als Prädiktor einbezogen. Die Operationalisierung variiert beträchtlich, so wird das Vorwissen in Studien teils über die selbsteingeschätzten Wissensbestände erfasst (z. B. bei Guo, 2022a, Kiili, Brante et al., 2020, Tarchi & Villalón, 2022, und van Steendam et al., 2022), was Validitätsprobleme mit sich bringt, da nicht im Ansatz geklärt ist, in welcher Granularitätsgröße, Komplexität und überprüfbar vorhandenem Wissen tatsächlich Gedächtnisinhalte existieren (McCarthy & McNamara, 2021). Darum werden in diesem Teilkapitel nur Studien berücksichtigt, die mittels Vorwissenstest mit mehreren offenen oder geschlossenen Multiple-Choice-Fragen das Wissen der Personen erfasst haben, welches in der Domäne der jeweiligen Experimentalstudie auch Gegenstand der Dokumentensets war.

5.3.2.1 Produkte

Beginnend mit den Produkten lässt sich zunächst konstatieren, dass Vorwissen und Produktmaße zwar durchaus oft in Studien zueinander statistisch in Bezug gesetzt wurden. Generell ist die Befundlage zu den Vorwissenseffekten allerdings gemischt. Denn die Zusammenhänge zwischen dem thematischen Vorwissen und Merkmalen der Diskurssynthesen sind nicht immer eindeutig. Es gibt demnach in der Forschung Befunde sowohl für positive Beiträge bzw. Korrelationen des Wissens mit Eigenheiten der Diskurssynthese als auch Nullergebnisse. Dies wird aus der systematischen Gegenüberstellung von Studien deutlich, die in Tabelle 29 vorgenommen wurde. Vertikal folgen die Produktindikatoren jener Reihenfolge, die im Teilkapitel [5.2.1.2](#) vorgestellt wurden. In der Tabelle 29 kommen teils andere Studien als im Teilkapitel [5.2.2–5.2.4](#) vor, die teils nur auf Produktmaße fokussierten.

Tabelle 29: Überblick über (Null-)Effekte des Vorwissens bei verschiedenen Produktmaßen der Diskurssynthese

Indikator der Diskurssynthese	Positive statistische Effekte	Statistische Nulleffekte
Referenzierungen	• Anzahl richtig gelöster MC-Fragen (Kiili et al., 2021) • Anzahl korrekter Begriffsdefinitionen (Du & List, 2020; List & Du, 2021) • Korrekte Antworten auf offene Fragen (List, Du & Wang, 2019)	• Anzahl richtig gelöster MC-Fragen (Bråten, Brante & Strømsø, 2018; Florit, Cain & Mason, 2020; Salmerón, Gil & Bråten, 2018a) • Anzahl korrekter Begriffsdefinitionen (List & Alexander, 2020; List et al., 2017) • Korrekte Antworten auf offene Fragen (Bråten, Ferguson et al., 2014; Kessler et al., 2021; List, Oaxaca et al., 2021; Mason et al., 2017; Mason, Scrimin, Zaccoletti et al., 2018)
Switches	–	Anzahl richtig gelöster MC-Fragen (Bråten, Brante & Strømsø, 2018)
Diskurskonnektoren	–	Anzahl korrekter Begriffsdefinitionen (Du & List, 2020)
Abdeckungsgrad von Inhalten	• Anzahl richtig gelöster MC-Fragen (Castells et al., 2023) • Korrekte Antworten auf offene Fragen (Braasch, Bråten, Strømsø & Anmarkrud, 2014; Du & List, 2022, Studie 2; List, Du & Wang, 2019, Thema 2)	• Anzahl richtig gelöster MC-Fragen (Barzilai & Eshet-Alkalai, 2015; Bråten, Brante & Strømsø, 2018; Casado-Ledesma & Tarchi, 2024; Florit, Cain & Mason, 2020; Peterson & Alexander, 2020) • Korrekte Antworten auf offene Fragen (Kessler et al., 2021; List, Du & Wang, 2019, Thema 1)
Lokale Kombinationen	–	• Anzahl richtig gelöster MC-Fragen (Florit, Cain & Mason, 2020; Peterson & Alexander, 2020; Salmerón, Gil & Bråten, 2018a) • Anzahl korrekter Begriffsdefinitionen (Du & List, 2021; List & Alexander, 2020) • Korrekte Antworten auf offene Fragen (List, Oaxaca et al., 2021)
Textlänge	• Anzahl richtig gelöster MC-Fragen (Casado-Ledesma & Tarchi, 2024; McNamara et al., 2024; Tarchi, in Druck; Tarchi & Mason, 2020) • Anzahl korrekter Begriffsdefinitionen (Anzahl von Begründungen: List et al., 2017)	• Anzahl richtig gelöster MC-Fragen (Bråten, Brante & Strømsø, 2018; Kammerer et al., 2021; Tarchi, 2021b; Tarchi, Casado-Ledesma et al., in Druck) • Anzahl korrekter Begriffsdefinitionen (Anzahl von Argumenten: List et al., 2017; Anzahl Sätze: List & Du, 2021)

Indikator der Diskurs-synthese	Positive statistische Effekte	Statistische Nulleffekte
Textstrukturmaße	• Anzahl richtig gelöster MC-Fragen (Bräten, Anmarkrud et al., 2014; Bräten, Ferguson, Anmarkrud & Strømsø, 2013; Castells et al., 2023; Florit, Carli et al., 2020, Thema 1; Haverkamp et al., 2024; List & Oaxaca, 2024) • Korrekte Antworten auf offene Fragen (Iordanou et al., 2020; Mason, Scrimin, Zaccoletti et al., 2018; Mason et al., 2020)	• Anzahl richtig gelöster MC-Fragen (Andresen et al., 2019; Barzilai & Eshet-Alkalai, 2015; Casado-Ledesma & Tarchi, 2024; Florit, Cain & Mason, 2020; Florit, Carli et al., 2020, Thema 2; Strømsø et al., 2016; Tarchi, 2021a, 2021b, in Druck; Tarchi, Casado-Ledesma et al., in Druck; Tarchi & Mason, 2020) • Korrekte Antworten auf offene Fragen (Barzilai & Ka'adan, 2017; Kobayashi, 2009; List, Du & Wang, 2019; Mason et al., 2017)
Textqualität	• Anzahl richtig gelöster MC-Fragen (McCarthy et al., 2022; McNamara et al., 2024; Raković, Fan et al., 2022) • Anzahl korrekter Begriffsdefinitionen (Du & List, 2020; Lee & List, 2023b; List, Stephens & Alexander, 2019; Ying Wang & List, 2019)	• Anzahl richtig gelöster MC-Fragen (van der Graaf et al., 2022; Vongkulluksn et al., 2023) • Anzahl korrekter Begriffsdefinitionen (List & Alexander, 2020; List & Du, 2021) • Korrekte Antworten auf offene Fragen (Bräten, Ferguson et al., 2014; McCrudden et al., 2024)

(Legende: MC = Multiple Choice)

Die mehr als 30 Studien, die in Tabelle 29 kompiliert sind, wurden danach differenziert, ob sie das Vorwissen mittels geschlossener Multiple-Choice-Fragen, mit offenen Begriffsdefinitionsfragen oder mit offenen Fragen operationalisierten. Hintergrund der Differenzierung in der Darstellung ist, dass sich beim Sourcing (s. Teikap. 3.3.2) in deskriptiver Sicht bereits Häufungen ergeben haben (Anmarkrud et al., 2022). Im Forschungsüberblick von Anmarkrud et al. (2022) wurden positive Zusammenhänge vor allem in Studien vorgefunden, die mittels Multiple-Choice-Fragen das Vorwissen maßen. Dieses Muster zeichnet sich in den Daten aus Tabelle 29 so nicht ab. Dafür sind andere Muster auffällig:

- *Keine Vorteile höheren Vorwissens:* Vernachlässigbar wirkende Effekte des Vorwissens lassen sich bei den *Referenzierungen* attestieren: Die Ausprägung des Vorwissens führte in der Mehrheit der Studien zu weder ausgeprägterem noch weniger stattgefundenem Belegen von Quellen in den Diskurssynthesen. Vorwissen hat zudem – erwartungswidrig mit Blick auf die Ergebnisse aus dem Teilkapitel 4.3.1.1 – auch nicht mit einem Mehr an inferenzbasierten *lokalen Kombinationen* in den Diskurssynthesen zu tun.
- *Empirisch unklare Befunde wegen widersprüchlicher Ergebnisse:* Auffällig ist, dass mehrere Produktmaße in mehr oder minder starkem Mischverhältnis der Befunde studienübergreifend divergierende Ergebnisse haben: Der *Abdeckungsgrad von Inhalten* sowie die *Textlänge*, *-strukturmaße* und

-*qualität* zeigten in einzelnen Studien positive Zusammenhänge mit den Vorwissensmaßen, teils auch in regressionsanalytischen, also multivariaten Auswertungen. Teils bestanden keine Korrelationen.

- *Empirisch wenig klare Befunde wegen Forschungslücken:* Forschungsdesiderate bestehen bei den Switches und der Anzahl der Diskurskonnektoren. Hier lässt sich momentan anhand der kaum vorhandenen Empirie am wenigsten etwas über Vorwissenseffekte aussagen.

Es zeigt sich also, dass Vorwissenseffekte differenziert nach Produktmaßen beurteilt werden sollten, selbst wenn die Befundlage gemischt ist. Auch das ist ein Ergebnis: Aktuell lässt sich wenig über den Status des Vorwissens als allgemeine Ressource bzw. allgemeingültige Determinante für Diskursynthesen als Produkte substantzieren, jedenfalls wenn es um die direkten, linearen Zusammenhänge geht. Das lässt sich durchaus theoretisch darüber erklären, dass Vorwissen als Ressource in aktuellen Modellen benannt wird (Britt et al., 2018; List & Alexander, 2019; Rouet & Britt, 2011; van den Broek & Kendeou, 2022), dass es aber zugleich als eher distaler Faktor zu behandeln ist. Das heißt: Im Zusammenspiel von weiteren in- und externalen Faktoren und der Ausgestaltung von Lese- und Schreibprozessen sind vermittelnde Wirkungen anzunehmen. Dies impliziert, dass das Zusammenspiel von Vorwissen und Prozessen beispielsweise stärker empirisch ausgeleuchtet werden sollte.

5.3.2.2 Prozesse

Ein generelles Problem der Forschung zum Herstellen von Diskurssynthesen in Bezug auf Vorwissenseffekte bei Prozessen besteht in zweierlei. Es existiert im Vergleich zu produktbezogenen Studien ein größeres Desiderat im Allgemeinen. Daneben wird – im Besonderen – mitunter das Vorwissen zwar teils erhoben, aber dann nicht in Beziehung zu Prozessen gesetzt (Anmarkrud et al., 2014; Du & List, 2020; Strømsø & Bråten, 2014; Strømsø et al., 2013). Die Befundlage ist gemischt:

- *Keine bzw. keine positiven empirischen Zusammenhänge:* Mehrere Studien konnten keine Zusammenhänge zwischen Vorwissen und Prozessmaßen feststellen, sei bei Lese- (Bråten, Brante & Strømsø, 2018; Peterson & Alexander, 2020) und Schreibzeiten (Bråten, Brante & Strømsø, 2018), sei es bei Notizen (Castells et al., 2023; Hagen et al., 2014), bei rekursiven Lesevorgängen (Bråten, Ferguson, Anmarkrud & Strømsø, 2013), bei epistemischen Kognitionen in Lautdenk-Protokollen (Bråten, Ferguson et al., 2014) und bei Eye-Tracking-Daten (Mason et al., 2020).
- *Positive empirische Zusammenhänge:* Teils ließen sich positive Korrelationen nachweisen, zum Beispiel von höherem Vorwissen und einem stärker an der Textoberfläche orientiertem Vorgehen mit stärkerem Hervorheben

oder Relektüren (van der Graaf et al., 2022) bzw. der stärkeren Zuwendung zu digitalen Dokumenten (Peterson & Alexander, 2020).

- *Komplexere Zusammenhänge von Wissen, Prozessen und Produkten:* Den Zusammenhang von Vorwissen, Prozessen und Produkten haben bislang nur wenige Untersuchungen zum Gegenstand erkoren. Die Befundlage ist hierbei uneindeutig. Zusammenhänge haben in ihrer Untersuchung Bräten, Anmarkrud et al. (2014) demonstrieren können. Sie erfassten mittels eines Multiple-Choice-Tests unter den teilnehmenden Studierenden das Wissen zum Thema des Dokumentensets als Vorwissensmaß, den retrospektiven Einsatz von intertextuellen Integrationsstrategien als Prozessindikator sowie die ausgewogene Darstellung des kontroversen Sachverhalts mit mehreren offenen Fragen als Produktvariable. Unter statistischer Kontrolle weiterer personenbezogener Merkmale ergab sich in Strukturgleichungsanalysen ein direkter und ein indirekter positiver Effekt des Vorwissens. Direkt war der Effekt, weil es unmittelbare Zusammenhänge zwischen mehr inhaltlichem Wissen und ausgewogenerer Darstellung der Kontroverse gab. Indirekt und zusätzlich war ein Zusatzeffekt: Je mehr Wissen die Personen hatten, desto mehr intertextuelle Strategien gaben sie an, für intertextuelle Inferenzbemühungen verwendet zu haben, und dieses Vorgehen korrespondierte mit ausgewogeneren Diskurssynthesen. Während Bräten, Anmarkrud et al. (2014) bei einem Textstrukturmaß medierende Effekte der Prozesse auf das Produkt vorfanden, stellten van der Graaf et al. (2022) derlei für die Textqualität nicht fest. In ihrer Studie ergaben sich keinerlei Mediationseffekte von verschiedenen Prozessgruppen.

Ähnlich wie bei der Produktperspektive stellt sich die Forschungslage zum Vorwissen bei den Prozessindikatoren disparat dar. Grundsätzlich ist der Bereich deutlich weniger erforscht, sodass die Datenbasis als eher dünn bezeichnet werden muss. In den bisherigen Studien sind enge Beziehungen zwischen Prozessindikatoren und Vorwissen bisher nicht nachgewiesen bzw. fokussiert worden. Darum bedarf es mehr Forschung und zudem genauerer Explikationen, auf welche Weise sich das Vorwissen auf die Prozesse auswirken soll.

5.3.2.3 Abschlussbemerkung

Inhaltliches Vorwissen über den Gegenstand, zu dem Personen eigene Diskurssynthesen verfassen, ist im Vergleich von produkt- und prozessbezogenen Studien häufiger mit Blick auf Produktmaße untersucht worden. Weder bei den Prozessen noch bei den Produkten stellt sich die Rolle des Vorwissens als sehr eindeutig dar, was im Falle der Produktmaße dann auffällt, wenn genauer differenziert wird, welche Zusammenhänge mit den Produktindikatoren untersucht wurden. Bei den am meisten fokussierten Maßen scheint eine paradoxe Beobachtung zu

bestehen: Je mehr Studien vorliegen, desto eher ist die empirisch beobachtbare Tendenz, dass sich die Befunde widersprechen. Speziell Struktur- und Qualitätsindikatoren bei Diskurssynthesen sind von diesem Muster betroffen, und zwar durchaus innerhalb der Studienkorpora, die aus den identischen Forschungsgruppen stammen. Das Phänomen solcher unklar wirkenden Auftretensmuster von (mangelnden) Zusammenhängen wirft die Frage nach dem Zustandekommen auf. Auffälligkeiten in der Art der Erfassung des Vorwissens mit geschlossenen oder offenen Frageformaten sind bezüglich der Art der Ergebnisse nicht direkt erkennbar. Prozessmaße, die dabei helfen könnten, die Genese der (Null-)Ergebnisse zu erklären, sind zum einen noch recht wenig mit Vorwissensmaßen gemeinsam analysiert worden. Zum anderen sind selbst dort die Ergebnisse alles andere als eindeutig, wo derartige Analysen vorgenommen wurden. Damit schließt sich der Kreis in Bezug auf die bis heute wenig klaren Befundmuster. Es ist demnach ein großer Forschungsbedarf anzumelden, um die bisherigen Ergebnisse zu plausibilisieren und den Kenntnisstand weiter auszubauen.

5.3.3 Thematische Überzeugungen (Produktperspektive)

Thematische Überzeugungen werden im Zusammenhang mit dem automatischen Validieren von neuen Informationen beim Lesen und den Folgen für das Leseverstehen aktuell diskutiert (O'Brien & Cook, 2016a; Richter & Maier, 2017; Wolfe & Griffin, 2017). Überzeugungen lassen sich als Positionen oder Haltungen definieren, die Individuen gegenüber Propositionen oder Aussagen hinsichtlich des Wahrheitsgehalts haben (Bråten & Strømsø, 2020). Die Grundannahme ist, dass Leserinnen und Leser Informationen nicht neutral verarbeiten, sondern diese anhand verschiedener innerer Vergleichsmaßstäbe – darunter Vorwissen und Überzeugungen – auf ihre Plausibilität hin automatisch überprüfen. Sind diese Inhalte konsistent mit den eigenen Überzeugungen, ist die Wahrscheinlichkeit höher, dass sie in die mentalen Repräsentationen integriert und tiefer verarbeitet werden. Widersprechen die Informationen aus Texten hingegen den Überzeugungen, erfolgt keine automatische Validierung und die Informationen werden eher unterdrückt. Das Ergebnis ist dann – ohne Korrektur – eine einseitig verzerrte, den eigenen Überzeugungen entsprechende Repräsentation des Sachverhaltes. Bei kontroversen Themen mit verschiedenen Perspektiven und widersprüchlichen Aussagen besteht ergo die Gefahr, dass unkorrigierte Überzeugungen der Komplexität des Themas nicht gerecht werden. Richter und Maier (2017) nennen dieses Phänomen den „*Text-Überzeugungskonsistenz-Effekt*“, den sie auf breiter empirischer Basis in ihrem Forschungsüberblick durchgängig vorfanden, darunter in Studien, in denen eine Diskurssynthese zu verfassen war. Ein neueres Beispiel, das aufgrund seiner Aktualität nicht bei Richter und Maier (2017) berücksichtigt werden konnte, ist im Kasten „Illustrierendes Beispiel 22“ enthalten.

Illustrierendes Beispiel 22

Berichte mir deine Einstellung, und ich sage dir, welche Ausrichtung deine Diskurssynthese haben wird – systematische Bias-Effekte bei Jugendlichen

A) Was war das Erkenntnisinteresse?

Spielen thematische Überzeugungen eine Rolle, wenn Jugendliche divergierende Texte lesen und einen Standpunkt in einer Diskurssynthese dazu einnehmen sollen? Dies untersuchte eine amerikanische Experimentalstudie (Braasch et al., 2022).

B) Wer wurde untersucht?

Das Sample der Studie bestand aus 44, im Mittel 12 Jahre alten Jugendlichen, davon waren 60 Prozent weiblich, 43 Prozent waren afrikanische, 46 Prozent europäische Amerikanerinnen und Amerikaner. Die Jugendlichen besuchten eine Schule im mittleren Süden der USA, die Herkunft der Probandinnen und Probanden kann aufgrund der Einkommenssituation der Familien in der Schule als sozioökonomisch schwach bezeichnet werden.

C) Welche Instrumente kamen zum Einsatz?

Die Studie erfasste eine Woche vor der eigentlichen Aufgabe mit der Diskurssynthese mit einem kurzen Fragebogen die individuellen Überzeugungen zum Vegetarismus. Diese mit einigem Vorlauf erfassten Überzeugungen bildeten in diversen Regressionsanalysen den Prädiktor für vier Produktindikatoren und ein prozessnahe Maß. Eine Woche nach dem Absolvieren des Überzeugungsfragebogens schrieben die Jugendlichen eine argumentative Diskurssynthese, bei der sie eine Position einnehmen und begründen sollten. Sechs Dokumente wurden eigens für das Experiment erstellt, diese behandelten drei Subthemen zum Vegetarismus aus einer Pro- und Contra-Perspektive und waren hinsichtlich der intertextuellen Bezüge und Aussagen ausbalanciert. Die Themen korrespondierten mit Aussagen aus der Skala mit den thematischen Überzeugungen. Die Texte dienten als Grundlage dafür, eine eigene Diskurssynthese zu verfassen, in der sich die Jugendlichen auf der Basis des Gelesenen begründet für oder gegen Vegetarismus entscheiden und nach Möglichkeit Belege aus den Dokumenten explizit einfügen sollten. In den Diskurssynthesen wurde neben der allgemeinen Haltung pro, contra oder neutral gegenüber Vegetarismus noch gezählt, wie viele der Aussagen aus den Dokumenten übernommen wurden und wie viele Referenzierungen die Diskurssynthesen enthielten. Als ein Prozessindikator wurde ein kurzer Fragebogen direkt nach dem Lesen, aber vor dem Schreiben der Diskurssynthesen vorgelegt, welcher das Konstrukt metakognitive Bewusstheit erfasste. Der Fragebogen beinhaltete Aussagen zum Wissen über Vorgehensweisen und zur Regulation der Leseprozesse beim Lesen der sechs Texte.

D) Was waren die Hauptbefunde?

Der Effekt der Einstellungen schlug sich in allen Analysen nieder. Wer positive Einstellungen zu Vegetarismus hegte,

- hatte mit höherer Wahrscheinlichkeit eine Diskurssynthese verfasst, welche pro Vegetarismus argumentierte (statt eine neutrale Position einzunehmen oder eine nicht-vegetarische Ernährung als erstrebenswert zu postulieren),
- dessen Diskurssynthese enthielt mehr zur eigenen Einstellung passende Aussagen aus den gelesenen Dokumenten,
- beinhaltete zudem weniger einstellungsinkompatible Aussagen aus den Dokumenten,
- referenzierte weniger – betrieb also weniger Sourcing – und
- gab mehr metakognitive Bewusstheit beim Lesen an.

E) Diskussion

Die Studie illustriert, dass sich die Einstellungen als durchgängiger Prädiktor vor allem für die Produktmaße niederschlugen. Je stärker positiv die Einstellungen waren, desto affirmativer gestalteten sich die Diskurssynthesen in dieser Hinsicht. Aufgenommen wurden hierbei vor allem unterstützende Aussagen, und die gegenläufigen Inhalte, welche aus Sicht der Theorie auch mehr Quellenangaben im Sinne von Referenzierungen erwartbar gemacht hätten, waren weniger stark in den Diskurssynthesen aufzufinden. Die prozess-näheren Maße der metakognitiven Bewusstheit waren ebenfalls stärker ausgeprägt, wenn die Jugendlichen positivere Überzeugungen zum Vegetarismus aufwiesen. Dieser letzte Befund wirkt jedoch insofern implausibel, als eine gewisse Einseitigkeit im Sinne automatisch ablaufender Validierungsprozesse aus den Produkten spricht. Möglicherweise ist hier also ein Problem in der adäquaten Einschätzung der eigenen metakognitiven Aktivitäten ursächlich.

Aus den bisherigen Studien, darunter auch solche, in denen Lesemaße erhoben wurden, lassen verschiedene Befundmuster zu den Diskurssynthesen primär zu Produkt-, teils auch zu Prozessmaßen extrahieren, wobei die Prozessmaße in den Studien zu teils schwer interpretierbaren Ergebnissen führten (s. z. B. Kobayashi, 2014). Textuelle Grundlage dieser Studien waren mindestens zwei Dokumente, die zu einen Sachverhalt gegensätzliche Positionen vertraten. Wer auf dieser textuellen Basis die jeweilige Aufgabe bearbeitete und zuvor erfasste (starke) thematische Überzeugungen hatte, verfasste Diskurssynthesen

- mit *klarerer inhaltlicher Ausrichtung zugunsten der eigenen Position* (Braasch et al., 2022; Karimi & Mousavi-Moghadam, 2022; Kessler et al., 2019; Kobayashi, 2014; Mateos et al., 2011; van Strien et al., 2014),
- mit *mehr gelesenen Inhalten, die der eigenen Überzeugung entsprachen* (Braasch et al., 2022; Karimi & Mousavi-Moghadam, 2022; Maier et al., 2018; van Strien et al., 2016),

- mit *weniger Inhalten, welche der eigenen Überzeugung widersprechen* (Braasch et al., 2022; Kobayashi, 2009, Experimente 1 und 2; Maier et al., 2018),
- mit *mehr ergänzten Inhalten*, welche nicht aus den gelesenen Dokumenten stammten, dafür aber teils unter geringerer Abdeckung von Inhalten aus den Dokumenten (van Strien et al., 2014),
- mit der *Übernahme von mehr inhaltlich korrekten Aussagen aus den gelesenen Dokumenten* bzw. weniger Übernahmen von falschen Aussagen (Kessler et al., 2021),
- die *in ihrem Umfang (= der Anzahl der Wörter) geringer ausfielen* (Tarchi, 2021b; Tarchi et al., 2022), was deshalb von Belang ist, als der Umfang in einigen Studien Prädiktor der argumentativen Ausgewogenheit war (Casado-Ledesma & Tarchi, 2024; Tarchi & Mason, 2020; Tarchi & Villalón, 2021) resp. substanziell mit ihr korrelierte (Tarchi, in Druck; Tarchi, Casado-Ledesma et al., in Druck; Tarchi & Villalón, 2022) – in diesen letztgenannten Studien ließen sich jedoch keine Effekte der Einstellungen auf die Produktmaße eruieren,
- mit *weniger stark zweiseitig argumentierenden Inhalten* (Tarchi, 2021b),
- mit *weniger kohärenten Argumentationen* (Mason et al., 2023),
- mit *weniger Referenzierungen* (Braasch et al., 2022), wobei dieser Befund je nach Überzeugungen zu divergieren scheint: Positive Einstellungen führen zu mehr, negative hingegen eher zu weniger Referenzierungen (Kessler et al., 2021).

Die Ergebnisse unterstreichen, dass Personen selektiv zugunsten ihrer eigenen Überzeugungen gelesene Dokumente nutzen und eigene Gedächtnisinhalte ergänzten. Dies spiegelt rein lesebezogene Ergebnisse wider, nach denen Verzerrungen vor allem auf höheren Verstehensebenen bestehen (Abendroth & Richter, 2020; Karimi & Richter, 2021a; 2021b).

Ein Teil der Studien mit Diskurssynthesen widmete sich zudem der Frage, ob *Einstellungen und Beurteilung von einzelnen Aussagen bzw. Dokumenten zusammenhängen*. Das ist deshalb von Belang, da das Auswählen von Inhalten einen zentralen Bestandteil des Verfassens von Diskurssynthesen bildet (s. Teilkap. 5.1.2.1). Studien, die nach Plausibilitätseinschätzungen fragten, zeigten ebenfalls Muster. Personen schätzten je nach ihrer eigenen Überzeugung

- die gelesenen Inhalte als plausibler ein, wenn diese ihrer eigenen Überzeugung entsprachen (Karimi & Mousavi-Moghadam, 2022; Kobayashi, 2009, Experiment 1; Kobayashi, 2014; van Strien et al., 2016), sowie
- zudem teils objektiv besser ein (Kessler et al., 2021).

Der „Text-Überzeugungskonsistenz-Effekt“ ist demnach ein gut belegtes Phänomen, das kognitiv effektive, automatisch ablaufende Validierungsprozesse erlaubt (Singer, 2019), allerdings zum Preis der geringeren Ausgewogenheit bei

kontroversen Themen. Die Folge sind einseitig verzerrte, unausgewogene Repräsentationen bzw. Textprodukte in Form der Diskurssynthesen, die nur einen Teil des Sachverhalts abbilden.

5.3.4 Lesefähigkeiten (Produktperspektive)

Wer eine Diskurssynthese herstellen soll, ist darauf angewiesen, die zu verarbeitenden Dokumente inhaltlich verstanden zu haben. Solche Fähigkeiten, die spezifisch bei intertextuellen Integrationsprozessen bereits andernorts Gegenstand waren (s. Teilkap. 4.2.1.1), werden an dieser Stelle nicht fokussiert. Vielmehr stehen generische Fähigkeiten im Vordergrund dieses Teilkapitels. Da Lesefähigkeiten ein sehr breites Spektrum abdecken, welches vom Wort- und Satzverstehen über längere Texteinheiten, ganze Texte und multiple Dokumente reicht (Perfetti, 1997), und da sich die Forschung zu Diskurssynthesen als heterogen darstellt, ist es nötig, bei den Fähigkeiten zu differenzieren. In Anlehnung an das Vorgehen von Anmarkrud et al. (2022) werden zwei Gruppen unterschieden: basale Lesefähigkeiten (Leseflüssigkeit sowie Verstehensleistungen auf Satzebene) und Leseverstehen. Die Tabelle 30 differenziert entsprechend diese Fähigkeiten und nutzt als Ordnungsprinzip die produktbezogenen Indikatoren, welche in Teilkapitel 5.2.1.2 etabliert wurden.

Tabelle 30: Überblick über (Null-)Effekte der Lesefähigkeiten bei verschiedenen Produktmaßen der Diskurssynthese

Indikator der Diskurssynthese	Positive statistische Effekte	Statistische Nulleffekte
Referenzierungen	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 1; Kiili, Bråten et al., 2020; Kiili et al., 2021) • Leseverstehen (Kiili, Bråten et al., 2020; Lee et al., in Druck)	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 2) • Leseverstehen (Mason et al., 2017; Mason, Scrimin, Zaccoletti et al., 2018)
Switches	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 1) • Leseverstehen: –	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 2) • Leseverstehen: –
Diskurskonnectoren	• Basale Lesefähigkeiten: – • Leseverstehen: –	• Basale Lesefähigkeiten: – • Leseverstehen: –
Abdeckungsgrad von Inhalten	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 1; Kiili, Bråten et al., 2020) • Leseverstehen (Kiili, Bråten et al., 2020; Schoor et al., 2020)	• Basale Lesefähigkeiten (Braasch, Bråten, Strømsø & Anmarkrud, 2014; Bråten, Brante & Strømsø, 2018, Thema 2) • Leseverstehen: –

Indikator der Diskurs-synthese	Positive statistische Effekte	Statistische Nulleffekte
Lokale Kombinationen	• Basale Lesefähigkeiten (Latini et al., 2019; Latini et al., 2021) • Leseverstehen (García et al., 2021; Hattan et al., 2023; Lee et al., in Druck)	• Basale Lesefähigkeiten: – • Leseverstehen: –
Textlänge	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 1; Latini et al., 2019) • Leseverstehen (MacArthur et al., 2023; Mason, Tarchi et al., 2022)	• Basale Lesefähigkeiten (Bråten, Brante & Strømsø, 2018, Thema 2) • Leseverstehen: –
Textstrukturmaße	• Basale Lesefähigkeiten (Andresen et al., 2019; Bråten, Ferguson, Anmarkrud & Strømsø, 2013; Florit, Cain & Mason, 2020, Thema 2; Florit, Carli et al., 2020, Thema 2; Haverkamp et al., 2024; Racanello et al., 2022) • Leseverstehen (Florit, Cain & Mason, 2020, Themen 1 und 2; Florit, Carli et al., 2020, Thema 2; Kullberg et al., 2023; Mason, Scrimin, Zaccoletti et al., 2018; Mason, Tarchi et al., 2022; Mason et al., 2020)	• Basale Lesefähigkeiten (Florit, Cain & Mason, 2020, Thema 1; Florit, Carli et al., 2020, Thema 1) • Leseverstehen (Florit, Carli et al., 2020, Thema 1; Mason et al., 2017)
Textqualität	• Basale Lesefähigkeiten: – • Leseverstehen (Cheong et al., 2019; Cheong et al., 2018; Cheong et al., 2022; Collins et al., 2021; MacArthur et al., 2023; McCarthy et al., 2022; Payant et al., 2019; Risemberg, 1996; Shin & Ewert, 2015; Trites & McGroarty, 2005; Weston-Sementelli et al., 2018; Zhu et al., 2021)	• Basale Lesefähigkeiten: – • Leseverstehen: –

(Legende: – = keine Studien berichtet)

Über 30 Studien bilden den Grundstock der Darstellung in Tabelle 30. Muster lassen sich primär in den beiden Produktmaßen im unteren Ende der Tabelle ausmachen: Wer Texte gut versteht, ist dazu in der Lage, besser strukturierte, d.h. darstellerisch ausgewogenere Diskurssynthesen sowie qualitativ höherwertige Texte zu schreiben. Ein zweites Muster besteht darin, dass generell wesentlich mehr Studien auf positive Zusammenhänge der Lesefähigkeiten und Diskurssynthesemaße verweisen als auf keinerlei solcher Zusammenhänge.

Nicht nur korrelative Zusammenhänge zwischen lesebezogenen Leistungen und den Diskurssynthesemaßen kommen in der Empirie zum Tragen. Mitunter erweisen sich die Zusammenhänge als weniger eindeutig, sobald mehr Variablen berücksichtigt werden. In Regressionsanalysen waren beispielsweise teils basale Lesefähigkeiten prädiktiv, während etwa das Leseverstehen keine zusätzlichen Varianzanteile aufklärte (Florit, Carli et al., 2020). Auch das inverse Befundmuster fand sich, bezeichnenderweise in einer Arbeit der gleichen Forschungsgruppe,

mit den gleichen Instrumenten (Florit, Cain & Mason, 2020). Andere Studien, in denen verschiedene Arten von Lesetests zum Einsatz kamen, ergaben ebenfalls nicht immer eindeutige Befunde, welche Lesefähigkeiten für die Diskurssynthesemaße prädiktiv waren (Kullberg et al., 2023; Liao et al., 2021; McCarthy et al., 2022). Hinzu kommt, dass je nach Berücksichtigung von anderen Variablen (basale) Lesefähigkeiten keine zusätzliche Varianz aufklären (Bräten, Brante & Strømsø, 2018). Dies spricht dafür, verschiedene Facetten der allgemeinen Lesefähigkeiten und ihr Zusammenspiel mit anderen Variablen genauer in ihrem Ertrag abzuschätzen und zu untersuchen.

Ein eher unterbelichteter Forschungszweig befasst sich mit der Frage, *wie nah oder fern die Lesefähigkeiten in Bezug auf das Herstellen der Diskurssynthesen erfasst werden*. Die Lesefähigkeiten werden im Fall des Leseverstehens nur selten proximal erfasst, d. h. mittels der Texte, welche die Basis der Aufgabe bei den Diskurssynthesen bildeten. Erste Befunde verdeutlichen die Eminenz dieser proximal gemessenen Leistungen: Je besser die Personen die jeweiligen Texte verstanden hatten (intratextuelles Verstehen), über die sie später schrieben, desto mehr waren Personen dazu in der Lage, die kontroverse Thematik ausgewogen darzustellen (Kobayashi, 2009; List & Oaxaca, 2024), mehr intertextuelle Inferenzen zu bilden (Griffin et al., in Druck; Hattan et al., 2023; Lee et al., in Druck), mehr Inhalte in die Texte zu übernehmen (Griffin et al., 2012, in Druck), mehr Referenzierungen im eigenen Text vorzunehmen (Lee et al., in Druck) sowie qualitativ hochwertigere Diskurssynthesen zu schreiben (Payant et al., 2019). Zudem erwies sich das intertextuelle Leseverstehen als Prädiktor der Textqualität (Cheong et al., 2019) bzw. korrelierte mit der Anzahl übernommener zentraler Konzepte (McCrudden et al., 2024; Schoor et al., 2020). Hierin deutet sich an, dass das Leseverstehen einerseits allgemein (mit distalen Instrumenten, die nichts mit der später zu lösenden hybriden Diskurssynthese-Aufgabe zu tun haben) und andererseits proximal (mit Verstehensleistungen auf der Basis der zu verarbeitenden Dokumente) erfasst werden kann. Hinzu kommt, dass beide Arten von Leseverstehensleistungen mehrheitlich mit Indikatoren der Textqualität korrespondieren. Dieser Befund ist für das Leseverstehen allgemein und die Textqualität sehr gut empirisch belegt (Ahmed & Wagner, 2020; Kent & Wanzek, 2016; Kim et al., 2024).

Ein letztes Wort verdienen *Prozessmaße*. Prozessdaten sind bislang nur wenig erforscht worden, sie betreffen als unspezifische Leseprozessmaße zum Beispiel wechselnde Fixationen von Informationseinheiten bei Eye-Tracking-Studien (Latini et al., 2021), Lesezeit (Bräten, Brante & Strømsø, 2018) sowie Schreibzeit (Bräten, Brante & Strømsø, 2018). Die Zusammenhänge sind allerdings studienübergreifend nicht eindeutig; sie wären es wert, zukünftig in der Grundlagenforschung vertieft untersucht zu werden.

Zusammengefasst: Basale Lesefähigkeiten, vor allem aber das Leseverstehen längerer Texte wirken im Licht der empirischen Befunde überwiegend als

Ressource speziell für jene Leistungen bei der Diskurssynthese, die auf der Gesamtstruktur des Textes bzw. dessen umfassender Qualität basieren. Darum ist ein gutes Leseverstehen, zumindest bei bivariaten Analysen, mit besseren Textleistungen assoziiert. In komplexeren Analysen bleibt dieser Zusammenhang nicht immer erhalten, zudem ist über die Mediation der Leseleistung durch spezifische Lese- und Schreibprozesse noch zu wenig bekannt.

5.3.5 Schreibfähigkeiten (Produktperspektive)

Das Herstellen von Diskurssynthesen als genuin hybrider Lese- und Schreibfähigkeiten erforderndes Vorgehen (Spivey, 1990) setzt sowohl die im Teilkapitel [5.3.4](#) behandelten Lesefähigkeiten voraus als auch jene im Schreiben. Dies ist in der Forschung indirekt anerkannt, indem das für die Schreibforschung dominante Konstrukt Textqualität auch besonders häufig in Studien mit Lese- und Schreibprozessen betrachtet wurde (s. Teilkap. [5.2.1.2](#)). Dass Schreibfähigkeiten bezogen auf Textqualitätsmerkmale bei einzelnen Aufgaben aber als eigenständiger Prädiktor in Regressions- (Collins et al., 2021; McCarthy et al., 2022; McNamara et al., 2024) sowie Strukturgleichungsanalysen (Cheong et al., 2022; Liao et al., 2021), als weitere Variable zur Klärung von Korrelationen von Schreibleistungen und Diskurssynthese-Merkmalen (Shin & Ewert, 2015) und schließlich als Kovariate in Interventionsstudien (Weston-Sementelli et al., 2018) berücksichtigt wurde, ist bisher noch selten. Die aufgeführten sieben Studien wurden mit älteren Personen (Personen aus der höheren Sekundarstufe bis hin aus Universitäten) durchgeführt, in den Experimenten wurden argumentative Diskurssynthesen verlangt, und die Studien erfassten allesamt die Leseleistungen der Testpersonen.

Das Ergebnis aller Studien ging in die gleiche Richtung: *Je besser die Schreibleistungen jenseits der verlangten Diskurssynthesen waren, desto besser war auch die über mehrdimensionale Maße erfasste Textqualität der Diskurssynthesen.* Von besonderem Belang ist in diesem Zusammenhang die prozessbezogene Mediationsanalyse von Liao et al. (2021): Bessere Schreibleistungen waren einerseits direkte Prädiktoren der Diskurssynthesenqualität. Andererseits wurden sie auch über den berichteten Einsatz von einem Verbund von verschiedenen Strategien vermittelt: Wer besser schrieb, wendete mehr Strategien an, welche wiederum bessere Diskurssynthesen nach sich zogen. Dieser singuläre Befund illustriert die Ressource Schreibfähigkeit besonders und deutet auf komplexe Zusammenhänge hin. Insgesamt ist die Datenbasis also zwar noch dünn, wirkt bei all ihrer Vorläufigkeit jedoch konsistent.

5.4 Probleme

Hybride Lese- und Schreibprozesse wie im Falle des Herstellens von Diskurssynthesen weisen eine Vielzahl schwierigkeitsgenerierender Merkmale auf, und entsprechend sind Schwierigkeiten aus der empirischen Forschung seit geraumer Zeit belegt (Cumming et al., 2016; Philipp, 2020a). In diesem Teilkapitel werden einige der Schwierigkeiten gebündelt dargestellt, als Strukturgeber fungieren zwei der drei Transformationsprozesse Auswählen, Organisieren und Verbinden, die im Teilkapitel [5.1.2.1](#) dargestellt wurden. Gleichwohl sind die Produktperspektiven in diesem Teilkapitel leitend, denn Prozessmaße sind in den Studien zum einen noch zu heterogen erfasst, zum anderen lassen sich derzeit kaum sinnvolle Benchmarks formulieren, ab wann das Fehlen von Prozessen bzw. das Unterschreiten von Vergleichswerten als kritisch gelten (s. z. B. van Steendam et al., 2022).

Produktbezogene Indikatoren wurden in vielen Studien berichtet, sie werden hier zu zwei Bereichen aggregiert. Erstens schöpfen Personen das Potenzial der Inhalte aus Dokumenten nicht aus, was meint, dass zentrale Inhalte nicht ausgewählt werden resp. nicht in den Diskurssynthesen rekonstruierbar sind ([5.4.1](#)). Dieser erste Problembereich betrifft das Auswählen. Zweitens gibt es darüber hinausgehende Defizite, welche das Integrieren und damit das Organisieren und Restrukturieren der gelesenen Inhalte im eigenen Text betreffen ([5.4.2](#)). Dieser Problembereich versammelt unter sich verschiedene Phänomene mangelnder Organisations- und Syntheseleistungen, von denen drei fokussiert werden: additive Listen von Inhalten, unausgewogen einseitige Darstellungen kontroverser Sachverhalte und zu guter Letzt Diskurssynthesen, in denen neben mangelnder Organisation der Inhalte weitere Auffälligkeiten bestanden und die sich zu Gruppenprofilen verdichten ließen – die hohen Anteile ungenügender Diskurssynthesen markieren mithin einen Kristallisationspunkt multipler Probleme, den viele Schreiberinnen und Schreiber offenkundig nicht zu lösen imstande sind.

5.4.1 Problemfeld 1: geringe Abdeckung von Inhalten in Diskurssynthesen (Probleme im Auswählen)

Der Abdeckungsgrad von Inhalten in Diskurssynthesen bildet einen relativ gut untersuchten Bereich der Forschung (s. Teilkap. [5.2.1.2](#)). Die Grundlogik vieler Studien, in denen die Testpersonen Diskurssynthesen verfassten, besteht darin, dass die Informationen, welche den Sachverhalt charakterisieren, über mehrere Dokumente verteilt sind. Daraus ergibt sich die Anforderung an die lesende und schreibende Person, die Inhalte aus multiplen Dokumenten ausreichend im eigenen Text abzudecken. Diese Abdeckungsgrade lassen sich verschiedentlich

operationalisieren, wobei eine Bedingung ist, dass geklärt ist, wie viele der Informationseinheiten als essenziell für den Aufgabenerfolg gelten, um dadurch eine Quote zu bestimmen.

Die Informationseinheiten zur Ermittlung von Abdeckungsquoten können hierbei variieren. Es gibt erstens Studien, in denen aufwändig sämtliche in den Texten vorkommenden Inhalte in sogenannte *Inhaltseinheiten* („Content Units“; Spivey, 1991) bzw. in Propositionen segmentiert werden, und es wird dann in den Diskurssynthesen ausgezählt, wie viele dieser Inhaltseinheiten/Propositionen vorgekommen sind. Deutlich selektiver sind Ansätze, in denen *Kernkonzepte* (Blaum et al., 2017; Braasch et al., 2013; Graesser et al., 2007, Experiment 2; Griffin et al., in Druck; Wiley et al., 2020), Argumente pro bzw. contra einer strittigen Thematik (Maier et al., 2018; Marttunen & Kiili, 2022) oder die Abdeckung von Teilthemen (Lyu et al., 2023) bestimmt wurden, deren Vorhandensein in den Diskurssynthesen ausgewertet wird. Dieser zweite Ansatz fokussiert damit deutlich auf Makropropositionen. Eine Zwischenposition im Spektrum von Auszählung aller potenziell verwendbaren Propositionen und spezifisch für die Aufgabebearbeitung erforderlicher Makropropositionen nimmt ein dritter Ansatz ein: eine differenzierte Auswertung der Übernahmen von Inhalten auf verschiedenen Ebenen der Makrostruktur der einzelnen Texte und bei Fokussierung auf spezifische Übernahmen aus einzelnen Texten (Rickert & Philipp, 2022). Diese drei Kategorien lassen sich bei aller Differenz in der fokalen Einheit gleichwohl einer übergeordneten Auswertbarkeit zuordnen: Immer geht es um die *Quotienten von inhaltlichen Übernahmen*.

Zu guter Letzt sind für die Belange dieses Teilkapitels noch als vierte Kategorie jene Studien anzuführen, welche konkrete Inhalte der Texte vernachlässigen und sich stattdessen auf die *Abdeckung der Dokumente* in vorgegebenen Dokumentensets als Bezugseinheit stützen (List & Du, 2021). Dies lässt sich zum Beispiel darüber legitimieren, dass das Leseverstehen multipler Dokumente nicht nur auf propositionale Abdeckungsgrade als Vorstufe für integrierte mentale Modelle angewiesen ist, sondern dezidiert auch Dokumente als Repräsentationseinheiten (im Intertextmodell des Dokumentenmodells etwa, s. Teilkap. [4.1.2.3](#)) vorgesehen sind.

- Zunächst zu den *Abdeckungsgraden von Inhalten*: Auffällig ist, dass eine Vielzahl von sowohl Interventions- als auch Grundlagenforschungsstudien berichtet, dass die in ihnen untersuchten (nicht eigens geförderten) Personen maximal ein (gutes) Drittel jener Informationseinheiten in ihren Diskurssynthesen aufgreifen, die aus der Sicht des jeweiligen Forschungsteams verwertbar bzw. verwertungsnotwendig wirkten (Braasch, Bräten, Strømsø & Anmarkrud, 2014; Braasch et al., 2013, 2022; Britt & Sommer, 2004, Experiment 2; Casado-Ledesma et al., 2021; Daher & Kiewra, 2016; Graesser et al., 2007, Experiment 2; Griffin et al., 2012, in Druck; Hattan et al., 2023;

Kammerer & Gerjets, 2014b, Studien 1 und 2; Linderholm et al., 2016; Luna et al., 2020; Luo & Kiewra, 2019, Experiment 1; Maier et al., 2018; Marttunen & Kiili, 2022; Mateos et al., 2018; Mateos et al., 2020; Merkt et al., 2017; Raković et al., 2023; Spivey, 1991; Voss & Wiley, 1997; Wiley et al., 2009, Experiment 1; Wiley et al., 2020; Wiley & Voss, 1996). Einige Untersuchungen geben höhere Quoten von rund der Hälfte an (Boscolo et al., 2007; Linderholm et al., 2014, Studie 2; Marttunen & Kiili, 2022; Reynolds & Perin, 2009); Studien, in denen Personen deutlich über die Hälfte zentraler Hauptinhalte in ihren Diskurssynthesen abdeckten, sind selten, kommen aber vereinzelt vor (Blaum et al., 2017; Lyu et al., 2023; Primor et al., 2021, Experiment 1). Diese allgemeinen Befunde werden von ersten Studienergebnissen ergänzt, in denen die Textschwierigkeit berücksichtigt wird: Je besser verständlich ein Text aus dem Dokumentenset ist, desto eher (und desto korrekter zudem) übernehmen Personen Inhalte (Rickert & Philipp, 2022). Außerdem scheinen sich makrostrukturelle Unterschiede dahingehend zu ergeben, ob unterstützende Details bzw. Hauptargumente übernommen werden (Marttunen & Kiili, 2022).

Weniger häufig ist das Gegenstück zu den verstehensnotwendigen Informationen bzw. Informationseinheiten empirisch untersucht worden. *Fehlerhafte oder irrelevante Informationen* sind jedoch deshalb von Belang, weil es Hinweise darauf gibt, dass sie zulasten korrekter bzw. relevanter Informationen gehen (Kessler et al., 2021; Wiley et al., 2009). Nur sporadisch vorliegende empirische Zugänge zeigen jedoch, dass Diskurssynthesen teils erhebliche Anteile von irrelevanten oder sogar inhaltlich falschen Informationen enthalten, wie es Castells et al. (2023) und – auf geringerem Niveau – Braasch et al. (2013) sowie Graesser et al. (2007, Experiment 2) ermittelten. Solche fehlerhaften Übernahmen scheinen von der Textschwierigkeit und der Salienz der Informationen innerhalb der textuellen Makrostruktur abzuhängen: Leicht rekonstruierbare Makropropositionen in nachvollziehbar strukturierten Texten übernehmen Personen korrekter und mit höherer Ausschöpfungsquote (Rickert & Philipp, 2022).

- Der Anteil der in den Diskurssynthesen verarbeiteten Dokumente ist weniger gut untersucht. Die Mittelwerte der bisher nur punktuell vorgelegten Studien, welche sich der Anzahl der rekonstruierbaren Dokumente widmeten, liegen unter der Hälfte der zur Verfügung gestellten Dokumente, seien es vier (Barzilai et al., 2015) oder sechs (List & Du, 2021; Weston-Sementelli et al., 2018). Damit ist der Anteil höher als bei den Inhalten, was sich darüber erklären lässt, dass sich die Grundgesamtheit an Texten deutlich von jener der Informationseinheiten als Analysekatgorie unterscheidet, sprich: niedriger liegt. Gleichwohl ist es kritisch, dass Dokumente nicht erkennbar verwendet wurden, die eigens in experimentellen Studien ausgewählt wurden, um die Gesamtheit einer komplexen Thematik zu transportieren.

Damit lässt sich aus mehr als zwei Dutzend Studien ein Muster abstrahieren: Mit unterschiedlichen Auswertungseinheiten und Fokussen ist ein relativ durchgängiger Befund, dass von den potenziell für die eigene Diskurssynthese verwertbaren Informationen bzw. Informationsträgern im Falle der Dokumente nur ein teils sehr geringer Anteil Eingang findet. Zwar ist die Selektion von Informationen kein Garant für glückender Diskurssynthesen, doch wirken Quoten bei den Propositionen von unter einem Drittel ausgesprochen gering, wenn der Anspruch besteht, Sachverhalte in ihrer inhärenten Komplexität schriftlich zu verarbeiten. Kritisch ist an diesem Befund außerdem, dass empirische Hinweise vorliegen, dass die Wiedergabe von Inhalten noch am häufigsten als kognitive Operation in Diskurssynthesen vorkommt, während Organisations- und Verknüpfungsleistungen seltener zu beobachten sind (Raković et al., 2021; Spivey, 1991). Wenn also hier schon Defizite zu beobachten sind, dann sind Folgeprobleme erwartbar.

5.4.2 Problemfeld 2: geringe Integrationsleistungen in Diskurssynthesen (Probleme im Organisieren)

Das zweite Problemfeld betrifft die Art, wie Personen gelesene Informationen strukturieren und miteinander kombinieren. Dabei lassen sich drei neuralgische Punkte ausmachen: Erstens neigen Schreiberinnen und Schreiber dazu, Inhalte sequenziell aneinanderzureihen ([5.4.2.1](#)). Damit missglückt es, dem mit echten Diskurssynthesen verbundenen Anspruch zu genügen, Informationen in eine angemessene Restrukturierung zu überführen. Doch selbst wenn nicht bereits die Auflistung von einzelnen Dokumenteninhalten als Knowledge-Telling-Äquivalent (Segev-Miller, 2007) gewählt wird, existieren weitere gut dokumentierte Problembereiche. Der zweite Bereich betrifft die inhaltlich verzerrte Darstellung, gerade wenn der Gegenstand aus verschiedenen Perspektiven entfaltet wird und kontroverser Natur ist. Hier besteht die Tendenz, dass die Diskurssynthesen an einer inhaltlichen Unausgewogenheit leiden ([5.4.2.2](#)). Ein weiterer Problembereich nimmt mehrere Indikatoren in den Blick: Diskurssynthesen weisen nicht nur in einzelnen Bereichen wie dem Abdeckungsgrad von Informationseinheiten Defizite auf bzw. sind einseitig – mitunter treten gleich mehrere Probleme gleichzeitig auf. Solche Problembereiche lassen sich zu Gruppen verdichten, und ihnen widmet sich der hier dritte Schritt der Darstellung ([5.4.2.3](#)).

5.4.2.1 Additive Sequenzen von Inhalten statt echter Synthesen

Ein in der Forschung noch recht selten vorzufindender Analyseansatz besteht darin, dass die Diskurssynthesen in ihrer Struktur ausgewertet werden. Dabei steht in einem ersten Analysezugang *die Reihenfolge der gelesenen Dokumente nebst deren Verwendung* im Vordergrund, nämlich, inwieweit sich die Personen

von der thematischen Struktur der Inhalte lösen bzw. ob sie eine eigene Syntheseleistung erbringen. Dies haben einige Studien mit kleinen Stichproben untersucht (Britt et al., 1994; Castells et al., 2023; Mateos et al., 2008; Mateos & Solé, 2009; Rickert & Philipp, 2022; Segev-Miller, 2007; Solé et al., 2013). Laut diesen Studien gelingt es nur einer Minderheit der getesteten (überwiegend erwachsenen) Personen, Inhalte aus den gelesenen Texten zu reorganisieren und zu synthetisieren. Vielmehr überwiegt ein sequenziell wirkendes Vorgehen, bei dem die Inhalte textweise wiedergegeben werden. Zum Problembereich eines solchen additiven Vorgehens gehört ferner, dass es zu *geringen Restrukturierungsleistungen* kommt, wenn Personen die Dokumente eher sequenziell lesend bearbeiten und sie in der Sequenz der rezeptiven Aneignung in ihrem eigenen Text verarbeiten. Stahl et al. (1996) ermittelten beispielsweise, dass nur ein Achtel der Personen in ihrer Studie die Textinhalte gelesener historischer Dokumente nicht in der Originalreihenfolge nutzten.

Eine zweite Analysevariante zur organisierenden Integrationsleistungen orientiert sich an der *inhaltlich-thematischen Integrierbarkeit bzw. entlang der Textsortenerfordernisse in den Diskurssynthesen*. Mangelnde Integrationsleistungen manifestieren sich darin, dass je nach Studie ohne explizite Hinweise mindestens ein Drittel der Personen (teils sogar erheblich mehr) eine Textstruktur wählte, die nicht dem kommunikativen Zweck bzw. der inhärenten kausalen Sachlogik entsprach (Greene, 1993; Hemmerich & Wiley, 2002, Studie 2; Voss & Wiley, 1997; Wiley et al., 2009; Wiley & Voss, 1996), was insbesondere für vergleichende Sachtexte untersucht wurde (Kirkpatrick & Klein, 2009; Luo & Kiewra, 2019, Experimente 1 und 2; Spivey, 1991). Das Problem, welches sich in diesen Studien zeigt, ist das Unvermögen eines Großteils von Personen, eine eigene Struktur für die eigene Diskurssynthese zu finden, die sich von der Wiedergabe des Gelesenen in der Reihenfolge des Gelesenen löst. Anders gesagt: Die Personen verharren, sofern sich dies aus den Produktmaßen extrapolieren lässt, im Modus des Knowledge Tellings; insbesondere gelingt es ihnen nicht, den rhetorischen Problemlöseraum für eine angemessene Reorganisation der Inhalte zu verwenden, wie es das Knowledge Transforming erfordern würde (s. dazu auch die hohen Anteile von separaten Repräsentationen in Tabelle 27 in Teilkap. 4.4.1.1).

Integrationsleistungen bzw. Probleme im Integrieren lassen sich selbst dort nachweisen, wo es vorderhand nicht um das Verfassen eines längeren Textes ging. Ein solches Beispiel für geringe Integrationsleistungen stammt aus der Interventionsstudie von Firetto und van Meter (2018). In ihr sollten Studierende zwei Texte zu zwei Kreislaufsystemen im menschlichen Körper lesen und danach mindestens fünf Minuten ihre Erinnerungen schriftlich darlegen. Kodiert wurde, ob die Personen Verbindungen zwischen den beiden Kreisläufen benannt und damit eine intertextuelle Integrationsleistung erbracht hatten. Damit fokussierte diese Studie ein lediglich sehr grobes Maß der intertextuellen Integration. Solcherlei Leistungen spiegelten sich in maximal fünf Prozent all jener Antworten

der Studierenden aus der Kontrollgruppe bzw. einer zweiten Interventionsgruppe wider. Selbst in der Interventionsgruppe mit expliziten Hinweisen lag die Quote bei lediglich knapp einem Drittel.

5.4.2.2 Einseitige Darstellungen bei kontroversen Themen in argumentativen Diskurssynthesen

Für das Verfassen von Diskurssynthesen ist es typisch, dass Personen auf der Basis inhaltlich kontroverser Inhalte, die auf eigens zusammengestellten Dokumentensets basieren, den Sachverhalt in seiner Vielschichtigkeit, allem voran aber in seiner Widersprüchlichkeit darstellen sollen. Hierunter fällt, dass Positionen untereinander zu vergleichen und abzuwägen sind, vor allem aber dass sie direkt aufeinander zu beziehen sind, um bei einem strittigen Sachverhalt zu einer begründeten und begründbaren Positionierung zu gelangen (Nussbaum, 2021).

Das intertextuelle Integrieren von konfligierenden Aussagen stellt sowohl in der Theorie zur „Multiple Documents Literacy“ im Allgemeinen (Saux, Britt et al., 2021) als auch bei schriftlichen Diskurssynthesen in der empirischen Forschung im Besonderen (Barzilai et al., 2018; Primor & Katzir, 2018) einen markanten Schwerpunkt dar. Entsprechend gibt es inzwischen diverse Studien, die mittels Analyseraster erfassen, wie ausgewogen und vollständig die Argumentationen in ihrer Textstruktur sind. Für die Belange der Thematik der Ausgewogenheit geht es jedoch hier vor allem um die produktbezogenen Maße, welche Auskunft erteilen über die Zweiseitigkeit der Pro- und Contra-Argumente aus den Texten und ihre intertextuellen Bezüge, die herzustellen Aufgabe der Testpersonen war.

Diverse Studien, seien es Grundlagenstudien (Anmarkrud et al., 2013; Barzilai & Eshet-Alkalai, 2015; Barzilai et al., 2015; Bråten, Ferguson, Anmarkrud & Strømsø, 2013; Kullberg et al., 2023; Mateos et al., 2011; Tarchi, 2021a, 2021b; Tarchi & Mason, 2020; Tarchi & Villalón, 2021, 2022; Tarchi et al., 2022), seien es Interventionsstudien – und hier vor allem bei den Kontrollgruppen bzw. bei den Ausgangswerten vor einer Förderung (Barzilai & Ka’adan, 2017; Casado-Ledesma et al., 2021; Cuevas et al., 2016; González-Lamas et al., 2016; Granado-Peinado et al., 2019, 2023; Luna et al., 2020, 2023; Mateos et al., 2018, 2020) – kommen zu ähnlich lautenden Resultaten. Den überwiegend erwachsenen Probandinnen und Probanden gelingt es – gemessen an den Mittelwerten der Gruppen – selten, in ihren Texten *die kontroversen Standpunkte aus multiplen Dokumenten paritätisch darzustellen*. Stattdessen überwiegt allenfalls eine einseitige Rechtfertigung einer Position. Entsprechend sind die getesteten Personen, teils sogar jene in den geförderten Experimentalgruppen nach Ende der Fördermaßnahme, nicht dazu in der Lage, mehrzügige Argumente schriftlich darzulegen und die Kontroverse ausgewogen zu behandeln. Anders gesagt: Gemessen an dem Anspruch, den das schriftliche Argumentieren auf der Basis multipler, inhaltlich divergierender

Dokumente an lesende und schreibende Personen stellt, erreichen Personen laut mehr als 20 Studien oftmals nur Vorstufen.

Hinzu kommt, was bereits im Teilkapitel [5.3.3](#) beschrieben wurde: Haben die Personen starke inhaltliche Überzeugungen zu dem Thema, neigen sie ohnehin zu einer einseitigen Darstellung der Thematik im Sinne des „Text-Überzeugungskonsistenz-Effekts“. Das heißt: Die eigenen Überzeugungen erschweren eine balancierte schriftliche Darlegung eines mehrperspektivischen bzw. kontroversen Sachverhaltes. Wenn sie nicht korrigiert werden, erfolgt mithin eine verzerrt-selektive und letzten Endes inadäquate Nutzung von Perspektiven aus multiplen Dokumenten.

5.4.2.3 Oberflächliche Verarbeitung – multiple Probleme in Gruppen von Personen

Diskurssynthesen werden nicht nur auf Abdeckungsgrade und textstrukturelle bzw. intertextuell basierte Makrostrukturen hin analysiert, denn dies sind nur Ausschnitte aus einem breiten Spektrum an Leistungen (Philipp, 2022a; Primor & Katzir, 2018). Ein weiterer methodischer Zugang besteht darin, dass Einzelindikatoren von Texten als Teildimensionen separat ausgewertet werden, um auf dieser Basis dann einen Gesamtscore für die Textqualität zu bilden (s. Teilkap. [5.2.1.2](#)). Während der erste Ansatz spezifische Teilleistungen fokussiert und dadurch übergeordnete Muster ausblendet, trifft dies spiegelbildlich auf Daten aus Mittelwertbildungen mit Teildimensionen zu. Hier gehen vereinzelt einige Studien einen Mittelweg, indem sie variablenzentriert die Ausprägungen multipler Indikatoren betrachten. In zwei finnischen Studien (Hyytinen et al., 2017; Hyytinen et al., 2021) wurden in szenariobasierten, argumentativen Schreibaufträgen die Diskurssynthesen von Testpersonen nach Art der Beantwortung in drei Gruppen klassifiziert; in einer weiteren Studien wurden die Testpersonen aufgrund ihrer demonstrierten Fähigkeiten in der schriftlichen Diskurssynthese in drei Leistungsgruppen unterteilt (Hyytinen et al., 2018).

Entscheidend für die Klassifikation in den drei Studien war, wie ausgewogen und kommunikativ erfolgreich die Diskurssynthesen waren. Trotz der Unterschiede in Samples, den Aufgaben und den Anteilen der Zugehörigkeiten in den Einzelstudien gab es in den drei Gruppen Ähnlichkeiten, die im Folgenden studienübergreifend gebündelt werden:

- 1) *Oberflächliche Diskurssynthesen*: Solche Texte zeichneten sich durch Probleme in der Argumentation aus. Unklare Standpunkte, wenig abgestützte Schlussfolgerungen und Sammlungen von isolierten Aussagen waren typisch. Auch die Nutzung der Dokumente erwies sich als problematisch: Es überwogen oftmals wortwörtliche Übernahmen und/oder solche mit wenig Referenzierungen der Aussagen. Solche Texte fielen zu guter Letzt kurz aus.

Die Anteile in diesen Gruppen machten in den drei Studien zwischen einem Viertel (Hyytinen et al., 2021) bzw. einem Drittel aus (Hyytinen et al., 2017, 2018).

- 2) *In Ansätzen geglückte Diskurssynthesen*: Auch diese Textprodukte sind in der Argumentation problematisch, weil sie sich mehr auf einzelne Dokumente denn auf das Set der Dokumente stützen, und die Schreiberinnen und Schreiber verfassten Argumentationen in variierender Qualität. Tendenziell waren sie in der Lage, eine sinnvolle Schlussfolgerung zu präsentieren. Die Quellenverwendung erwies sich als eher patchworkartig. Die Anteile lagen zwischen 42 und 53 Prozent.
- 3) *Tief verarbeitete Diskurssynthesen*: Solche Produkte zeugten von einem tiefen Textverstehen, indem propositionale intertextuelle Auswahl-, Organisations- und Verknüpfungsprozesse geglückt waren. Ebenfalls hatten die Personen Informationen kritisch und angemessen beurteilt, ihre Argumentationen waren kohärent und überzeugend, und die Quellen schließlich hatten sie sinnvoll wiedergegeben und überwiegend belegt. Die Anteile lagen bei den akademisch am wenigsten sozialisierten Personen bei einem Siebtel (Hyytinen et al., 2017), in den anderen Studien mit älteren Studierenden bei mehr als einem Viertel (Hyytinen et al., 2021) bzw. fast einem Drittel (Hyytinen et al., 2018).

Die drei finnischen Studien zeigen in ihrem Gesamt, dass es mindestens jeder dritten Person nicht gelingt, multiple Dokumente in einer eigenen argumentativen Diskurssynthese angemessen, sondern nur oberflächlich zu verarbeiten und dabei weder die Dokumente adäquat zu verwenden noch eine stringente Argumentation zu entwickeln. Als normativ zwar besser, gleichwohl immer noch eher als oberflächlich ist das Vorgehen der größten Gruppe in den Studien zu bezeichnen. Diese Personen weisen in Teilen der Dokumentennutzung und der Fähigkeit, die Inhalte angemessen wiederzugeben, Vorsprünge auf. Allerdings sind sie nicht fähig, das Gesamt der Dokumente für eine argumentative Problemlösung zu verwenden. Mindestens zwei Drittel der untersuchten Personen sind demnach nicht ausreichend dazu in der Lage, Diskurssynthesen zu verfassen, die sich über zwei Hauptkategorien (mit diversen Unterkategorien) als qualitativ hochwertig klassifizieren lassen. Ähnliches ließ sich ebenfalls bei Jugendlichen und in eher erklärenden Sachtexten als Zieltextsorte wiederfinden (s. Kasten „Illustrierendes Beispiel 23“).

Illustrierendes Beispiel 23

Von Minimalist:innen, Replikator:innen und Transformierenden – Gruppen von expositorischen Diskursynthesen

A) Was war das Erkenntnisinteresse?

In einer Studie zur Testentwicklung interessierte sich das Forschungsteam dafür, wie verschiedene Indikatoren von kognitiven Prozessen, die über Diskurssynthesen rekonstruiert wurden, dazu geeignet sind, unter Jugendlichen Leistungsgruppen zu bestimmen (Goldman et al., 2013).

B) Wer wurde untersucht?

Das Sample bildeten 211 US-amerikanische Schülerinnen und Schüler der Klassenstufen 5–7, die Klassenstufe 6 war mit 43 Prozent überrepräsentiert. Die Publikation beschreibt das Sample nicht genauer hinsichtlich weiterer soziodemografischer oder anderer Merkmale.

C) Welche Instrumente kamen zum Einsatz?

In der Studie wurde ein digitales Assessment entwickelt, welches drei kognitive Prozesse (Analyse, Synthese und Integration) aktivieren und deren Produkte erfassbar machen sollte. Hierfür wurden in einer digitalen Lese- und Schreibumgebung drei komplementäre Dokumente zu einer gesellschaftswissenschaftlichen Frage – dem Erfolg von Bürgerrechten in den 1950er und 1960er Jahren – zur Verfügung gestellt. Die Schülerinnen und Schüler erhielten den Auftrag, eine Erklärung zu schreiben, und bekamen Zusatzinstruktionen zum Vorgehen. Es gab eine 10–15 Minuten lange Lese- und eine 15- bis 30-minütige Schreibphase. Die Texte waren in der Schreibphase verfügbar, ein Kopieren und Einfügen der Texte war jedoch technisch ausgeschlossen.

Im Fokus der Analyse standen die Diskurssynthesen, die mehrdimensional kodiert wurden. Zentral waren die drei kognitiven Prozesse, die über Produktmaße operationalisiert wurden. Die 1) Analyse wurde über die Übernahme von Propositionen aus den Dokumenten erfasst. Die 2) Syntheseleistungen wurden über die Zahl der genutzten Dokumente abgebildet. Für die 3) Integration standen intertextuelle Inferenzen bzw. die Verknüpfungen mit Vorwissen. Integrations- und Analyseleistungen wurden in der Studie so ausgewertet, dass die Inhalte der Diskurssynthesen prozentual den Kategorien zugeteilt wurden. Diese Produktmaße wurden mit Clusteranalysen dazu verwendet, Gruppen zu bilden, die Indikatoren der Synthese wurden nicht verwendet. Diese wurden zusammen mit Textlängen und Prozessmaße im Sinne von Bearbeitungszeiten bei der Beschreibung der drei ermittelten Gruppen berichtet.

D) Was waren die Hauptbefunde?

Die Clusteranalyse ergab drei Gruppen, die als Minimalist:innen, Replikator:innen und Transformierende bezeichnet wurden, die sich über alle Klassenstufen hinweg ähnlich verteilten. Minimalist:innen machten mit 48 Prozent die größte Gruppe aus, gefolgt von Transformierenden (31 %) und Replikator:innen (21 %). Zu den Gruppen im Einzelnen:

- *Minimalist:innen* schrieben kurze Texte und für kurze Zeit. Sie nutzten zu 60 Prozent nur ein Dokument, integrierten entsprechend ebenfalls wenig. Ihre Diskurssynthesen bestanden zu zwei Dritteln aus übernommenen Inhalten aus den Dokumenten, zu fast einem weiteren Fünftel aus verzerrt wiedergegebenen Inhalten.
- *Replikator:innen* verfassten doppelt so lange Texte und schrieben am längsten. Vier Fünftel der Inhalte ihrer Diskurssynthesen waren Übernahmen von Dokumenteninhalten (und davon drei Viertel: direkte). Die Inhalte waren weitestgehend korrekte Wiedergaben, Inferenzen kamen selten vor. Sie nutzten als Gesamtgruppe im Durchschnitt zwei Dokumente.
- *Transformierende* nahmen in Umfang und der Schreibzeit eine mittlere Position zwischen den beiden obigen Gruppen ein. Allerdings fielen sie dadurch auf, dass ihre inhaltlichen Übernahmen vor allem Paraphrasen waren. Sie hatten die wenigsten inhaltlichen Fehler und die höchsten Anteile von Inferenzen, wenngleich immer noch in geringen prozentualen Anteilen. Sie verwendeten zwei bis drei Dokumente, schöpften also das Dokumentenset aus.

E) Diskussion

Die Studie hatte zum Hauptergebnis, dass nur ein Drittel der getesteten Personen dazu in der Lage war, Integrations- und Syntheseleistungen in schriftlichen Produkten zu demonstrieren und dabei Informationen angemessen zu transformieren. Besonders der hohe Anteil von entweder rein reproduzierenden oder mit geringem Aufwand oberflächliche Texte produzierenden Personen – letztgenannte Gruppe bestand aus jeder zweiten Testperson – illustriert, dass Jugendliche bei Diskurssynthesen Unterstützungsbedarf haben.

5.5 Fazit

Das Lesen von analogen und digitalen Texten erfolgt nicht nur, um Inhalte zu verstehen, sondern kann durchaus einer Verwertungsabsicht wegen erfolgen, etwa dem Produzieren von einem eigenen Text. Aufgabenprodukte sind von der lesenden Person auf der Basis des Gelesenen sowie des eigenen Wissens und für einen (durch eine Aufgabe) definierten Zweck hergestellte externale Darstellungen des Gelesenen. Darunter fallen mündliche Vorträge, grafische Darstellungen oder die in diesem Kapitel primär behandelten Diskurssynthesen. Diskurssynthesen sind eigene (expositorische) Texte, die auf hybriden Lese- und Schreibaktivitäten basieren, welche dazu dienen, aus gelesenen Dokumenten über Prozesse der Informationsauswahl, -organisation und -verknüpfung ein eigenes Textprodukt zu

generieren. Als ein idealtypisches Vorgehen lässt sich aus der Schreibforschung das Transformieren des Wissens („Knowledge Transforming“) benennen, in dem inhaltliche und rhetorische Transformationsprozesse interagieren (Raković & Winne, 2022). Hierfür müssen Personen selbstreguliert die angestrebte kommunikative Absicht ihres eigenen expressiven Aufgabenprodukts ebenso nutzen, wie sie gefordert sind, in der rezeptiven Verwendung der Dokumente den Gegenstand überhaupt erst einmal sukzessive zu konstruieren und zu konstituieren. Auch wenn es also aus darstellerischen Gründen in diesem Kapitel nicht eigens so fokussiert dargestellt wurde: De facto ist das Herstellen einer Diskurssynthese ein rekursiver, selbstregulatorische Fähigkeiten voraussetzender bzw. doch mindestens einfordernder Prozessverbund (Vandermeulen et al., 2023).

Damit ist benannt, dass das Schreiben von Diskurssynthesen als *Verbund interagierender Lese- und Schreibprozesse* zu konzeptualisieren ist. Empirisch zeigt sich, dass sich solche Lese- und Schreibprozesse differenziert erfassen lassen. Zwei Gruppen lassen sich grob unterscheiden, nämlich die unspezifischen Lese- und Schreibprozesse, welche Hinweise darauf geben, dass vor allem temporal bestimmbare Aktivitäten des Umgangs mit schriftlichen Informationen erfolgen, und die spezifischen Prozesse. Diese letztgenannte Gruppe ist präziser, weil sie die Arten und Bezugspunkte sowie die Funktionen der Prozesse klarer fasst und differenziert. Beide unterschiedlichen Datenquellen, die sowohl bei der stärker erforschten Domäne Lesen als auch beim Schreiben vorgekommen sind, hängen partiell mit Produktmaßen zusammen.

Produktmaße bei den Diskurssynthesen werden aktuell ähnlich heterogen operationalisiert wie die Prozessindikatoren. Aus der Vielzahl von Textmaßen ließen sich stellvertretend acht Gruppen von Variablen aus jenen Studien extrahieren und aggregieren, in denen neben den Produktmaßen auch mindestens ein weiterer Prozessindikator oder sogar mehrere erfasst wurden. Gesamthaft existieren in der Forschung weitaus mehr Produktmaße (s. Primor & Katzir, 2018, für einen Teilüberblick jener Studien, in denen ausschließlich das intertextuelle Integrieren mittels Diskurssynthesen und weiteren schriftlichen Äußerungen von Testpersonen operationalisiert wurde).

Der in diesem Kapitel vorgenommene Forschungsüberblick hat vor allem zum Hauptergebnis, dass der Kenntnisstand zum Zusammenhang von Prozess- und Produktmaßen fragmentarisch ist und metaphorisch gesprochen einer noch wenig gefüllten Leinwand gleicht. Da zudem viele Studien unterschiedliche Aufgaben mit unterschiedlichen Analysefokussen einsetzten, fällt es schwer, studienübergreifend konsistente Muster auszumachen. Gleichwohl bestehen Tendenzen bei den positiven Zusammenhängen. Stark gebündelt lassen sie sich so benennen: *Zeitlich ausführlichere, von Rekursionen geprägte Bearbeitungen, zumal solche, in denen Informationen aus gelesenen Dokumenten bereits einer Selektion und Restrukturierung mit Blick auf das verlangte spätere eigene Textprodukt unterliegen, sind vorteilhaft.* Als idealisiertes Vorgehen umfasst das Verfassen von

Diskurssynthesen demnach planerische, also metakognitiv einer Zielerfüllung dienende Lese- und Schreibprozesse, aber es inkludiert ebenso adaptive, auf den dynamischen Prozess produktiv reagierende Prozessverbünde. Lese- und Schreibstrategien interagieren demnach günstigstenfalls ergebniszentriert und -offen, was sich aus der Notwendigkeit ergibt, schritt- und etappenweise einen Problemlöseraum zu queren.

Wie gut dieses Unterfangen glückt, hängt von individuellen Merkmalen ab, von denen abgesehen von Leistungsgruppendifferenzen vier ausführlicher dargestellt wurden. Dabei ist zunächst ein Ungleichgewicht zu attestieren: Produktmaße überwiegen in der Forschung, was in der Ergebnisdarstellung dazu geführt hat, dass bei drei der vier Individualmaße – thematischen Überzeugungen sowie den Lese- und Schreibfähigkeiten – nur Produktmaße ausreichend vorhanden sind, um studienübergreifende Muster zu erkennen. Lediglich das inhaltliche Vorwissen ist in einer Reihe von Studien mit Prozessindikatoren in den Blick der Forschung geraten. Allerdings hat sich bei dieser Variable eine besonders große Uneindeutigkeit in der konsultierten Empirie ergeben: Das Vorwissen scheint kein Garant dafür zu sein, dass Prozesse bzw. Produkte besser ausfallen. Die Forschungsergebnisse werfen diesbezüglich aktuell mehr Fragen auf, als sie beantworten. Eindeutiger sind hingegen die Befunde bei den verbleibenden drei Variablen. Insgesamt gerieten Diskurssynthesen mehrheitlich dann besser, zumal in anspruchsvollen Maßen der Struktur und der Qualität, wenn Personen bessere rezeptive und produktive Leistungen in anderen produktbezogenen Tests erbracht hatten. Hohe Lese- und Schreibkompetenz korrespondieren also mit besseren Leistungen bei den Diskurssynthesen. Ebenfalls ist es günstig für Diskurssynthesen hinsichtlich ihrer argumentativen Ausgewogenheit von Positionen, wenn die Testpersonen keine allzu starken inhaltlichen Überzeugungen aufweisen. Dies leitet bereits über zu Problemlagen beim Verfassen von Diskurssynthesen.

Die Forschung zeigt bei aller Divergenz der vielen Zugänge in den Studien, dass es in zwei Bereichen zu relativ konsistenten ungünstigen Mustern kommt. Das erste Problem besteht darin, dass in den Diskurssynthesen nur wenige jener Informationen übernommen wurden, welche für die inhaltliche Darstellung des Sachverhalts zentral wären. Ein weiteres, zweites Problem ist in einer zu geringen Organisation bzw. Strukturierung der Inhalte zu lokalisieren. Dieses Problem ist vielgestaltig. Es besteht darin, dass echte Syntheseleistungen aus zweierlei Gründen nicht erkennbar sind. Zum einen ist ein Problem, dass lesende und schreibende Personen dokumentenweise vorgehen und damit additiv Inhalte in einer Sequenz darstellen, die eher dem Prinzip der Reihenfolge der Aneignung (Knowledge Telling) denn einer thematisch sinnvollen Synthetisierung (Knowledge Transforming) folgen. Zum anderen missglückt es häufig, kontroverse Themen ausgewogen wiederzugeben. Stattdessen besteht die Tendenz, einseitige Darstellungen vorzunehmen. Die oben schon erwähnten thematischen Überzeugungen und mit ihnen der „Text-Überzeugungskonsistenz-Effekt“, der als

unerwünschter Filter eine Wiedergabe der überzeugungskongruenten Inhalte fördert, sind eine Erklärung dafür. Zudem deuten Studien noch zusätzlich darauf hin, dass die analytisch trennbaren Problembereiche bei der Auswahl und der Integration von Inhalten miteinander auftreten. Anders gesagt: Personen stehen durchaus vor mehreren Herausforderungen, die sie in ihrem Gesamt mehr oder weniger meistern.

Die Aufgabenprodukte des (digitalen) Lesens sind insgesamt vielgestaltiger, als es der hier eingeschränkte Fokus auf Diskurssynthesen nahelegt. Gleichwohl zeigt sich in den Diskurssynthesen prototypisch, worin die Anforderungen bei der zielbezogenen Verwertung von Informationen und Dokumenten bestehen: Es geht um eine adäquate Überführung der vorgefundenen und auf verschiedene Weisen zieladäquat zu transformierenden und zu prozessierenden Informationen. Dafür sind ganze Verbünde und hohe Kompetenzen in rezeptiven und produktiven Fähigkeiten nötig, welche zudem ausreichend gesteuert und koordiniert werden müssen. Dieser hohen Bedeutung der Selbstregulation geht nun das folgende und letzte Kapitel nach.

6 Komponente 5: Selbstregulation

Darum geht es

Digitales Lesen erfordert die Koordination verschiedenster Komponenten, damit Personen in einem insgesamt offenen und selbstständig zu strukturierenden Prozess(-verbund) Erfolg haben. Deshalb ist es naheliegend, in diesem Kapitel ein Konzept ausführlich zu behandeln, das seit geraumer Zeit in der Bildungsforschung einen wichtigen Gegenstand bildet: Selbstregulation bzw. dessen schon auf den Erwerb von Wissen enggeführten Ausschnitt des selbstregulierten Lernens. Selbstregulation ist ein psychologisches Rahmenkonstrukt, das die eigenaktive Steuerung von verschiedenen Prozessen und Aktivitäten betrifft, welche gesamthaft darauf ausgerichtet sind, einen Zielzustand zu erreichen. Die dafür nötigen Ressourcen und Komponenten werden je nach theoretischem Modell anders akzentuiert, doch die Perspektive auf einen komplexen Prozess, der metakognitiv gesteuert wird und des Einsatzes von Strategien bedarf, bildet eine wichtige Gemeinsamkeit der Modellierungen. Dies ist eines der Hauptergebnisse eines mehrdimensionalen Vergleichs von acht Modellen des selbstregulierten Lernens, der zu Beginn des Kapitels erfolgt. Hinzu kommt, dass Selbstregulation biologische Grundlagen hat: Die sogenannten „exekutiven Funktionen“, die als Basis für Kognitionen höherer Ordnung wie die Selbstregulation gelten, beinhalten Aufmerksamkeitslenkung, Wechsel zwischen Aufgaben und die Abschirmung vor interferierenden Impulsen. Angereichert um generische bzw. domänenspezifische Metakognitionen als deklaratives Wissen oder als dessen prozedurale Ausführung fungieren somit diverse Prozesse im Verbund gleichsam als kognitive Werkzeuge ([6.1](#)).

Entscheidend für gelingende Selbstregulation ist, dass Personen die Diskrepanz zwischen angestrebtem Ziel und gegenwärtigem Zustand erkennen und die adäquaten Vorgehensweisen wählen, um die Diskrepanz zu minimieren. Hierfür spielen kognitive (strategische) Prozesse auf der Objektebene des Lesens eine Rolle, die von darüberliegenden Metakognitionen auf einer Metaebene im Sinne von Top-down-Prozessen geplant werden (Kontrolle) und deren Erfolg bei der Umsetzung auf der Objektebene bottom-up-basiert einem Monitoring unterliegen (Überwachung). Beide Prozesse der Metakognitionen zwischen den Ebenen laufen reziprok und dynamisch ab. Diese Dynamik bildet entsprechend auch das Herzstück dessen, was in dem exemplarisch entfalteten Modell COPEs als Komponentenmodell zur prozessualen Ausgestaltung beschrieben wird. Ihm werden empirische Befunde zur Seite gestellt, welche darauf aufmerksam machen, dass selbstreguliertes Lernen im Allgemeinen und selbstreguliertes Lesen im Besonderen durch das Zusammenspiel von verschiedenen kognitiven und metakognitiven Strategien gekennzeichnet sind ([6.2](#)).

Als Verbund multipler Prozesse ist es bei der Selbstregulation erwartbar, dass diverse individuelle und kontextuelle Merkmale Folgen für die konkrete Ausgestaltung und Dynamik des Prozesses haben. Drei Merkmale fokussiert das Kapitel, und sämtliche dieser Merkmale sind kognitiver bzw. metakognitiver Provenienz. Als bedeutsam, wenn auch nicht immer in der

Deutlichkeit konsistent erwiesen sich bislang Leistungsgruppendifferenzen. Ihnen zufolge trennt erfolgreiche von weniger erfolgreichen Personen, dass die erstgenannte Gruppe mehr vorwissensbasierte Elaborationsstrategien und mehr metakognitive Überwachungsstrategien praktiziert und zudem Strategien dynamischer bzw. adaptiver einsetzt. Dieses Grundmuster von besserer Nutzung von Vorwissen und Metakognition beim selbstregulierten Lesen lässt sich folgerichtig bei einer weiteren zentralen Einflussgröße nachzeichnen: dem Vorwissen. Personen mit mehr Wissen in der Domäne nutzen dieses besser und stärker metakognitiv überwacht. Schließlich sind noch subjektive Theorien bzw. genauer: die epistemischen Überzeugungen eine bedeutende Variable: Fortgeschrittene Überzeugungen dazu, was als Wissen gilt und wie solches Wissen produziert wird, fungieren als interne Standards vor allem bei komplexen Aufgaben (6.3).

Selbstregulation als kontingenter Prozess mit vorgängig, synchron und nachträglich zu koordinierenden Komponenten verlangt den Personen einiges ab. Darum lässt sich annehmen, dass Selbstregulation voller Sollbruchstellen ist. Drei Problembereiche sind, dass erstens Personen ihre Aufmerksamkeit zu wenig regulierend steuern, was Prozesse erschwert und Produkte beeinträchtigt. Ein zweites, metaanalytisch breit abgestütztes Problemfeld ist, dass Leserinnen und Leser leider nur ungenau metakognitiv repräsentieren, wie gut sie Texte verstanden haben. Das macht die auf präzisen metakognitiven Überwachungen basierenden Entscheidungen fehleranfällig, weil eine mangelhafte Kalibrierung zu optimistischen Sichtweisen verleitet und zu einem oberflächlichen Lesen führt. Ein drittes Problemfeld schließlich ist darin zu lokalisieren, dass Leserinnen und Leser von sich aus zu wenig ein gesamthaft lernstrategisches, auf Tiefenverstehen abzielendes Vorgehen demonstrieren. Hier ist also der vernetzte Verbund von kognitiven und metakognitiven Strategien defizitär (6.4).

6.1 Funktionen

Die Ausführungen zu den ersten vier Komponenten des digitalen Lesens aus den Kapiteln 2 bis 5 haben immer wieder deutlich gemacht, dass viele der Aktivitäten beim Lesen digitaler Texte durch eine teils große Offenheit gekennzeichnet sind. Sie erfordern damit von den Personen Koordination, komplexe Aktivitäten selbstständig so auszuführen, dass die angestrebten Ziele aus eigener Kraft erreicht werden. Paradigmatisch ist damit ein Konstrukt angesprochen, welches in der Lehr-Lernforschung seit geraumer Zeit erhöhte Aufmerksamkeit auf sich gezogen hat: die *Selbstregulation*, häufig auch in der Kombination mit dem Lernen, also das *selbstregulierte Lernen* (Dinsmore et al., 2008).

Selbstregulation ist ein Sammelbegriff – ergo ein Meta-Konstrukt – aus der Psychologie, der in unterschiedlichen Teildisziplinen jeweils leicht unterschiedlich gefüllt wird (Inzlicht et al., 2021). Im Bereich der pädagogischen Psychologie wird Selbstregulation im Kontext des Lernens betrachtet, also in einem Ausschnitt all dessen, wozu selbstregulatorische Prozesse benötigt werden. Eine weit gefasste Definition zeigt die Breite des Dachkonzepts selbstreguliertes Lernen (SRL):

„Beim SRL geht es darum, wie Lernende versuchen, Wissen aufzubauen, indem sie mentale Inhalte und Faktoren in der äußeren Umgebung metakognitiv überwachen und regulieren. Zu den Objekten der metakognitiven Kontrolle gehören Kognition, Entscheidungsfindung, Planung, Motivation, Affekt und Emotion. Zu den Zielen der metakognitiven Kontrolle zählen: (1) grundlegende kognitive Operationen wie Üben oder Übersetzen sowie umfassendere Lernstrategien; (2) motivational orientierte Gedanken über die Gründe für das Lernen (z. B. Orientierung an Beherrschungs- oder Leistungszielen) und die Erklärungen (Attributionen) und den Wert (die Werte) der durch das Lernen erzielten Ergebnisse; (3) die Wahrnehmung von Gefühlen (Affekt, Emotion); und (4) Faktoren in der äußeren Umgebung, z. B. Möglichkeiten, Hilfe zu erhalten, und die Frage, ob Informationen extern in Notizen gespeichert werden sollen. Nachdem die Weichen für eine Lernepisode gestellt wurden, verändern die Entscheidungen, die der Lernende im Rahmen von SRL trifft, im Laufe der Episode diese Bedingungen. SRL ist ein dynamischer Prozess der Anpassung zur Unterstützung des Lernens.“ (Winne, 2015, S. 535 f.)

An dieser Definition fällt auf, a) dass sie sich auf verschiedene Bereiche erstreckt (Motivation, Affekte, Kognitionen, Umwelt und (eher implizit) Verhalten), b) dass sich die Prozesse auf das Erreichen von Zielen beziehen, c) dass diese Prozesse von der Person selbst initiiert und aufrechterhalten werden müssen und d) dass es eine zeitliche Perspektive mit Rückkopplungen gibt, die als Episoden bezeichnet werden. Damit sind Aspekte angesprochen, die auch für die offene Suche nach Informationen in digitalen Texten und Suchmaschinen nötig sind, um ein Aufgabenprodukt zu erstellen und dabei die nötigen Zwischenschritte selbstständig absolvieren zu können (Cho & Afflerbach, 2017; Wylie et al., 2018). Darin zeigt sich, dass Selbstregulation als ein theoretisches Meta-Konstrukt dazu fungiert, die vielfältigen Prozesse selbstständiger Vorgehensweise zu modellieren und in ihrer Gesamtheit in den Blick zu nehmen. Mit dem Ausdruck „Meta-Konstrukt“ ist angesprochen, dass sich mikro-theoretische Vertiefungen und Bestandteile unter dem Dachkonzept Selbstregulation versammeln lassen, da eine einzige Theorie – z. B. in den klassischen Bereichen Kognition und Motivation – nicht genügt, um selbstregulatorische Prozesse abzudecken. Es ist deshalb auch kein Zufall, dass eine Vielzahl von Modellierungen von Selbstregulation vorgelegt wurde, was der Komplexität des Gegenstands durch Spezifikationen der Prozesse gerecht wird. In ihrem Gesamt wirkt das Konstrukt dadurch jedoch unübersichtlicher, sodass an dieser Stelle eine begriffliche und konzeptionelle Annäherung erfolgen muss.

Die Darstellung und Entfaltung dieser Thematik erfolgt in drei Hauptschritten. Zunächst verdient das Konstrukt Selbstregulation ausführliche Beachtung und Betrachtung. Es wird anhand prototypischer Merkmale geschärft und entlang eines Vergleichs von acht verschiedenen theoretischen Modellierungen herausgearbeitet, was Konsens in der Selbstregulationskonzeptualisierung ist ([6.1.1](#)).

Eine wichtige Essenz liegt darin, dass Selbstregulation in hohem Maße durch prozedurale Metakognition geprägt ist, die sich sowohl durch Kontroll- und Steuerungs- sowie Überwachungsprozesse auszeichnet als auch zielbezogene Kognitionen bei der Beschäftigung mit dem Lern-/Lesegegenstand im Sinne von Lernstrategien nutzt. Deshalb wird dieses Verhältnis von objekt- und darüber befindlichen metakognitiven Repräsentationen und ihrer prozeduralen Verknüpfung in einem weiteren Schritt thematisiert (6.1.2). Zu guter Letzt wird noch ein Bereich angesprochen, der für die Selbstregulation eine Basis darstellt, nämlich die Nutzung der exekutiven Funktionen. Dieser Verbund verschiedener, vor allem für die Nutzung der Aufmerksamkeit im Sinne ihrer Lenkung, Abschirmung, aber auch flexiblen Lenkung nötiger Komponenten wird als biologische Grundlage der Selbstregulation erachtet (6.1.3).

6.1.1 Begriffliche und konzeptuelle Schärfungen zum Konstrukt „Selbstregulation“

Ein so ertragreiches und weites Feld der Lernforschung wie jenes zur Selbstregulation hat zu einer außerordentlich hohen Aktivität in Theorie und Empirie geführt. Zur Wahrheit gehört, dass es dem selbstregulierten Lernen als ohnehin metatheoretischem Konstrukt inhärent ist, verschiedene Konstrukte explizit in sich zu vereinen. Dies erschwert naturgemäß die Abgrenzung. Das beginnt bereits auf der begrifflichen Ebene. „Selbstregulation“, „selbstreguliertes Lernen“ und „Metakognition“ etwa sind drei konzeptionell verwandte Konstrukte, was mit ihrem gemeinsamen Kern – zum einen der Bewusstheit eines Individuums für seine eigenen Prozesse und zum anderen der Intentionalität, selbst zu handeln (Dinsmore et al., 2008; Kaplan, 2008) – zu tun hat. Selbiges gilt für einen weiteren Bereich, der mit den drei genannten Termini zu tun hat, nämlich den „Strategien“, welche insbesondere für das selbstregulierte Lernen von Belang sind. Auch hier werden verschiedene Begriffsdefinitionen und Taxonomien vorgeschlagen (Dinsmore, 2017b; Dinsmore & Hattan, 2020; Dumas, 2020; Gu, 2012; Hattie & Donoghue, 2016). Damit geht einher, dass es selbst in der Wissenschaft zu einer teils erheblichen begrifflichen Unschärfe kommt (Dinsmore, 2017b; Dinsmore et al., 2008). Es sind daher also einige Klärungen nötig.

Eine solche Schärfung erfolgt in diesem Teilkapitel anhand eines Modellvergleichs. Anhand der Extraktion von Merkmalen acht verschiedener, in der Forschung etablierter Modellierungen mit unterschiedlichen Zielsetzungen und Erklärungsansätzen, wie Selbstregulation funktioniert, sollen Gemeinsamkeiten und Unterschiede deutlicher profiliert werden. In der Folge wird klarer, was den Kernbestand gegenwärtiger Modellierungen ausmacht und wo Eigenheiten und Schwerpunkte bestehen. Eine gesonderte Darstellung der einzelnen Modelle erfolgt aus darstellungsökonomischen Gründen nicht.

6.1.1.1 Typische Bestandteile und Merkmale von Modellen selbstregulierten Lernens – ein mehrdimensionaler Vergleich

Eine erfreuliche Entwicklung in der Forschung zur Selbstregulation und zum selbstregulierten Lernen ist die Genese von Theorien zum Konstrukt. Diese Entwicklung hat beispielsweise dazu geführt, dass mehrere einflussreiche Modelle entwickelt und erforscht worden sind (für Überblicke über jeweils einige Modelle s. Panadero, 2017; Puustinen & Pulkkinen, 2001; Winne, 2015; Modellvergleiche entlang verschiedener Dimensionen nehmen Panadero, 2017, sowie Puustinen und Pulkkinen, 2001, vor). Insbesondere der Vergleich von mehreren einflussreichen Modellen des selbstregulierten Lernens von Panadero (2017) und die Vorarbeiten von Puustinen und Pulkkinen (2001) sind instruktiv, weil sie Systematisierungen von verschiedenen Modell-Dimensionen vorgenommen haben.

Mehrdimensionale Vergleiche und damit auch Achsen der Ähnlichkeit und Differenz sind also schon seit Längerem Gegenstand der theoretischen Modellierung und szientifischen Wissensakkumulation zur Selbstregulation im Kontext des Lernens. Mit der Genese neuerer Modelle und einer fortschreitenden Tendenz, Synopsen von Modellen bzw. Rahmenmodelle vorzulegen, erhöht sich die Komplexität und damit die Notwendigkeit, diese Modelle zu überblicken. Darum erfolgt an dieser Stelle eine solche knappe Synopse.

Eine umfassende Darstellung verschiedener Modelle liegt (allein aus Platzgründen) außerhalb der Ziele dieses Kapitels, dies leisten die bereits erwähnten Überblicksbeiträge bereits. Stattdessen werden in diesem Teilkapitel in mehreren Schritten acht verschiedene, teils bereits sehr etablierte, teils neuere Modelle mit neuen Schwerpunkten betrachtet. Ausgangspunkt der vergleichenden Analyse ist eine tabellarische Gegenüberstellung dieser Modelle entlang mehrerer, teils bereits etablierter Vergleichsdimensionen (s. Tabelle 31). Es folgen dann mehrdimensionale Vergleiche zur theoretischen Verortung, zu einzelnen Komponenten über die Modelle hinweg und zur zeitlichen Logik innerhalb der Modelle.

6.1.1.2 Gesamtschau: eine tabellarische Übersicht über acht Modellierungen

Als erster Zugang, die Schnittmengen zwischen Modellen zu bestimmen, enthält die nachstehende Tabelle 31 diverse deskriptive Informationen zu den einzelnen Modellen. Diese Modelle sind bereits überwiegend von Panadero (2017) sowie Puustinen und Pulkkinen (2001) ebenfalls analysiert worden, dann jedoch mit divergierendem Erkenntnisinteresse. Nicht Gegenstand der vergleichenden Perspektiven in den beiden genannten theoretischen Überblicksarbeiten waren die Modelle von Muis et al. (2018) und Hadwin et al. (2018). Während das erstgenannte Modell aus einer zeitlichen Perspektive in den beiden früheren Überblicksbeiträgen nicht berücksichtigt werden konnte, trifft dies auf zweitgenannte

nur partiell zu. Panadero (2017) widmete sich einer ersten Fassung des Modells (Hadwin et al., 2011), die sich in ihrem Aufbau jedoch in der aktuellen Fassung deutlich von der ersten Variante unterscheidet. Deshalb weicht die Darstellung in Tabelle 31 von dem ab, was Panadero (2017) über das Modell berichten konnte.

Zum Aufbau und zur Darstellung der Modelle in Tabelle 31: Die Tabelle differenziert zwischen allgemeinen Merkmalen der Modelle (A), geht dann auf die Repräsentationen zentraler Parameter ein (B) und widmet sich im letzten Block den Phasen und den darin beteiligten Komponenten (C). Der tabellari-schen Sammlung von Informationen folgen mehrere Analyseschritte, in denen modellübergreifende Vergleiche stattfinden. Quervergleiche der Modelle erfolgen im Anschluss an die Tabelle in gesonderten Abschnitten.

Tabelle 31: Vergleich von Modellen selbstregulierten Lernens

Vergleichsdimension	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Modelle			Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides	Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al.	Hadwins et al. Modell der Selbstregulation in kooperativen Settings
A) Kurzschilderung des Modells										
A1) Hauptpublikationen zur Modellexplikation	Winne & Hadwin (1998) (Vorläufer: Butler & Winne, 1995, s. Teilkap. 6.2.1)	Borkowski et al. (2000)	Zimmerman (2000) (genauere Darstellung: Panadero & Alonso-Tapia, 2014, sowie Zimmerman & Moylan, 2009)	Pintrich (2000)	Boekaerts (2011) (Vorläufer: Boekaerts, 2006 – trotz gleichen Modells verschiedene Foki: 2006: Motivation, 2011: Emotion)	Efklides (2011)	Muis et al. (2018) (Vorläufer: Muis, 2007)	Hadwin et al. (2018) (Vorläufer: Hadwin et al., 2011, mit starker Modifikation des Ursprungsmodells)		
A2) Hauptanliegen des Modells	Modellierung des individuellen selbstregulierten Lernens aufgrund eigener Ziele als dynamisch-rekursives Vorgehen mit Nutzen der gleichen Modellkomponenten	Erwerbsperspektive von mehreren kognitiven Strategien über metakognitives Planen und Überwachen der Strategien hin zur Genese von motivationalen Überzeugungen als mehrstufige Langfristspektive modellieren	Modellierung des strukturalen Zusammenspiels günstiger selbstregulatorischer Prozesse als Zyklus	Taxonomische Synthese von Modellen im Sinne eines allgemeinen Rahmenmodells von selbstreguliertem Lernen, dargestellt als Tabelle, die Phasen und Bereiche der Selbstregulation systematisiert	Selbstregulation als Zwei-Plad-Modellierung, in der der Zuwachs an kognitiven Fähigkeiten und Wissen („Wachstumspfad“) bzw. Erhalt des affektiven Wohlbefindens („Wohlbefindenspfad“) austariert werden, dabei Interaktionen und Wechsel zwischen den beiden Pfaden möglich	Modellierung des Zusammenspiels von (meta-)kognitiven, motivationalen und affektiven Personmerkmalen auf einer Makroebene (Personenebene) und der konkreten Ausgestaltung auf einer Mikroebene (Aufgabe-x-Personenebene) im dynamischen Zusammenspiel mit der konkreten Aufgabe	Synthese bestehender Modelle (vor allem COPES und Pintrichs Modell) mit zusätzlicher Anreicherung um epistemische Überzeugungen und Emotionen	Spezifikation der Regulationsprozesse auf individual- und Gruppenebene bei kooperativem Lernen, das selbstregulatorische Prozesse und Prozesse der sozialen Regulation (sozial geteilte und Co-Regulation) inkludiert		

Vergleichsdimension	Modelle					
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Bockaerts' Zweifels-Modell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides
A3) Paradigmatische Verortung	Metakognition beim Lernen, Modelle der Informationsverarbeitung	Idealbild der gut informierten, arbeitenden bzw. lernstrategischen Person („Good Information Processor“)	Sozial-kognitive Lerntheorie (speziell unter Erwerbsperspektive)	Motivations-theoretische Perspektive, v.a. zu Zielorientierung; deutliche Nähe zu sozial-kognitiver Lerntheorie	Primär pädagogisch-psychologische Perspektive mit Synthese verschiedener Theorien, v.a. in puncto akademischer Emotionen	Metakognition, aber ergänzt um Affekte zur Klärung des Zusammenhangs von metakognitiven und affektiven Prozessen
						Epistemische Kognitionen und Motivation, bei selbstreguliertem Lernen Synopse von Modell-Architekturen
						Winnes und Hadwins (1998) Modell mit Erweiterung um soziokognitive Perspektiven, um soziale Regulation zu fassen
B) Metakognitive und kognitive Repräsentation zentraler Elemente						
B1) Kontext	Nur impliziter Bestandteil, z. B. über externe Rückmeldungen, keine eigene dezidierte Berücksichtigung im Modell	Impliziter Bestandteil, da vor allem günstige Merkmale des Erwerbskontexts nicht im Modell selbst, sondern aus Erwerbsperspektive benannt werden	Nicht spezifiziert	Als eigener Bestandteil der Selbstregulation enthalten, v.a. bzgl. der im Modell kombinierten Aufgaben- und Kontextbedingungen und -merkmale	Zusammen mit Aufgabe als „Aufgabe innerhalb des Kontexts“ zusammengefasst	Nicht spezifiziert
						(Impliziter) Bestandteil als Teil der externalen Umgebungen, im Modell aus Komplexitätsreduktionsgründen ausgespart, in Phase 1 als Bestandteil der Aufgabendefinition
						Kontext zentrale Bedingung der Regulation, der über Prozesse der geteilten Regulation auf Individualebene einwirkt; dabei werden reziproke Effekte angenommen

Vergleichsdimension	Modelle					
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Boekaerts' Zweiprozess-Modell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides
B2) Aufgabe	Einerseits als externe Bedingung und damit explizit vorgesehene Auslöser im Modell, andererseits auch als internalisierte metakognitive repräsentierter Bestandteil (Aufgabendefinition)	Unterspezifiziert im Modell, wird vor allem über die Aufgabenanalyse als Bestandteil der metakognitiven Prozesse behandelt und in Verbindung mit konditionalem Strategiewissen thematisiert	Implizit im Modell enthalten als „Aufgabenanalyse“ in der Planungsphase, die zu metakognitiven Analysen und Strategiewahlen führt und der Mobilisierung motivationaler Überzeugungen dient	Als kombinierter Bestandteil mit dem Kontext, vor allem in der Wahrnehmung, Überwachung, Veränderung und Evaluation der Aufgabenanforderungen über alle Phasen enthalten	Aufgabe ist externe Bedingung, die interpretiert und mit eigenen Wissensbeständen und Überzeugungen abgeglichen wird, sie ist damit Auslöser interner Analysen, welche zur (mehr oder minder bewussten) Wahl eines der Pfade führen	Als Bestandteil der externalen Bedingungen und vor allem als durch metakognitive, motivationale und affektive interne Merkmale beeinflusste Interpretation der Anforderungen Aufgabenverständnis dynamisch modelliert
						Als eigenständiger Bestandteil, der beide Ebenen des Modells beeinflusst Wahrnehmung der Aufgabenanforderungen als Bedingung postuliert, wobei automatische und analytische Zugänge je nach metakognitivem Wissen unterschieden werden
						Als Bestandteil der externalen Bedingungen und vor allem als durch metakognitive, motivationale und affektive interne Merkmale beeinflusste Interpretation der Anforderungen Aufgabenverständnis dynamisch modelliert
						Aufgabenrepräsentation in Phase 1 auf Individualebene und auf Ebene der Gruppe verortet, wo sie durch verbale Austauschungsprozesse abgestimmt werden, um Zielerreichung auf beiden Ebenen zu ermöglichen

Vergleichsdimension	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Modelle Bockaerts' Zweiprocess-Modell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides	Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al.	Hadwins et al. Modell der Selbstregulation in kooperativen Settings
B3) Ziele	• Zieldefinition als eigene Phase im Modell (Phase 2), in der eigene Ziele formiert werden, die nicht notwendigerweise (vollumfänglich) der extern gestellten Aufgabe entsprechen • Ziele dienen der Auswahl von Strategien	Impliziter Bestandteil der Metakognition im Modell, nicht ausgeführt im Modell	Setzen von Zielen als Bestandteil der Planungsphase zur Klärung der Ergebnisse (Ziele setzen) und der Auswahl adäquater Strategien (strategisches Planen)	Zieldefinition in der Phase 1 als Teil der Planung in Kombination mit der Auswahl passender Strategien (Aktivieren metakognitiven Wissens)	Ziele formieren sich aus den Aufgabenmerkmalen, der Einschätzung der eigenen kognitiven Leistungsfähigkeit (Domänen- und Strategiewissen) und den motivationalen Überzeugungen, wobei die Kongruenz mit eigenen internen Zielstrukturen den Pfad beim Vorgehen bestimmt, nämlich zum Lernen des Gegenstands (Top-down-Ziele) oder zum Schutz des eigenen Wohlbefindens (Bottom-up-Ziele)	Spezifische Zieldefinition als Teil der Phase 1 (Aufgabenrepräsentation), Strategieauswahl und kalibrierte Zielformierung von metakognitivem Wissen und Automatisierung abhängig	Zieldefinition als eigene Phase (Phase 2), Ziele – unterschieden in epistemische (zum Wissenserwerb) und weitere Ziele – als Profil von Standards für späteren metakognitiven Vergleich, kombiniert mit Auswahl von Strategien	• (Individuelle und überindividuelle) Ziele sowohl auf der Individual- als auch auf der Gruppenebene lokalisiert, die sich wechselseitig beeinflussen und über Aushandlungsprozesse und soziale Regulation abgestimmt werden • Funktionen der Ziele analog zu COPES

Vergleichsdimension	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Modelle Bockaerts' Zweiphasen-Modell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides	Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al.	Hadwins et al. Modell der Selbstregulation in kooperativen Settings
C) Phasen und Komponenten im Modell								
C1a) Generelle Rolle von Phasen	• Drei dynamische Phasen (1–3), eine fakultative Phase (4) • Phasen sind als Abfolge deklariert, aber interagieren stark, evolvierte Ergebnisse (Produkte) aus vorherigen Phasen bilden Ausgangspunkt für Evaluationen und fortschreitende Selbstregulation im Lernprozess	• Phasen im Modell nicht als distinkte Abfolge definiert, sondern eher implizit angedeutet • Modell enthält fünf Varianten des Zusammenspiels von verschiedenen Komponenten; mit zunehmender Genese der umfassenden Selbstregulation höhere Anzahl von Komponenten und Rückkopplungen	• Drei zeitlich geordnete Phasen mit einzelnen Prozessen und Komponenten • Phasen werden als Zyklus durchlaufen • Selbstregulationszyklen werden als Dynamik für den längerfristigen Erwerb von kalibrierten Selbstregulationseigenschaften modelliert	• Vier zeitlich geordnete Phasen • Phasen als idealtypische Sequenz, aber kein Anspruch auf Notwendigkeit aller Phasen – Phasen explizit als „Heuristik“ deklariert • Pro Phase vier Bezugspunkte der Regulation: Kognition, Motivation/Affekt, Verhalten, Kontext	• Drei implizite Phasen im Modell, die allerdings nicht als Phasen deklariert werden und als dynamisch rekursive Sequenzen angelegt sind • Phasen vor allem zeichnet, welche prototypischen Verläufe im Sinne der beiden Pfade (und dynamischen Wechsel zwischen Pfaden) stattdessen finden	• Drei Phasen postuliert, diese allerdings eher als Kontinuum denn als distinkte Phasen angenommen • Phasen werden unter (Meta-) Kognitionen subsumiert, allerdings mit affektiven Zuständen im Modell deutlich postuliert • Phasen vor allem danach analysiert, welche metakognitiven Erfahrungen vorkommen	• Vier Phasen im Modell postuliert, Phasen als rekursiv und nur wenig in ihrer Sequenz vorgegeben • Komponenten im Modell, die Phasen zugeordnet wurden, sind in den Phasen jedoch reziprok und dynamisch vernetzt	• Wegen Übernahme von COPES- Architektur auf der Individualebene ähnliche Funktionsweise • Auf Ebene der Gruppen prinzipiell ähnliche Funktionsweise der Phasen postuliert • Phasen mit Prozessen und Produkte auf Gruppenebene sind Bedingungen auf Individualebene und umgekehrt

Vergleichsdimensionen	Modelle					
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Bockaerts' Zweiphasen-Modell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides
C1b) Konkrete Phasen	• Phase 1: Aufgabendeinition • Phase 2: Ziele und Plan/Pläne • Phase 3: Durchführung • Phase 4: Adaptationen	• Phase 1: Aufgabenanalyse und Auswahl von Strategien • Phase 2: metakognitiv überwachte Strategienutzung und Leistungserbringung • Phase 3: internes Feedback zur erbrachten Leistung	• Phase 1: Planungsphase • Phase 2: Durchführungsphase/volitionale Kontrolle • Phase 3: Selbstreflexionsphase	• Phase 1: Planungs- und Aktivierungsphase • Phase 2: Überwachen • Phase 3: Kontrollen • Phase 4: Reaktion und Reflexion	• Implizite Phase 1: Aufgabenwahrnehmung und Zielformierung • Implizite Phase 2: Ziele anstreben • Implizite Phase 3: leistungsbezogene Rückmeldung	• Phase 1: Aufgabenrepräsentation • Phase 2: kognitive Verarbeitung • Phase 3: Leistungsphase (und Ergebnisevaluation)
					Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al. • Phase 1: Aufgabendefinition • Phase 2: Planen und Ziele setzen • Phase 3: Durchführung • Phase 4: Evaluation	Hadwins et al. Modell der Selbstregulation in kooperativen Settings • Phase 1: Aufgabendefinition • Phase 2: Ziele und Plan/Pläne • Phase 3: Durchführung • Phase 4: kleine und große Adaptationen

Vergleichsdimension	Modelle					
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Bockaerts' Zweiphasenmodell der Selbstregulation	Metakognitiv-affektives Modell des selbstregulierten Lernens von Efklides
C2) Motivation	Bestandteil der internalen Bedingungen einer Person und Grundlage bei Evaluationen der Prozesse und Produkte über vorgängige und nachgängige Einschätzungen	- Bestandteil erst dann im Modell, wenn kalibrierte metakognitive Prozesse und Wissen erworben sind - Motivationale Variablen (vor allem Attributionen, Selbstwirksamkeit, intrinsische Motivation, Zielorientierungen) als Folge von internalen Feedbackschlaufen	Bestandteil von Phase 1 (Selbstwirksamkeit, Ergebniswartungen, intrinsischer Wert, Zielorientierungen) und Phase 3 (Kausalattributionen)	Eigener Bereich der Regulation über alle vier Phasen hinweg mit mehreren Konstrukten (Zielorientierungen, Selbstwirksamkeit, Aufgabenwert, Interesse – jeweils in Phase 1 – und in Phase 4: Kausalattributionen)	- Als „motivationale Überzeugungen“ im Modell primär dem Wohlbefindenspfad zugeordnet und mit Emotionen verquickt - In der Publikation von 2006 intrinsische Motivation, Interesse, Erwartungs-x-Wert-Kognitionen explizit angeführt	Bestandteil der Personenebene (Zielorientierungen, Erwartungs-mal-Wert-Kognitionen, Kausalattributionen, Selbstkonzept, Kontrollüberzeugungen) im Sinne einer Bedingung, wird allerdings nicht deutlich fokussiert
						Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al.
						Bestandteil im Modell als Bedingung, in der 2007er Publikation einzelne motivationale Konstrukte (Selbstwirksamkeit, Zielorientierungen, Interesse und Aufgabenwert) aufgeführt, aber nicht weiter im Modell spezifiziert
						Als bedeutsam betont, aber nicht ausgeführt
						Hadwins et al. Modell der Selbstregulation in kooperativen Settings

Vergleichsdimensionen	Modelle					Hadwins et al. Modell der Selbstregulation in kooperativen Settings
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Bockeaerts' Zweiphasen-Modell der Selbstregulation	
C3) Emotionen und Affekte	Im Originalmodell vernachlässigt	Im Modell vernachlässigt, aber als ein Merkmal des Idealtypus' des guten Informationsverarbeiters benannt	Wenig explizierter Bestandteil von Phase 3 (Einschätzung der Zufriedenheit und Affekte)	Kombiniert mit motivationalen Merkmalen, expliziert in den Phasen 2-4, angedeutet in Phase 1 am Beispiel Angst	Dominanter Bestandteil im Modell in der Darstellung aus dem Jahr 2011, hierbei vor allem in den auf das Selbst fokussierten Aktivitäten und als Teil der Emotionsregulation beim Lernen, wenn aus der Umwelt Hinweise bestehen, die Bedürfnisse, Ziele und Aspirationen gefährden, um dann eine emotional stabilisierende Regulation vorzunehmen (auf dem „Wohlbefindenspfad“)	Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al. • Epistemische Emotionen explizit enthalten und als Ergebnis des Zusammenspiels von Passung von kognitiven Informationen und existierenden Wissensbeständen und Überzeugungen bzw. neu verarbeiteten Informationen in Phase 1 beschrieben • Direkte Zusammenhänge zwischen epistemischen Emotionen und weiteren Bestandteilen, sei es über Vorläufer oder Folgen von Emotionen

Vergleichsdimension	Modelle					Integratives Modell der epistemischen Kognitionen und des selbstregulierten Lernens von Muis et al.	Hadwins et al. Modell der Selbstregulation in kooperativen Settings
	Winnes und Hadwins COPES-Modell	Prozessorientiertes Modell der Metakognition	Zyklisches Modell der Selbstregulation von Zimmerman	Pintrichs Rahmenmodell selbstregulierten Lernens	Boeckerts' Zweifelsmodell der Selbstregulation		
C4) Metakognitionen	Zentrales Element des Modells, da Überwachen und Kontrolle als metakognitive Regulationsprozesse dazu dienen, die internen Feedback-Schleifen über die fünf namensgebenden Bestandteile (Conditions, Operations, Products, Evaluations, Standards) metakognitiv zu regulieren	- Im Modell als „exekutive Prozesse“ deklariert – mit doppelter Funktion: metakognitive Planung und Analyse sowie Regulation - Zentrales Element des Modells, auf dessen Grundlage das deklarative und exekutive Strategiewissen zusammenwirken	Explizites Element in der Phase 1 bei der Zielsetzung und Strategieauswahl, in Phase 2 der Durchführung und volitionalen Kontrolle und Teil der Phase 3 der Selbstreflexion bei den Vergleichen der Ergebnisse mit diversen Evaluationsstandards	Kombiniert mit weiteren kognitiven Bestandteilen eines eigenen Bereichs der Selbstregulation in allen vier Phasen, aber auch essenzieller Bestandteil der überwachenden (Phase 2) und kontrollierenden prozeduralen Metakognition (Phase 3), daher metakognitive Regulation von Kognitionen, Motivation/Affekt, Verhalten und Kontext explizit enthalten	Bestandteil in der Wahrnehmung der Aufgabe und Passung mit eigenen Zielen, die als Gradmesser für das weitere Vorgehen fungieren Metakognitiv gesteuerte Strategieauswahl und Vorgehensweise sind Bestandteil des „Wachstumspfads“, setzen allerdings voraus, dass interne Ziele und Aufgabe kongruent sind, sodass die nicht explizierten metakognitiven Prozessen optimal genutzt werden	- Durchgängiger Bestandteil im Modell, der in allen Phasen Bestandteil ist und vor allem in den Phasen 1, 3 und 4 zentral wirkt - Deklaratives und konditionales Strategiewissen (in Phase 1) als Basis für Strategieauswahl in Phase 2, in Phase 4 als metakognitive Überwachung und Regulation	- Auf Individual-ebene ähnliche Prämissen wie bei COPES – hier sind vor allem Metakognitionen über eigene Person und Aufgabenerfordernisse angesiedelt - Metakognitionen auf Gruppenebene beziehen sich auf die Wahrnehmung der Spezifika der Gruppe und ihrer Merkmale, die sich direkt oder indirekt als metakognitive Regulationsmechanismen auf das individuelle Vorgehen auswirken, sodass es zu simultaner und reziproker metakognitiver Regulation innerhalb und zwischen den Gruppenmitgliedern kommt

(eigene Darstellung, basierend auf und teils Inhalte übernehmend aus Li & Lajoie, 2022; Panadero, 2017; Puustinen & Pulkkinen, 2001, sowie Rouet et al., 2017)

6.1.1.3 Analyse: modellübergreifende Vergleiche

Vergleich der theoretischen Verortung: welche Differenzen sich in den Zielen und Geltungsbereichen aus der theoretischen Provenienz ergeben

Die Selbstregulationsforschung speist sich aus verschiedenen Theorieströmungen. Eine grundsätzliche Schwierigkeit bei der klaren Zuteilung von einzelnen Modellen zu einer Kategorie von Modellen besteht deshalb darin, dass sich neuere Modelle dezidiert als integrative Modellierungen verstehen. Da Selbstregulation per se als mindestens ohnehin kognitiver, metakognitiver sowie emotional-affektiver Prozessverbund modelliert wird und sich vorhandene Modelle bereits in aller Regel mehrerer Theorien bedienen, ist eine exklusive Situierung in einzelnen Theoriefamilien nicht ohne weiteres möglich. Integrative Modellierungen gehen hierbei noch einen Schritt weiter, sodass eindeutige Verortungen mindestens erschwert sind. Dies ist eine nötige Vorbemerkung, auch um auf die inhärente mangelnde Trennschärfe eingangs deutlich hingewiesen zu haben.

Die Modelle haben eigene Herkünfte, die nicht immer vollständig explizit gemacht werden, was beispielsweise Puustinen und Pulkkinen (2001) zu einer nur sehr knappen Bestimmung der Ähnlichkeit im direkten Modellvergleich untereinander bewogen hat. Sie nutzten eine vierstufige Skala zur Explikation der Ähnlichkeit, machen aber nicht explizit, woran sich (Un-)Ähnlichkeit bemisst. Gesamthaft gelangen sie zur Einsicht, dass die fünf von ihnen analysierten Modelle, die auch in diesem Teilkapitel fokussiert werden, einer von zwei idealtypischen Kategorien zuordenbar sind. Sie nennen die erste Kategorie „zielorientierte Definitionen“ selbstregulierten Lernens und die zweite Kategorie „metakognitiv gewichtete Definitionen“.

- Die *erste, zielorientierte Definition* fokussiert darauf, dass persönliche Ziele das selbstregulatorische Geschehen beeinflussen, und diese Modelle weisen als zweites Merkmal eine Breite an nicht-metakognitiven Komponenten auf, darunter motivationale, affektive, behaviorale und soziale Merkmale. Puustinen und Pulkkinen (2001) subsumieren dieser Gruppe die Modelle von Boekaerts (2011), Pintrich (2000) und Zimmerman (2000). Mit seiner Perspektive auf soziale Prozesse, die Selbstregulationen rahmen und beeinflussen, ist das von Hadwin et al. (2018) vorgelegte Modell ebenfalls Vertreter dieser Konzeptualisierungen des selbstregulierten Lernens. Selbiges trifft auf das integrative Modell zu, welches Muis et al. (2018) eingeführt haben.
- Demgegenüber sind Modelle mit *metakognitiver Gewichtung* enger gefasst: Sie akzentuieren die Steuerung und Überwachung von kognitiven Strategien zum Zweck des Lernens. Solche Modelle sind also kognitivistisch in ihrem Zugang und akzentuieren vor allem die metakognitive Selbstregulation. Puustinen und Pulkkinen (2001) haben hier ursprünglich die Modelle von Borkowski

et al. (2000) sowie Winne und Hadwin (1998) dieser Kategorie zugeordnet. Als drittes Modell ist Efklides' (2011) Modell mit seiner stark metakognitiven Ausrichtung zu ergänzen.

Mit dieser groben Zuteilung zu einer von zwei Gruppen in den Modellen wird nicht nur eine theoretische Verortung erleichtert, sondern auch zugleich der Anspruchsgrad deutlicher: Die eher eng gefassten Modelle mit metakognitiver Gewichtung sind stärker fokussiert auf das Zusammenwirken primär kognitiver Prozesse und fallen in dieser Hinsicht in ihrem höheren Detaillierungsgrad bei den metakognitiven Regulationsprozessen auf. Diese Prozesse werden in den drei Modellen um zwei Bestandteile gruppiert: erstens die metakognitive Steuerung und Überwachung von Prozessen, die sich auf der Objektebene des Gegenstands vollziehen, und zweitens in ihrem Zusammenspiel mit Aufgabenmerkmalen und sich entwickelndem Aufgabenprodukt, die fortlaufend auf die Adäquanz metakognitiv beurteilt werden. Es sind also die internalen Kognitionen in ihrer Interaktion mit dem entstehenden kognitiven Produkt, die solche Modelle präzise in den Blick nehmen.

Die andere Gruppe mit breiter Zielorientierung und weitere individuelle Merkmale sowie teils den Kontext erheblich weiter berücksichtigenden Modelle zeichnen sich dadurch aus, dass sie diese Bestandteile nicht nur integrieren, sondern auch das prozessuale Zusammenwirken dieser Bestandteile erklären müssen. Dieser breitere Geltungsbereich bringt es mit sich, dass solche Modellierungen metatheoretischen Charakter haben, wenn sie nicht wie Zimmermans (2000) Modell relativ eindeutig einer Theoriefamilie zuzurechnen sind.

Vergleich struktureller Bestandteile: welche Komponenten die Modelle enthalten und wie sie diese gewichten

Die dichotome und damit schematische Gegenüberstellung der theoretischen Provenienz der Modelle im Abschnitt zuvor könnte auf den ersten Blick als eine homogenisierende Aussage gewertet werden. Tatsächlich aber zeigt sich in den Modellen eine Heterogenität, die sich über die Zwecke der Modelle und deren theoretischer Verortung erklären lässt. Panadero (2017) hat diese Heterogenität danach systematisiert, inwiefern die Modelle die drei Bestandteile (Meta-) Kognitionen, Motivation und Emotion gewichten, sodass sich Modelle anhand dieser Schwerpunkte profilieren lassen. Einer etwas anders gelagerten Logik folgt die sich anschließende Darstellung in diesem Kapitel, indem sie Zentrum und Peripherie der Modelle danach bestimmt, worauf sie fokussieren und die Modelle nicht danach reiht, wo sie jeweils besondere Akzente setzen. Sie greift aber Panaderos (2017) Grundidee auf, indem verschiedene Bereiche der Profilierung nacheinander vom Zentrum hin zur Peripherie über alle Modelle hinweg behandelt werden.

- *Die Modelle legen gesamthaft einen Fokus auf Metakognition:* Metakognitive Prozesse kommen in allen Modellen vor, selbst wenn die Modelle deren Komponenten unterschiedlich stark akzentuieren. Einige Modelle wie jene von Borkowski et al. (2000), Efklides (2011), Hadwin et al. (2018) sowie Winne und Hadwin (1998) betonen deutlich Metakognitionen, andere – insbesondere das von Boekaerts (2011) entwickelte Modell – nivellieren diesen Bereich eher. Die übrigen Modelle (Muis et al., 2018; Pintrich, 2000; Zimmerman, 2000) enthalten Metakognitionen als einen Bestandteil unter mehreren. Diese divergierende Gewichtung ergibt sich partiell aus der theoretischen Herkunft, aber auch daraus, dass die Modelle jenseits der deutlichen Fokussierung auf Metakognition Akzente bei den motivational-affektiven Bestandteilen setzen. Es gibt in diesen Modellen mithin jeweils mindestens einen weiteren Fokus, der sich auf die Regulation von Motivation (Boekaerts, 2006; Pintrich, 2000; Zimmerman, 2000), Emotionen (Boekaerts, 2011; Muis et al., 2018) und teils Verhalten (Pintrich, 2000) erstreckt. Insofern sind die Schwerpunkte im Bereich der Metakognition dadurch zu erklären, dass die Modelle verschiedene Bereiche der Regulation systematisieren, doch selbst dort, wo die Metakognition wenig akzentuiert zu sein scheint, nämlich im Modell von Boekaerts (2011) kommt sie in Form eines eigenen Pfades, nämlich der metakognitiv gesteuerten und lernstrategischen inhaltsorientierten Auseinandersetzung, explizit vor. Metakognitive Prozesse der Steuerung und Überwachung sind ebenfalls in den Modellen als dynamisch interagierende Bestandteile enthalten. Die metakognitiv repräsentierten, aus der Aufgabe extrahierten und mit eigenen Wissensbeständen angereicherten Ziele auf Individualebene fungieren als Standards bei der kontinuierlichen Fortschrittsüberwachung und finaler Gesamteinschätzung, sei es in puncto Prozesse, sei es bei den (entstehenden) Produkten. Hinzu kommen die metakognitiven Prozesse der Zielformulierung in den frühen Phasen des Geschehens in den Modellen: Hier dienen sie dazu, dass Personen adäquate Strategien auswählen, um das jeweilige Ziel der aktuellen Aufgabe zu erreichen. Prozesse der Ziel-Mittel-Analyse und Überwachung des Fortschritts sind demnach in den Modellen mehr oder minder deutliche Konstituenten (Stine-Morrow et al., 2006). Die metakognitiv repräsentierten Ziele können bei alldem auch deutliche Abweichungen von dem in der Aufgabe explizit formulierten Anforderungen haben. Winne und Hadwin (1998) machen darauf aufmerksam, dass Personen eigene Ziele verfolgen; Boekaerts (2011) macht sogar die (In-)Kongruenz mit eigenen motivationalen Orientierungen zum Gradmesser dessen, ob Personen external formulierte Ziele aufgreifen. Und bei Hadwin et al. (2018) sind Ziele sowohl auf der Individual- und der Gruppenebene abstimmungsbedürftig, wobei divergierende Individualziele dezidiert möglich sind. Dies wird durch die in den Modellen aufgeführten individuellen Bedingungen und teils über

den unterschiedlich systematisch einbezogenen Kontexteffekt nochmals erheblich dynamisiert.

- *Motivation ist in fast allen Modellen integraler Bestandteil, der Stellenwert und die Art der motivationalen Konstrukte variieren jedoch:* Motivation ist ein in diversen Modellen vorkommender Bestandteil (mit Ausnahme von Winne & Hadwin, 1998, und darauf aufbauend Hadwin et al., 2018). Hadwin et al. (2018) betonen wiederholt die Motivation als essenziell für die Selbstregulation auf Individual- und sozialer Ebene, aber sie bleiben unspezifisch, wie die Motivation in ihrem Modell wirken soll. Auffällig ist: Motivation als auch in der Leseforschung häufig vorkommender Sammelbegriff (Conradi et al., 2014) wird in den Modellen über mehrere Konstrukte abgedeckt, welche allgemein eine dominante Stellung in der Motivationspsychologie einnehmen (Cook & Artino, 2016; Hattie et al., 2020).

Diese motivationalen Konstrukte sind einerseits stabile Merkmale auf Personenebene wie das *Interesse* als eine positive Orientierung zu einem Thema oder Gegenstand oder – damit verwandt – die *gegenstandsspezifische intrinsische Motivation* als mit dem Sachgegenstand verbundener Anreiz der Auseinandersetzung mit ihm (Conradi et al., 2014). Diese beiden Konstrukte kommen in fast allen Modellen vor, die im Abschnitt oben zur theoretischen Verortung als zielbezogene Modelle klassifiziert wurden (Boekaerts, 2006; Muis et al., 2018; Pintrich, 2000; Zimmerman, 2000) sowie zusätzlich in dem Modell von Borkowski et al. (2000) in seiner finalen Stufe.

Ziele der Personen im motivationalen Sinne werden in den Modellen unterschiedlich behandelt. Auffällig ist, dass Boekaerts (2011) bei den Zielen zwei grundsätzliche Arten unterscheidet, wonach Personen streben. Je nachdem, welches Ziel priorisiert wird, ist auch der eingeschlagene Pfad zuordenbar, der „Wohlbefindenspfad“ und der „Wachstumspfad“. Im ersten Fall ist das handlungsleitende (implizite) Ziel, das eigene Wohlbefinden zu maximieren und Schäden daran zu vermeiden, was sich in den entsprechenden Verhaltensweisen manifestiert, die dazu dienen, dieses Ziel zu erreichen. Im letztgenannten Fall des Wachstumspfad steht das Ziel im Vordergrund, seine eigenen Fähigkeiten und Kompetenzen sowie sein Wissen zu erweitern. Dieser zweite Zugang weist eine Ähnlichkeit zu dem auf, was in den übrigen Modellen unter einem Teil des Konstrukts günstiger *Zielorientierungen* subsumiert wird (Borkowski et al., 2000; Efklides, 2011; Muis et al., 2018; Pintrich, 2000; Zimmerman, 2000). Zielorientierungen sind im allgemeinem Verständnis stabile Absichten bei Aufgabenbewältigungen, die sich entweder auf die Steigerung der eigenen Fähigkeiten beziehen, was an den Aufgabenerfordernissen oder an dem eigenen Leistungsvermögen gemessen wird (Lernorientierung), oder das Abschneiden bei Leistungen im Vergleich zu anderen (Performanzorientierung) betreffen, sich also an externalen Beurteilungskriterien bemessen (Urđan & Kaplan, 2020).

Zwei weitere motivationale Variablen sind erwartungs-x-wert-theoretischer Herkunft, nämlich die Wert-Kognitionen und die Ergebniserwartungen. Dabei überwiegt die *Wert*-Komponente in den Modellen (Boekaerts, 2006; Efklides, 2011; Muis et al., 2018; Pintrich, 2000; Zimmerman, 2000), während die *Erwartungskomponente* in weniger Modellen aufgeführt worden ist (Boekaerts, 2006; Efklides, 2011; Zimmerman, 2000). Erwartungs-x-Wert-Theorien zielen darauf ab, Verhalten auf der Basis von mehrdimensionalen Einschätzungen zum Wert einer inhaltlichen Auseinandersetzung mit einem Gegenstand und zur subjektiv eingeschätzten Erfolgseinschätzung zu erklären (Eccles & Wigfield, 2020). Je stärker eine Aktivität eine Gratifikation verspricht (Wertkomponente) und je größer die Aussicht auf Erfolg ist (Erwartungskomponente), desto wahrscheinlicher ist es, dass eine Person ihr nachgeht. Wichtig ist: Wert und Erwartung beziehen sich auf eine unmittelbar bevorstehende Aktivität, sie sind damit situativ zu modellieren.

Theoretisch verwandt, aber nicht deckungsgleich mit der Erwartungskomponente ist die *Selbstwirksamkeit*. Dieses Konstrukt bezeichnet die kognitive Selbsteinschätzung, in einer definierten Aufgabe, die unmittelbar bevorsteht, erfolgreich oder nicht erfolgreich zu sein (Schunk & DiBenedetto, 2020). Diese Komponente spielt in den Modellen der Selbstregulation eine prominente Rolle, die aus sozial-kognitiven Theorien schöpfen (Borkowski et al., 2000; Muis et al., 2018; Pintrich, 2000; Zimmerman, 2000). Selbstwirksamkeit gilt als Schlüsselvariable in sozial-kognitiven Theorien, weil sie als idealerweise kalibriertes metakognitives Urteil dabei hilft, realistische Erfolgsaussichten zu erkennen und sich selbst zu mobilisieren.

Eine andere Art von motivationalem Konstrukt, welches häufig in den Modellen vorkommt (Borkowski et al., 2000; Efklides, 2011; Pintrich, 2000; Zimmerman, 2000), betrifft den Umgang mit meist erwartungswidrigem Misserfolg. Es handelt sich um *Kausalattributionen* als individuelle Erklärungsansätze für Ergebnisse, die sich darin unterscheiden, ob der Grund dafür innerhalb oder außerhalb der eigenen Person liegt, wie stabil diese Faktoren sind und ob sie beeinflussbar wirken (Sa. Graham, 2020). Dabei gibt es günstigere und ungünstigere Konstellationen bei der Erklärung, die zum einen handlungsleitend für spätere ähnliche Situationen sind und zum anderen auch in akademische Emotionen münden.

Der Vollständigkeit halber sei schließlich noch erwähnt, dass einzelne motivationale Konstrukte nur vereinzelt in den Selbstregulationsmodellen auftauchen: Efklides (2011) benennt zwei Konstrukte auf der Personenebene: das *Selbstkonzept* als mehrdimensionale, die eigene Person zeitstabil betreffende Überzeugung zu domänenspezifischen Fähigkeiten inkl. affektiver Merkmale, und die Kontrollüberzeugungen, was in ihrer Darstellung als *Agency* bezeichnet werden kann, also das Ausmaß, in dem Personen sich als fähig erachten, Folgen ihres Handelns selbstständig herbeizuführen (Conradi et al., 2014).

Die Prävalenz einiger motivationaler Komponenten ist gemäß den bisherigen Ausführungen also deutlich. Hinzu kommt, dass diese Konstrukte einer *zeitlichen Perspektive* unterliegen. Die Modelle, in denen die Motivation explizit vorkommt, trennen motivationale Konstrukte in einer prä- und einer postaktionalen Phase. Präaktional, also vor dem Lernen zu lokalisierende Konstrukte sind: Selbstwirksamkeitsüberzeugungen, Erwartungs-x-Wertüberzeugungen und das Verhalten im Optimalfall auf tiefenorientierte Prozessierung energetisierende Variablen wie Selbstkonzept, Kontrollüberzeugungen, Interesse/intrinsische Motivation und Zielorientierungen. Postaktional, mithin nach dem Lernen zu verorten, sind die Kausalattributionen. Deren Ergebnis des Erklärens von (Miss-)Erfolg bildet den Ausgangspunkt für metakognitiv verarbeitete Prozessergebnisse, aus denen sich andere Komponenten speisen, etwa die Meisterschaftserfahrungen als wichtige Quelle für die Selbstwirksamkeit (Schunk & DiBenedetto, 2020).

- *Emotionen werden in den wenigen Modellen, in denen sie expliziter Gegenstand sind, vor allem als Folgen des Aufgabenfortschritts konzeptualisiert und haben aktivierende und deaktivierende Wirkung auf Kognition und Motivation.* Die Konzeptualisierungen und der Fokus in den Modellen ist recht ähnlich, nämlich bei den akademischen Emotionen, die sich als mehrere Prozesse umfassende Konstrukte auf Tätigkeiten oder Ergebnisse des Lernens bzw. Leistens beziehen und aus einem Abgleich von neuen Informationen und vorherigen Informationen resultieren (Muis et al., 2018).

Die Modelle divergieren in ihrem Fokus, wie sie diese Emotionen konzeptualisieren. Efklides (2011) beschreibt in ihrem Modell Emotionen als Konsequenz eines Abgleichs von Ist-Zuständen durch internale Feedbackschleifen bei der Fortschrittsüberwachung, welche zu angenehmen, neutralen oder unangenehmen affektiven Zuständen führen. Sie bringt Emotion und Motivation eng in Verbindung, trennt diese aber nicht sehr deutlich voneinander und bleibt hinsichtlich der Rolle von Emotionen recht vage. Boekaerts' (2011) Zugang in der theoretischen Modellierung bezieht sich auf den Umgang mit negativen Emotionen, welche sich daraus ergeben, dass jemand Hindernisse und Brüche dabei feststellt, wenn es um die eigenen Ziele und die Aufgabenstellungen geht. Damit fokussiert sie auf den „Wohlbefindenspfad“, der in Vorgängermodellen als „Coping-Pfad“ bezeichnet wurde, also als Bewältigung schwieriger Situationen (Boekaerts & Niemivirta, 2000). Diese Hindernisse haben eine potenziell negative Wirkung, die aktivierend oder deaktivierend ausfällt. Selbstregulation ist ihrem Verständnis zufolge die Fähigkeit, solche Emotionen unbewusst oder bewusst erfolgreich zu regulieren. Diese Fähigkeit ist auf eine der beiden Routen in ihrem Modell bezogen und damit auf einen Bereich, den andere Modelle so nicht enthalten. Den systematischsten Zugang zu den Emotionen wählen Muis et al. (2018). Sie beschreiben fünf Qualitäten von Vorläufern von Emotionen, die jederzeit im Verlauf von Selbstregulationsphasen zum Tragen

kommen können und das Geschehen beeinflussen. Diese erste Perspektive richtet sich darauf, was welche Emotionen hervorruft und wie beeinflusst. Die zweite Perspektive geht auf die Folgen von Emotionen ein, also wie diese Emotionen das Planen und Zielsetzen, die Strategienutzung, die Motivation, die Ergebnisse und schließlich auch diverse Überzeugungen beeinflussen.

- *Die Regulation von Verhalten und Kontext als eigene Bereiche der Selbstregulation ist randständig.* In der Tabelle 31 ist das Verhalten keine eigene Vergleichsdimension und stand also nicht im Fokus der Analyse. Jedoch ist auffällig, dass eine explizite Thematisierung des Verhaltens als eigener Bereich der Selbstregulation in den Modellen selbst eher selten ist. Pintrich (2000) spekuliert, dass es der Fokus auf internale Prozesse ist, der zu der Randständigkeit des beobachtbaren Agierens führt, und entsprechend ist sein ohnehin weit gefasstes Rahmenmodell eine Ausnahme. Er inkludiert damit das von außen beobachtbare Verhalten wie Zeitplanung, Festhalten von Erfolgen im Sinne einer Fortschrittsüberwachung mit Buchführungscharakter, die Regulation der Anstrengungsbereitschaft mittels motivierender Selbstgespräche, dem Hilfe-Suchen etc. Solche Verhaltensweisen werden in Taxonomien von Lernstrategien häufig als „Stützstrategien“ bezeichnet (Philipp, 2015). Triadische Selbstregulationsmodelle in der Tradition des sozial-kognitiven Lernens haben den Wert einer solchen verhaltensbezogenen Ressource deutlich gemacht (Schunk & DiBenedetto, 2020; Zimmerman, 2013). Weiterhin ist Pintrichs (2000) Modell wegen seiner inhärenten Breite dasjenige Modell, das bereits auf Individualebene berücksichtigt, dass Personen durch ihr Verhalten gezielt den Kontext verändern. Dies lässt sich über die im Absatz zuvor benannten triadischen Modelle der Selbstregulation abstützen, in denen neben internal-verdeckten Prozessen und von außen beobachtbarem Verhalten auch durch das eigene Verhalten hervorgerufene Modifikationen im Kontext der Zielerreichung dienen. Die Veränderung des Kontexts ist weiterhin im Modell von Hadwin et al. (2018) inkludiert und wird dort als Rückkopplung von Verhaltensweisen und Kommunikation auf die soziale Ebene der anderen Personen modelliert, wo sie die Bedingungen auf beiden Ebenen dynamisierend und reziprok beeinflussen.

Aus dieser Randständigkeit von Verhalten und dessen Auswirkungen auf die Umwelt wird implizit deutlich, dass die bislang vorgelegten Modelle zuallererst das Individuum und dessen internen regulatorischen Prozesse ins Zentrum stellen. Hier sind aktuelle Tendenzen zu nennen, die dem Kontext, zumal unter der Perspektive des Erwerbs und der Vermittlung selbstregulatorischer Fähigkeiten, erheblich mehr Bedeutung zumessen, als es in den dominierenden Modellen bisher der Fall ist (Ben-Eliyahu & Bernacki, 2015; Butler, 2021; Järvenoja et al., 2015).

Vergleich temporaler Aspekte: welche Funktionen die Phasen im Gesamt der Selbstregulationsmodelle haben

Selbst wenn die Modelle im Konkreten unterschiedliche Architekturen haben und sich unterschiedlicher Bestandteile bedienen, so besteht gesamthaft eine auffällige Übereinstimmung darin, dass drei Phasen voneinander getrennt werden, selbst wenn die darin beteiligten Komponenten und Konstrukte inhaltlich und terminologisch variieren. Diese drei Phasen sind, um mit Puustinen und Pulkkinen (2001) zu sprechen:

- 1) eine *vorgängige Planungsphase* – in dieser Phase werden die Aufgabe (und der Kontext) enkodiert, Ziele formiert und das Vorgehen im Sinne der metakognitiven Planung festgelegt; je nach Modell kommen affektive und motivationale Mobilisierungen hinzu;
- 2) eine *Leistungsphase* – in ihr findet die inhaltliche Auseinandersetzung mit dem Gegenstand statt, der über automatisiert ablaufende Prozesse und lernstrategische Vorgehensweisen inklusive ihrer metakognitiven Überwachung prozessiert wird, die Modelle ähneln sich darin, dass Kognitionen und Metakognitionen in dieser Phase den Hauptgegenstand bilden;
- 3) eine *abschließende Beurteilungsphase* – in dieser finalen Phase werden Prozesse und/oder Produkte von der lernenden Person evaluiert, in breiter gefassten Modellen kommt es zu motivationalen Einschätzungen und emotionalen Reaktionen; diese Reflexionen und Reaktionen bilden in den zyklischen Modellen den Ausgangspunkt weiterer Zyklen. Nur in einem Modell fehlt diese Phase (Efklides, 2011).

Die Modelle variieren darin, ob sie Aufgabenrepräsentation und Zielformulierung als eine einzelne Phase annehmen oder diese beiden Vorgänge als separate und distinkte Phasen modellieren. Letztgenannter Fall trifft auf Winnes und Hadwins (1998) Modellierung zu, die diese Trennung vornehmen, um damit die eigenaktive Zielsetzung in ihrer Phase 2 von der Wahrnehmung der Aufgabenerfordernisse zu separieren und die Umdeutung von Zielen betonen. Andere Modelle kombinieren dies in einer eigenen ersten Phase.

Trotz einer mehrheitlich großen Übereinstimmung im Postulat von Phasen in den Modellen besteht weitestgehend Einigkeit darüber, dass diese Phasen nicht als strikte Sequenz zu denken sind. Vielmehr wirken sie im Licht der einzelnen Modellbeschreibungen als Sammelgefäße von Prozessen mit spezifischen Funktionen, welche sich idealtypisch Zeitpunkten zuordnen lassen, die sich aus der Abfolge einzelner Aktivitäten in einem temporalen Kontinuum beschreiben und analysieren lassen. Nur wenige Modelle wie etwa das von Muis et al. (2018) treffen strikte zeitliche Trennungen, indem sie im konkreten Fall das Zustandekommen von Emotionen als Folge der Informationsverarbeitung in der mittleren

Leistungsphase postulieren. Andere Modelle gehen von ausgesprochen lose verknüpften Phasen aus, die eine hohe Dynamik aufweisen (Winne & Hadwin, 1998). Diese Variabilität und Dynamik lässt sich auch darüber erklären, dass allein die Leistungsphase in ihrer parallelen kognitiven Verarbeitung von Informationen und der Überwachung und etwaigen Veränderung der Kognitionen nicht-lineare Abfolgen distinkter Prozessgruppen aufweist.

Eine weitere Facette aus temporaler Sicht betrifft nicht nur die Phasen im einzelnen Auftreten von Episoden der Selbstregulation bei konkreten singulären und abgeschlossenen Anlässen. Hinsichtlich des längerfristigen Erwerbs von selbstregulatorischen Fähigkeiten sind die Modelle unterschiedlich explizit. In den Modellen besteht neben der Interaktion und Modifikation die grundsätzliche Auffassung, dass Rückkopplungen zum einen den je konkreten Anlass der Aufgabenbewältigung betreffen und dynamisieren. Zum anderen werden über die reflexiven Phasen am Ende in den jeweils letzten Phasen metakognitive Prozesse initiiert, welche auf weitere, zukünftige Selbstreflexionszyklen einwirken. Dies geschieht über sowohl motivationale Prozesse wie (un-)günstige Attribuierungen (Borkowski et al., 2000; Zimmerman, 2000) als auch über die Modifikation der metakognitiven Wissensbestände, indem zum Beispiel das Wissen über adäquate Strategieranwendung oder das Testen von komplexeren Strategien erweitert wird und dann zukünftig als Ressource zur Verfügung steht (Winne & Hadwin, 1998). Diese allgemeinen Mechanismen über durchlaufene Selbstregulationszyklen werden im Allgemeinen angenommen und im Besonderen nur in wenigen Modellen spezifischer entfaltet. Borkowski et al. (2000) gehen aus einer Erwerbsperspektive davon aus, dass eine umfassende Selbstregulation mit verschiedenen motivationalen und kognitiven Bestandteilen einen Zustand darstellt, der erst durch das Absolvieren verschiedener Zwischenstadien erreichbar ist, und sie betonen dabei die hohe Bedeutsamkeit eines Kontexts, in dem selbstregulatorische Fähigkeiten (explizit) vermittelt und eingefordert werden. Die Erwerbslogik im didaktisch ertragreich wirkenden Modell von Borkowski et al. (2000) besteht darin, dass einzelne Strategien zunächst erworben werden, um Handlungsfähigkeit aufzubauen. Es folgt dann ein sukzessives Erweitern des Strategiespektrums, welches zu einem erweiterten konditionalen Strategiewissen und dessen exekutiver Aktivierung im Lernprozess führt. Erst wenn das Strategiewissen ausreichend erworben und konsolidiert ist, wird gemäß diesem Modell die Motivation (hier: Attributionen) als zusätzliche mobilisierende Quelle für gelingende Selbstregulation angenommen. Diese Erwerbslogik ist stark metakognitiv geprägt, sie modelliert den Erwerb konsequent als sich ausweitendes Wissen über die adäquate Strategienutzung.

Modelle in der Tradition sozial-kognitiver Theorien – allen voran das von Zimmerman (2000, 2013) – nehmen ebenfalls Perspektiven auf den Erwerb ein (nicht dargestellt in Tabelle 31). Anders als das Modell von Borkowski et al. (2000) modelliert Zimmerman (2013) den Erwerb selbstregulatorischer Fähigkeiten als vierphasigen Prozess. In ihm sorgen 1) stellvertretendes Erleben

der zu erwerbenden selbstregulatorischen Fähigkeit und 2) die anschließende Verhaltensnachahmung mit sozialer Unterstützung dafür, dass diese Fähigkeiten zunächst internalisiert werden. Erst danach avancieren Selbstregulationsprozesse dadurch, dass sie mit 3) zunehmender Selbstkontrolle ausgeführt werden, zum Gegenstand von eigeninitiierten Modifikationen 4) in der finalen Phase der Selbstregulation. In dieser Modelllogik sind Interaktionen zwischen Personen mit verschiedenen Fähigkeitsausprägungen und Rollen entscheidend bei einem Verlauf von eher external zu lokalisierendem Ursprung der regulatorischen Fähigkeiten hin zu einer echten *Selbstregulation* mit dem Individuum die treibenden Kräfte. Die theoretische Provenienz macht deutlich, dass der soziale Kontext in dieser Modellierung eine zentrale Rolle einnimmt, da Modellpersonen und soziale Quellen für Rückmeldungen von hoher Bedeutung sind und sogar als Bedingungen in dem Erwerbsmodell in Phase 1 und 2 aufgeführt werden.

Mögen sich beide Erwerbsmodelle auf den ersten Blick unterscheiden, so weisen sie doch Gemeinsamkeiten auf: Aus Aneignungslogik stellen sie an den Anfang, dass zunächst einzelne Fähigkeiten von der lernenden Person in begrenztem Maße reproduziert werden. Umfassende Regulationsprozesse stellen hingegen ein Fernziel dar, bei dem zunehmend mehr interagierende Elemente wie Motivation, Strategiewissen und eine Sensitivität für die kontextuellen und Aufgabenmerkmale erkannt, koordiniert und möglichst zielführend in Einklang gebracht werden. Eine zunehmende, auf Erfahrungen in der Selbstregulation beruhende Wahrnehmung der eigenen Kompetenz – welche zum einen als Attribution als auf eigene Anstrengung rückführbare Erklärung für Erfolg gilt und zum anderen als durch kalibriertes Urteil über Standards des Erfolgs und deren Erreichen als Grundlage der Selbstwirksamkeit modelliert wird – bildet die motivationale Klammer (Cook & Artino, 2016).

6.1.1.4 Abschlussbemerkung

Die verschiedenen Modelle selbstregulierten Lernens haben auf den allgemeinen Nenner gebracht ein komplementäres Verhältnis. Dadurch, dass sie unterschiedliche Komponenten in einer mehr oder minder strikten zeitlich trennbaren Phasenlogik kombinieren und dabei komplexe Wechselwirkungen annehmen, sind modellübergreifende Aussagen nur auf abstrakter Ebene leistbar. Es ist deshalb auch immer wieder in der Sekundärliteratur beobachtbar, dass Modellvergleiche über Kontinua von Dimensionen und die Explikation der (Un-)Ähnlichkeit erfolgen (Kaplan, 2008; Panadero, 2017; Puustinen & Pulkkinen, 2001). Die Essenz des Modellvergleichs lässt sich folgendermaßen bündeln:

- *Definitorisches:* Selbstregulation ist ein zielbezogenes Vorgehen, bei dem sich Personen in Bezug auf das Lernen – bzw. für die Zwecke dieses Buches: das Lesen digitaler Dokumente – selbstständig Ziele setzen. Sie nutzen verschiedene

kognitive, motivationale, affektive, behaviorale und teils soziale Prozesse, um diese Ziele adaptiv zu erreichen, und gleichen metakognitiv über Rückkopplungen den Fortschritt ab, den sie selbstständig koordinierend anstreben.

- *Strukturelles*: Die Modelle unterscheiden sich grosso modo weniger darin, dass sie allgemein verschiedene motivational-emotionale, kognitive und (eher randständig) behaviorale und kontextuelle Elemente beinhalten. Sie gewichten, priorisieren und akzentuieren diese allerdings spezifisch, weil sie ihrem Modellschwerpunkt folgend unterschiedliche Phänomene stärker beleuchten. Deshalb ist die zeitliche Ähnlichkeit als Oberflächenmerkmal einzustufen, da die differenzstiftenden Modellmerkmale auf einer Tiefenmerkmalsebene der interagierenden Komponenten lokalisierbar sind. Die essenziellen Unterschiede in den Modellen sind darin zu finden, wie sie das Interagieren von Komponenten modellieren. Metakognitive Prozesse des Planens und Überwachens nebst Regulation der Kognitionen sind mehr oder minder in allen Modellen erkennbare Bestandteile. Bei der Motivation bestehen Unterschiede, welche vor allem präaktionalen motivationalen Konstrukte als Ressource fungieren, und bei den Emotionen ist markant, wie stark sich die Modelle in Detailliertheit unterscheiden. Weitere Areale der Regulation wie Kontext und Verhalten sind in den Modellen selten und teils marginal.
- *Temporales*: Selbstregulation ist ein (phasenweise) ablaufender Prozess, in denen zu verschiedenen Zeitpunkten verschiedene Teilprozesse und Bestandteile interagieren. Auch wenn die Prozesse in einer Sequenz stattfinden, sind Phasen in den Modellen nicht als strikt zu durchlaufende Schrittabfolgen modelliert, sondern interagieren dynamischer. In diesem Punkt gibt es eine deutliche Konvergenz der Modelle, über welche die unterschiedlichen Bezeichnungen der Phasen nicht hinwegtäuschen. Unter diese zeitliche Perspektive fällt auch, dass die Wahrnehmung von Aufgabe und Kontext prototypisch in einer ersten Phase erfolgt und hierbei den Ausgangspunkt für top-down geleitete Analysen und das Aktivieren von metakognitiven Standards bildet. Solcherart präpariert nutzen Lernende in der aktionalen Phase Strategien zur Fortschrittserreichung und regulieren diese in Bottom-up- und Top-down-Prozessen. Am Ende, in einer dritten Phase, stehen metakognitive Evaluationen der verschiedenartigen Prozesse und Produkte.

Wenn es wie im Titel dieses Teilkapitels darum geht, begriffliche Schärfungen und konzeptionelle Klärungen vorzunehmen, so ist abstrakt gesehen in den Modellen selbstregulierten Lernens bei Definitionen und Phasen Einigkeit zu attestieren. Auch bei der Notwendigkeit, dass Prozesse und Produkte metakognitiv für ein spezifisches Ziel und mit Blick auf die aufgabenbezogen extrahierten und interpretierten Soll-Zustände geplant werden und ihr Ist-Zustand überwacht

wird, besteht Konsens. Markante Differenzen treten bei behavioralen, motivationalen, emotionalen und kontextuellen Bereichen auf, welche die Modelle teils erheblich anders konstruieren.

6.1.2 Metakognition und Strategien als zentrale Elemente der Selbstregulation beim Lesen (digitaler Texte)

Metakognition und metakognitive Selbstregulation sind Kernbestandteile des übergeordneten Konstrukts Selbstregulation (Tarricone, 2011). Zugleich ist Metakognition davon betroffen, dass es sich um einen nicht immer einheitlich verwendeten Terminus handelt, der von einigen hartnäckigen Unklarheiten geplagt wird, z. B. die Fragen danach, wie sich Kognitionen und die sich auf sie beziehenden, auf einer darüber befindlichen Ebene befindlichen *Meta*-Kognitionen trennscharf voneinander separieren lassen (Azevedo, 2020; Winne, 2018a). Deshalb ist dieser Aspekt eine kurze Vertiefung wert.

6.1.2.1 Metakognition – Konturen eines Konstrukts

Der bereits mehrfach gefallene Begriff „Metakognition“ bedarf einer Schärfung. Metakognition ist als Konstrukt alles andere als einheitlich definiert, aber es lassen sich Gemeinsamkeiten destillieren, wie es der Überblicksbeitrag von Norman et al. (2019) illustriert. Metakognition, so ihr Resümee bei einer Sichtung nur in der Psychologie, kennzeichnet

- eine *Bewusstheit über eigene kognitive Aktivitäten*, was einhergeht mit einer Überwachung (*Monitoring*) und einer Planbarkeit eigener kognitiver Prozesse (*Kontrolle*);
- eine *Vielzahl von Bezugspunkten* dieser Bewusstheit, darunter Wissen, Erfahrungen sowie Strategien;
- die Annahme, dass Metakognition *nützlich* – z. B. für das Lernen – ist und auf verschiedene Arten *erworben werden kann*.

In der Forschung besteht also Einigkeit darüber, dass Metakognition als Sammelbegriff für erlernbare und lernförderliche Kognitionen über Kognitionen fungiert, unter welchem sich verschiedene Konstrukte vereinen lassen, deren auf den allgemeinsten Nenner gebrachte Gemeinsamkeit darin besteht, dass sie kognitionsbezogene Kognitionen darstellen (Azevedo, 2020; Efklides, 2008; Flavell, 1979; Tarricone, 2011). Außerdem ist aus pädagogisch-psychologischer Perspektive relativ unstrittig, dass Metakognition ein integraler Bestandteil der

Selbstregulation ist (Norman et al., 2019; Tarricone, 2011; Winne, 2018a). Dies war bereits ein zentrales Ergebnis der Sichtung verschiedener Modelle in Teilkapitel [6.1.1](#).

Klassischerweise werden in der Forschung zur Metakognition zwei große Bereiche unterschieden: *Wissen* und *Regulation*, welche sich auf verschiedene Teilbereiche beziehen und systematisch feiner untergliedern lassen (s. hierfür im Detail die umfassende Taxonomie von Tarricone, 2011). Diese Unterteilung ist für die Selbstregulation insofern wichtig, als metakognitive Wissensbestände die Basis dafür bilden, in Prozessen der Regulation zur Anwendung zu gelangen. Metakognitives Wissen lässt sich gemäß Tarricones (2011) Taxonomie drei Wissensarten und drei Wissensbereichen zuordnen, die zudem miteinander kombiniert werden. Die Bereiche des Wissens bilden zudem die Bezugspunkte der eigentlichen Regulationsprozesse:

- *Metakognitives Wissen*: Die Wissensarten sind 1) deklaratives Wissen (Wissen über), 2) prozedurales Wissen (Wissen, wie) und 3) konditionales Wissen (Wissen, wann und warum). Die drei Wissensbereiche erstrecken sich auf a) personenbezogenes Wissen (Wissen über sich und andere als kognitive Wesen), b) aufgaben- und kontextbezogenes Wissen (Wissen über Aufgabenerfordernisse und -merkmale) sowie c) strategiebezogenes Wissen (Wissen über geeignete steuerbare Vorgehensweisen).
- *Metakognitive Regulation*: Die drei Wissensbereiche sind zudem jene, die in Tarricones (2011) Taxonomie als Ressource für die Ausgestaltung der Prozesse dienen (Regulation als exekutive Komponente der Metakognition) und Bezugspunkte planerischer Kontroll- und überwachender Monitoringprozesse bilden. Metakognitive Regulation meint damit das gezielte Abrufen des metakognitiven Wissens, um auf dieser Basis Vorgehensweisen top-down zu planen und bottom-up zu überwachen und ggf. zu intervenieren.

6.1.2.2 Objekt- und Metaebene und die sie verbindenden metakognitiven Prozesse

Eine der Grundschwierigkeiten der Metakognitionsforschung besteht darin, dass sie die analytische Trennung von Kognitionen und Kognitionen über Kognitionen (also die eigentliche Metakognition) theoretisch vornehmen muss. Einen bis heute stark beachteten und immer wieder aufs Neue aufgegriffenen Systemisierungsvorschlag (so z. B. bei Hacker, 1998, und Winne, 2018a) haben Nelson und Narens (1990) vorgelegt. Ihre Grundüberlegung stammt aus einem Modell aus der für die gesamte Metakognition prägenden Gedächtnisforschung und verbindet die beiden Ebenen der Kognition und der Metakognition über drei Prinzipien miteinander (Nelson & Narens, 1990):

- Prinzip 1: Die Prozesse werden gemäß ihren Bezugspunkten in zwei oder mehr spezifisch aufeinander bezogene Ebenen aufgeteilt, die Meta- und die Objektebene.
- Prinzip 2: Die Metaebene enthält ein dynamisches Modell (z. B. eine mentale Simulation) der Objektebene und ist bildlich gesprochen jener Ort, an dem metakognitive Prozesse auf der Basis von Wissensbeständen zu lokalisieren sind. Auf der Objektebene spielen sich die Prozesse des (strategischen) Lesens ab, nämlich die kognitiven Lesestrategien.
- Prinzip 3: Es gibt zwei Beziehungen zwischen den Ebenen, die als „Kontrolle“ (die Regulation) und „Überwachung“ (Monitoring) bezeichnet werden und die sich auf die Richtung des Informationsflusses zwischen der Metaebene und der Objektebene beziehen. Kontrolle ist als Top-down-Prozess zu denken, Monitoring als Bottom-up-Prozess. Kontrolle wird mitunter auch als Planung bezeichnet und umfasst gemäß Schraw und Moshman (1995) die Auswahl geeigneter Strategien sowie die Einteilung von Ressourcen, die sich auf die Leistung auswirken. Überwachung oder Monitoring kann definiert werden als parallele Wahrnehmung des eigenen Verständnisses und der Aufgabebearbeitung (Schraw & Moshman, 1995).

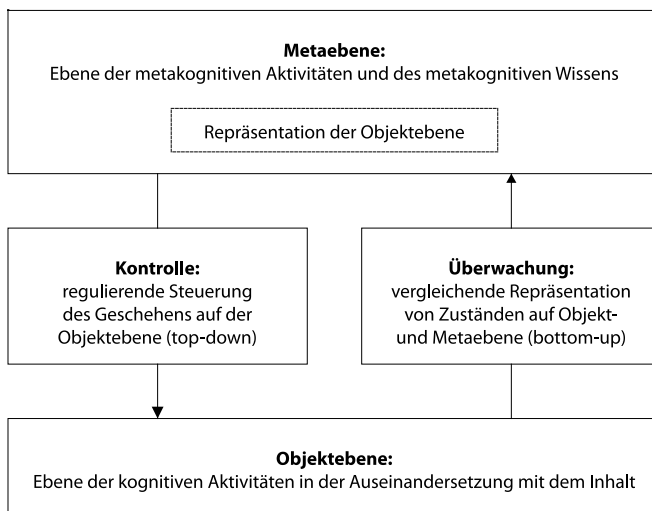


Abbildung 20: Meta- und Objektebene und die beiden konstitutiven Prozesse des Informationsflusses zwischen den Ebenen (eigene Darstellung, basierend auf de Bruin & van Gog, 2012, S. 246, sowie Nelson & Narens, 1990, S. 126)

Diese Überlegungen zur Funktionsweise von Kognition und Metakognition enthält grafisch zusammengefasst die Abbildung 20. Sie enthält unten in der Darstellung die Objekt- und zuoberst die Metaebene, verbunden werden sie mit der jeweiligen Richtung des Informationsflusses über die Kontroll- und Überwachungsprozesse.

Da Selbstregulation im Allgemeinen und selbstreguliertes Lernen im Besonderen als zielorientierte, koordinierte Aktivitäten modelliert werden, stellt sich die Frage, womit Personen ihren Fortschritt erreichen. Für das Leseverstehen wird beispielsweise in diversen Modellen ein ganzer Prozessverbund angenommen (McNamara & Magliano, 2009), und nur ein Teil davon lässt sich gemäß den Modellen der Selbstregulation als metakognitiv gezielt steuer- und modifizierbar und damit als zentral für die Selbstregulation annehmen (Panadero, 2017). Im Falle des Lesens sind damit als domänenspezifische Prozesse die Lesestrategien gemeint, im Falle des Lernens domänenübergreifende Lernstrategien. Die Granularität und die Spezifik variieren damit. Dennoch gibt es eine Gemeinsamkeit: Es handelt sich um (kognitive) Prozesse mit einer für die sie nutzenden und anwendenden Person erkennbaren Funktion. Anders gesagt: Strategien bilden jene auf verschiedenen Facetten des metakognitiven Wissens basierende Handlungspläne, die Lernende dafür verwenden, Fortschritte metakognitiv zu erzielen (über Kontrollprozesse) und diesen Weg auf dem Fortschritt zu überwachen (über Überwachungs-/Monitoringprozesse; Tarricone, 2011). Darum benötigen die Strategien, die gemäß der Darstellung aus Abbildung 20 als kognitive Strategien primär auf der Objektebene anzusiedeln sind, vertiefte Aufmerksamkeit.

6.1.2.3 Strategien als prozedurale Handlungspläne

Definition und drei Hauptmerkmale

Der Ausdruck „Strategien“ bzw. „Lernstrategien“ ist für das Konzept selbst-reguliertes Lernen von eminenter Wichtigkeit, da er ein zentrales Konstrukt bezeichnet. Der Terminus wird häufig und durchaus mit divergierendem Inhalt verwendet, sodass eine Definition nötig wirkt bzw. eine Eingrenzung anhand typischer Merkmale geboten erscheint (Alexander et al., 1998; Gu, 2012). Das gilt umso mehr, als „Lernstrategien“ in der Forschungsliteratur auch in Bezug auf das übergeordnete Konstrukt Metakognition sowie das nochmals darüber anzusiedelnde selbstregulierte Lernen nicht immer einheitlich verwendet werden. Da eine umfassende Diskussion und Herleitung des Konstrukts außerhalb der Ziele dieses Buches liegen, erfolgt stattdessen an dieser Stelle eine explizite und ihrerseits bereits deutlich synoptische Definition, welche zentrale Merkmale beinhaltet:

„Eine Strategie ist eine mit einem Aufwand verbundene und zielgerichtete Form von prozeduralem Wissen, welches im Langzeitgedächtnis gespeichert ist. Es gibt verschiedene Arten von Strategien, und diese Strategien müssen bei komplexen Aufgaben miteinander koordiniert werden.“ (van Meter & Stepanik, 2020, S. 87)

Drei Merkmale von Strategien sind nach van Meter und Stepanik (2020) bedeutsam an der obigen Definition: erstens „eine mit einem Aufwand verbundene und zielgerichtete Form von prozeduralem Wissen“, zweitens „verschiedene Arten“ und, drittens, dass sie „bei komplexen Aufgaben miteinander koordiniert werden“ müssen. Dies lässt sich knapp erläutern:

- 1) *Strategien als prozedurales Wissen, das zielgerichtet und das in der Ausführung mit Aufwand verbunden ist:* Strategien sind als Handlungspläne darauf ausgerichtet, dass mit ihnen ein Zustand erreicht wird, z. B. eine nötige Information aus dem Internet, die man gefunden haben will. Solche Ziele können dem konditionalen Wissen über Strategien zugeordnet werden, weil sie den angestrebten Zustand selbstregulatorischer Vorgehensweisen definieren und über die Angemessenheit der Vorgehensweisen angesichts der Ressourcen (etwa: Wie viel Zeit habe ich, um eine Information zu recherchieren? Wie genau muss ich verstanden haben, was diese Informationen bedeutet?) und der erwartbaren Effektivität des Vorgehens mitentscheiden. Strategien als steuer- und auswählbare Handlungsoptionen der Zielerreichung sind mit Aufwand verbunden, d. h. man muss sie aus einem Repertoire auswählen, sie absichtsvoll ausführen und die Zielerreichung auch überwachen – das macht sie anders als automatisierte Fähigkeiten zu einem kognitiven aufwändigen Unterfangen, was auch die kognitive Belastung erhöht. Ein letztes Merkmal ist das im Langzeitgedächtnis gespeicherte prozedurale Wissen, also das Wissen, welche Schritte und Handlungen in welcher Reihenfolge zielführend sind, z. B. das Spezifizieren des Rechercheziels, das Formulieren geeigneter Suchbegriffe bzw. Prompts in einer geeignet wirkenden Datenbank, der Umgang mit Suchergebnissen etc. Prozedurales Wissen ist dafür nötig, die einzelnen Teilschritte und ihre Reihenfolge erfolgreich zu bewältigen, um das Ziel zu erreichen, sie dienen also der Ausführung von Handlungsplänen wie der Internetrecherche.
- 2) *Typologien/Taxonomien von Strategien nach Bezugspunkten und Funktionen:* Ein weiteres zentrales Merkmal von Strategien ist, dass sie unterschiedliche Bezugspunkte und Funktionen aufweisen, was verschiedene Klassifikationen in Form von Typologien und Taxonomien hervorgebracht hat. In ihrem Überblick, der seinerseits eine Synopse verschiedener aktueller theoretischer Handbuchkapitel extrahieren van Meter und Stepanik (2020) mehrere Dimensionen: a) selbstregulatorische Strategien, die sich untergliedern lassen

in metakognitive, kognitive, emotional-affektive, behaviorale und soziale Strategien; b) die Spezifität, also ob Strategien breit einsetzbar sind, nur in einzelnen Wissensgebieten oder teils nur bei einzelnen Aufgaben; c) die Verarbeitungstiefe, was eine vor allem auf die kognitiven Strategien zutreffende Unterscheidung von oberflächen- oder tiefenorientierten Auseinandersetzungen mit den Inhalten darstellt. In dem Beispiel der Recherche nach Inhalten sind die Schritte der Klärung des Informationsbedürfnisses beispielsweise metakognitiven Strategien zuzuordnen, während das Formulieren von Suchbegriffen eher zu den kognitiven Strategien zählt. Typologien für Lesen digitaler (multipler) Dokumente legen beispielsweise Afflerbach und Cho (2009), Cho, Afflerbach und Han (2018) und List (2020) vor; einen ausführlicheren Überblick liefert Philipp (2015). Dies wird erneut im Teilkapitel [6.2.3](#) aufgegriffen.

- 3) *Koordinationsnotwendigkeit in komplexen Settings*: Insbesondere das digitale Lesen, das wegen seiner grundsätzlich hohen Offenheit nicht umsonst als Problemlöseprozess konzeptualisiert wird, macht es nötig, dass verschiedene Handlungen unter einem übergeordneten Ziel koordiniert werden müssen. Gerade komplexe Aufgaben, die verschiedene Lösungswege aufweisen und sich in verschiedene unterscheidbare, aber miteinander verknüpfte Strategien zergliedern lassen, erfordern es, dass Personen metakognitiv planen und überwachen, welche Handlungen sie wann wie ausführen. Diesem Punkt geht die nachstehende Fokussierung auf Strategien als adaptives, mehrere Schritte umfassendes Vorgehen nach.

Strategien als adaptives mehrschrittiges Vorgehen

Auf einen besonderen Charakter der Strategien macht Winne (2001) aufmerksam, indem er den zweiten Punkt aus der obigen Aufzählung behandelt. Er präzisiert den Begriff der „Strategie“, indem er ihn mit dem der „Taktik“ verbindet und diese Prozesse von angeborenen (bzw. wie er es nennt: „primitiven“; Winne, 2018a) Prozessen abgrenzt. Für die Zwecke dieses Buches sind Taktiken und Strategien besonders relevant. Dabei ist „Strategie“ das übergeordnete Konstrukt, bei dem die gezielte metakognitiv gesteuerte Nutzung der Taktiken den Ausschlag für ein lernstrategisches Verhalten gibt.

Taktiken bezeichnen eine kurze einzelne kognitive Operation oder eine kurze Sequenz von solchen Operationen, wobei diese sich auf die zielgerichtete Verarbeitung von Informationen beziehen und durchaus durch Übung automatisiert erfolgen (Winne & Marzouk, 2019). Winne (2001) zufolge handelt es sich bei den Taktiken um kognitive Schemata in Form von gedächtnisbasierten Regeln, die dem Muster Wenn-dann folgen und damit Bedingungen und einzelne Handlungen miteinander als Bedingungsgefüge postulieren, etwa wie folgt: „Wenn ich einen Begriff nicht kenne, dann recherchiere ich ihn mit einer Suchmaschine“. Die

Taktik ist die Recherche in der Suchmaschine, die Bedingung ist die mangelnde Kenntnis eines Begriffs. Solche Taktiken können zudem mehrere Bedingungen aneinander koppeln, z. B. „Wenn ich einen Begriff nicht kenne und er mir wichtig erscheint, dann recherchiere ich ihn mit einer Suchmaschine“. Je mehr solcher Bedingungen enthalten sind – darunter auch motivationale Komponenten wie Selbstwirksamkeitsüberzeugungen oder aufgabenbezogene Wert-Zuschreibungen –, desto sensibler ist eine Person für die Spezifika von Aufgabenerfordernissen. Das bedeutet: Das metakognitive konditionale Strategiewissen über die Bedingungen und die adäquate Handlung im Sinne der Taktik ist dann umso ausgeprägter, je akkurater die Bedingungen mit einer Handlungsoption verbunden sind. Zusammengefasst lässt sich also die Taktik als eine Handlungsoption im Sinne einer Konsequenz unter einer mehr oder minder großen Anzahl von Bedingungen terminologisch fassen.

Dem Konstrukt „Strategien“ kommt in der Modellierung von Winne (2001) im Vergleich zu den Taktiken eine übergeordnete Funktion und Rolle zu. Strategien sind hier als ebenfalls zielgerichtete und koordinierende Handlungspläne zu verstehen, in denen verschiedene Taktiken als Schritte von der lernenden Person flexibel dazu genutzt werden, das übergeordnete Ziel zu erreichen, etwa das Schreiben eines Textes über mehrere Internetdokumente. Das zuvor genannte Beispiel mit dem unbekannten Begriff wäre hierbei nur eine einzelne Episode in einem umfassenderen Geschehen, ein Teilproblem, dessen Lösung bzw. Nicht-Lösung seinerseits das weitere Vorgehen beeinflusst. Deshalb wird die Wenn-dann-Formulierung, die für eine Taktik noch ausreichte, nun erweitert, denn sie genügt nicht mehr. Das Muster ist daher „Wenn-dann-sonst“, womit es um Handlungsalternativen im Sinne weiterer Taktiken geht. Das bedeutet, dass Strategien darauf angewiesen sind, dass mehrere Taktiken verfügbar sein müssen und dass diese Taktiken modular zu denken sind, nämlich als einzelne Komponenten. Wer, um im Beispiel zu bleiben, einen unbekannten Begriff (Wenn-Bestandteil) mittels Suchmaschine sucht (Dann-Bestandteil), aber feststellt, dass das Problem nicht gelöst ist, weil etwa die Beschreibung des Begriffs unverständlich ist, muss sich strategisch entscheiden, wie er nun vorgeht (Sonst-Bestandteil). Daran wird deutlich, dass ein strategisches Vorgehen sich aus Kontrollprozessen ergibt, welche das Vorgehen im Sinne der Auswahl und Aktivierung adäquat wirkender Taktiken ebenso beinhaltet wie das über Monitoring ablaufende Verändern von Taktiken, sollten diese nicht zum Erfolg führen. Insofern sind Strategien mehrschrittige, Multi-Taktiken-Prozesse, die sensibel für die Zielerreichung sind und damit auch adaptiv erfolgen (Winne & Marzouk, 2019). Bemerkenswert ist, dass über die Erfahrung, dass strategisch gesteuerte Taktiken nicht das Ziel erreichen, sowohl das eigentliche Geschehen in der Aufgabenbearbeitung dynamisiert wird als auch über Rückkopplungen das eigene konditionale Strategiewissen modifiziert wird (Winne, 2001).

Durch diese Unterscheidung von Taktiken als modulare und in sich durchaus komplexe kurze Handlungen bzw. Handlungssequenzen und Strategien als übergeordnete Handlungspläne auf der Basis von Taktiken und Alternativen zwischen Taktiken wird es möglich, der Komplexität von so vielen Einzelhandlungen, wie sie für das Lesen von digitalen Dokumenten nötig sind (Afflerbach & Cho, 2009, 2010; Cho & Afflerbach, 2017), aus theoretischer Sicht gerecht zu werden. Zudem wird bei der Modellierung von Strategien als Handlungspläne mit metakognitiver Selbstregulation die enge Verknüpfung von Kognitionen und Metakognition angemessen plausibilisierbar, wie Kognitionen und Metakognitionen regelgeleitet, nämlich über die Wenn-dann-sonst-Logik, zusammenwirken, ohne dabei individuelle Unterschiede zu verwischen. Dieses Zusammenspiel wird in verschiedenen deskriptiven Modellen immer wieder behandelt (s. bspw. Cho, Afflerbach & Han, 2018).

6.1.3 Exekutive Funktionen

Lesen ist als ein zuvorderst kognitiver Prozessverbund und damit darauf angewiesen, dass Personen ihre kognitiven Ressourcen wie Wissensbestände zum einen (Smith et al., 2021), aber auch die Funktionsweise ihres gesamten Gedächtnisses mitsamt seinen Eigenschaften zum anderen möglichst optimal nutzen (Peng & Kievit, 2020). Empirisch zeigen sich in diversen Studien und ihren metaanalytischen Synopsen zudem regelmäßig positive Zusammenhänge zwischen exekutiven Funktionen und verschiedenen Leseleistungen (Ahmed & Wagner, 2020; Carretti et al., 2009; Cortés Pascual et al., 2019; Daneman & Merikle, 1996; Follmer, 2018; Follmer & Sperling, 2020; Jacob & Parkinson, 2015; Ober et al., 2020; Peng et al., 2018; Robson et al., 2020; Yeniad et al., 2013). Diese Zusammenhänge werden zunehmend im Kontext des digitalen Lesens thematisiert, und dabei wird hervorgehoben, dass das Lesen digitaler Texte erhöhte Anforderungen an die exekutiven Funktionen von Leserinnen und Lesern stellt (Wylie et al., 2018). Exekutive Funktionen sind allerdings kein homogenes Konstrukt.

6.1.3.1 Ein Sammelbegriff für die Kontrolle zielführender Kognitionen mit drei Komponenten

Mit dem Ausdruck „exekutive Funktionen“ bezeichnet Diamond (2013, S. 136) „eine Sammlung von Top-down-Steuerungsprozessen, die verwendet werden, wenn es unklug, unzureichend oder unmöglich wäre, sich auf automatisch ablaufende, instinktive oder intuitive Prozesse zu verlassen.“ Demnach sind exekutive Funktionen im Sinne der in Teilkapitel [6.1.2.2](#) eingeführten Differenzierung von Prozessebenen als allgemeine und fundamentale metakognitive

Steuerungsprozesse zu bezeichnen, und diese Prozesse gelten als Grundlage für metakognitive Prozesse höherer Ordnung, auch und gerade für den Erfolg in einzelnen Domänen wie dem Lesen (Diamond, 2013; Y.-S. G. Kim, 2020).

Die Dimensionalität des Konstrukts exekutive Funktionen ist nach wie vor Gegenstand von Kontroversen (Baggetta & Alexander, 2016; Follmer & Sperling, 2020; Jurado & Rosselli, 2007; Miyake & Friedman, 2012), doch es lassen sich drei Komponenten als Kernbestand festhalten: *Arbeitsgedächtnis*, *inhibitorische Kontrolle* sowie *kognitive Flexibilität*. Gemäß Diamond (2013) bauen sie aufeinander auf. Diese sequenzielle Perspektive lässt sich ebenfalls entwicklungspsychologisch demonstrieren, da die kognitive Flexibilität erst bei schulpflichtigen Kindern und Jugendlichen als eigenständiger Bestandteil der exekutiven Funktionen nachweisbar ist (Karr et al., 2018): 1) Das Arbeitsgedächtnis gilt als Basis für 2) die inhibitorische Kontrolle, und 1) das Arbeitsgedächtnis sowie 2) die inhibitorische Kontrolle stellen die Grundlage für 3) die kognitive Flexibilität dar. Im Einzelnen ist Folgendes mit den drei Komponenten gemeint:

- 1) *Arbeitsgedächtnis*: Das Arbeitsgedächtnis bezieht sich auf die „Mechanismen und Prozesse, welche die mentalen Repräsentationen, die derzeit für eine laufende kognitive Aufgabe am meisten benötigt werden, zur Verarbeitung bereithalten“ (Oberauer, 2019, S. 1). Es geht mithin darum, möglichst viele der Zielerreichung dienende Informationen auf möglichst günstige Weise kognitiv präsent zu halten und diese Informationen zu verarbeiten.
- 2) *Inhibitorische Kontrolle*: Die inhibitorische Kontrolle lässt sich definieren als „die Fähigkeit, die eigene Aufmerksamkeit, das eigene Verhalten, die eigenen Gedanken und/oder Emotionen zu kontrollieren, um sich über eine starke innere Neigung oder eine äußere Verlockung hinwegzusetzen und stattdessen das zu tun, was angemessener ist oder gebraucht wird“ (Diamond, 2013, S. 137). Hiermit ist also die Fähigkeit angesprochen, störende Reize und Impulse zielführend abzuschirmen und damit das genuine Aufgabenziel strikt zu verfolgen.
- 3) *Kognitive Flexibilität*: Sie bezeichnet die Fähigkeit, einen „Wechsel der Perspektive oder der Herangehensweise an ein Problem, eine flexible Anpassung an neue Anforderungen, Regeln oder Prioritäten (z. B. beim Wechsel zwischen Aufgaben) vorzunehmen“ (Diamond, 2013, S. 137). Diese Fähigkeit setzt voraus, dass Personen eine vorherige Perspektive unterdrücken (mittels inhibitorischer Kontrolle) und neue Informationen im Arbeitsgedächtnis re-präsentieren.

Einen Überblick über die drei Bestandteile der exekutiven Funktionen, die für die Ausführungen dieses Kapitels wichtig sind, gibt Tabelle 32. Diese drei Bestandteile werden nachstehend entfaltet.

Tabelle 32: Zusammenhänge zwischen exekutiven Funktionen und Selbstregulationsmechanismen

Exekutive Funktion	Selbstregulationsmechanismus	Bezug zum Lesen
Operationen im Arbeitsgedächtnis (6.1.3.2)	• aktive Repräsentation selbstregulativer Ziele und Standards • Top-down-Kontrolle der Aufmerksamkeit auf zielrelevante Informationen und fort von aufmerksamskeitsfordernden Reizen • Abschirmung von Zielen und Standards vor Störungen durch Fokussierung auf genuines Ziel • Unterdrückung von abschweifenden Gedanken durch Fokussierung auf zielführende Kognitionen • (Herunter-)Regulieren unerwünschter Affekte, Wünsche und Begehren	• Verfügbarhalten wichtiger bzw. relevanter Informationen zum Leseziel • Unterdrückung und Löschen irrelevanter bzw. unwichtiger Informationen für das Leseziel
Inhibitorische Kontrolle (6.1.3.3)	• aktive Hemmung vorherrschender Impulse und gewohnheitsmäßiger, „gedankenloser“ Verhaltensweisen	• Unterdrückung von Informationen im Arbeitsgedächtnis, die mit dem Verstehen (oder dem sonstigen Leseziel) interferieren könnten • Schutz des Arbeitsgedächtnisses vor kognitiver Überlastung durch irrelevante Informationen
Kognitive Flexibilität (6.1.3.4)	• flexibler Wechsel zwischen verschiedenen Maßnahmen, welche dem gleichen (selbstregulierten) Ziel dienen (Wechsel der Maßnahmen) • Wechseln zwischen mehreren Zielen (Zielverschiebung/-Balancieren)	• Basis dafür, verschiedene Leseprozesse auf mehreren Ebenen effektiv zu koordinieren

(Quelle: Hofmann et al., 2012, S. 175, mit leichten Ergänzungen der Originaldarstellung; Ergänzung der lesespezifischen Spalte auf der Basis von Butterfuss & Kendeou, 2018)

6.1.3.2 Arbeitsgedächtnis

Das Arbeitsgedächtnis gilt als Basis für die inhibitorische Kontrolle und darauf aufbauend die kognitive Flexibilität (s. dazu Diamond, 2013). Das Konstrukt Arbeitsgedächtnis ist uneinheitlich definiert (Adams et al., 2018; Baddeley, 2012; Cowan, 2017), dennoch gibt es bei aller Differenz drei Gemeinsamkeiten:

- sein temporal limitierter Charakter,
- sein Fokus auf Verarbeitungsprozesse außerhalb des Langzeitgedächtnisses und
- sein Verfügbarhalten von aktuell nötigen Informationen (Cowan, 2017).

Als ein Bestandteil des menschlichen kognitiven Systems hat das Arbeitsgedächtnis aus mindestens zwei Gründen starkes Interesse in Theorie und Empirie zur Selbstregulation bzw. ihrer Nachbargebiete auf sich gezogen. Grund Nummer

eins ist, dass es der Ort ist, an dem die Aufmerksamkeit auf die Verarbeitung von Informationen gelenkt wird, und da selbstreguliertes, gezieltes Handeln durchaus sehr selektiv stattfinden kann, stellt sich die Frage, was Aufmerksamkeit ist und wie sie mit dem Arbeitsgedächtnis zusammenhängt. Der zweite Grund lautet: Das Arbeitsgedächtnis hat limitierte Verarbeitungskapazitäten (Cowan, 2010), was bei komplexen, bewussten und dynamischen Prozessverbünden danach verlangt, diese Kapazitäten strategisch unter Berücksichtigung von aufgaben- und personenbezogenen Merkmalen einzusetzen (Reynolds, 2000; Wirth et al., 2020). Aus empirischer Sicht haben insbesondere die Kapazitäten des Arbeitsgedächtnisses in Verbindung mit dem Leseverstehen eine erhöhte Forschungsaktivität nach sich gezogen, welche substanzielle Zusammenhänge in Metaanalysen zur Folge hatte (s. z. B. Follmer, 2018, und Spiegel et al., 2021). An dieser Stelle wird jedoch nicht dies fokussiert, sondern ein anderer Akzent gesetzt. Denn es gibt aktuell sehr deutliche Bemühungen, die Theorie der kognitiven Belastung mit der Selbstregulationstheorie zu synthetisieren (de Bruin et al., 2020; Seufert, 2020; kritisch dazu Sweller & Paas, 2017). Ehe dies dargestellt wird, ist es jedoch noch nötig, kursorisch auf das Verhältnis von Aufmerksamkeit und Arbeitsgedächtnis einzugehen.

Aufmerksamkeit und Arbeitsgedächtnis

Wie aus der Tabelle 32 im Teilkapitel [6.1.3.1](#) hervorgeht, sind für die Selbstregulation die Aufmerksamkeitslenkung und -fokussierung im Arbeitsgedächtnis zentral. Aufmerksamkeit beschreibt die Selektivität bei der Verarbeitung von externalen sensorischen Informationen bzw. internal generierten Informationen (Chun et al., 2011) und hierbei wiederum mit einer begrenzten Anzahl von Informationen. Aufmerksamkeit wird seit Langem eng mit dem Arbeitsgedächtnis in Verbindung gebracht (Oberauer, 2019), speziell mit der *zentralen Exekutive* in Baddeleys einflussreichem Modell – in diesem Subsystem des Arbeitsgedächtnisses finden steuernde Prozesse der Verarbeitung von Informationen in weiteren Komponenten des Arbeitsgedächtnisses statt (Baddeley, 2012). Engle und Kane (2004) nennen die Fähigkeit, für eine Aufgabe zielführende Informationen im Arbeitsgedächtnis zu behalten und gegen andere, nicht-zielführende Informationen abzuschirmen, aus diesem Grund „exekutive Aufmerksamkeit“. Diese beinhaltet also sowohl das proaktive Aufrechterhalten eines Ziels als auch einen reagierenden Umgang mit Ablenkungen und Störungen.

Der genaue Zusammenhang von zu lenkender Aufmerksamkeit und Arbeitsgedächtnis ist in der Kognitionspsychologie Gegenstand von unterschiedlichen Konzeptualisierungen (Adams et al., 2018). Eine Synopse gegenwärtiger Theorien und empirischer Befunde hat Oberauer (2019) vorgenommen, und seine Zusammenfassung lässt sich folgendermaßen bündeln:

- *Das Arbeitsgedächtnis ist eine Form der kontrollierten, gefilterten Aufmerksamkeit:* Im Arbeitsgedächtnis wird nur ein Teil der Informationen kurzzeitig gespeichert. Selbst innerhalb eines Sets von im Arbeitsgedächtnis befindlichen Informationen kann sich die Aufmerksamkeit nur auf einen Teil der aktuell präsenten Informationen richten. Über das Verfügbarhalten und Unterdrücken der Informationen entscheiden die Ziele und Absichten der lesenden Person (kontrollierte Aufmerksamkeit). Informationen, welche der Zielerreichung nicht (mehr) dienen, werden hingegen aus dem Arbeitsgedächtnis getilgt.
- *Grundlagen der Entscheidung für oder gegen die Informationsverarbeitung sind selbst Teil der Repräsentationen im Arbeitsgedächtnis:* Um die Aufmerksamkeit zu lenken, ist es nötig, die Passung von benötigter Information bzw. deren Verarbeitung zu bestimmen. Dieses Ziel bzw. der Standard, an dem sich die Eignung bemisst, kann unter Umständen selbst verarbeitungsintensiv sein und muss deshalb ebenso im Arbeitsgedächtnis präsent gehalten werden. Das ist dann der Fall, wenn solche aufmerksamkeitslenkenden Repräsentationen entweder für die Person neu sind oder aber mit Inhalten des Langzeitgedächtnisses konfligieren.

Zusammengenommen lässt sich für die Zwecke dieses Buches festhalten: Das Arbeitsgedächtnis ist der Ort, an dem für eine begrenzte Zeit die Wahrnehmung auf eine limitierte Anzahl von Informationen gerichtet wird, um diese Informationen aktiv zu verarbeiten. Für die Selbstregulation beim Lesen ist dies wichtig, weil die Nutzung dieses ‚Flaschenhalses‘ der kognitiven Verarbeitung möglichst effektiv über die gesamte Dauer der selbstregulatorischen Episoden hinaus erfolgen sollte. Deshalb ist – zumindest bei neuen Anforderungen und nicht-automatisch ablaufenden Prozessen – die Verwendung von Strategien als so bedeutsam zu benennen, weil diese unter anderem dabei helfen, über Verlangsamungen Prozesse gezielt zu segmentieren und metakognitiv zu regulieren.

Arbeitsgedächtnis und Selbstregulation

Die Zusammenhänge von Arbeitsgedächtnis und Selbstregulation/selbstreguliertem Lernen werden in der Forschung aus zwei Richtungen betrachtet. Während ein erster Zugang die kognitiven Kosten der Selbstregulation betrachtet, versucht ein zweiter Zugang, Arbeitsgedächtnis und Selbstregulation zu integrieren.

Erster Zugang: Perspektiven der Theorien der kognitiven Belastung und der kognitiven Theorie des multimedialen Lernens

Die beiden Theorien der kognitiven Belastung und des multimedialen Lernens sind insofern verwandt, als sie auf grundsätzlich ähnlichen Annahmen basieren,

die die Funktionsweise des menschlichen Gedächtnisses und insbesondere des Arbeitsgedächtnisses betreffen. Besonders wichtig sind die Prämissen, dass Personen eine limitierte Verarbeitungskapazität des Arbeitsgedächtnisses aufweisen und dass für das Lernen zentrale Prozesse des Lernens eben dort stattfinden, sodass bei aller inhaltlicher Differenz bzw. Spezialisierung die Frage, womit und wie das Arbeitsgedächtnis im Dienste des Lernens genutzt wird, im Zentrum beider Theorien steht (Kalyuga, 2011; Mayer, 2014; Sweller et al., 2011, 2019; Vu et al., 2022). Die Theorie der kognitiven Belastung (Sweller et al., 2011) legt ihr Hauptaugenmerk darauf, wie lernförderliche bzw. überflüssige, lernhinderliche Belastungen des Arbeitsgedächtnisses zustande kommen, die um die Kapazitäten des Arbeitsgedächtnisses konkurrieren und es beanspruchen. Die Theorie des multimedialen Lernens (Mayer, 2021) hat sich zunehmend darauf konzentriert, Prinzipien über essenzielle, günstige bzw. erschwerende Prozesse des Lernens von sprachlich und visuell kombiniert dargebotenen Inhalten zu abstrahieren. Dies lenkt den Blick darauf, welche schwierigkeitsgenerierenden Merkmale beim digitalen Lesen aus der Warte der optimalen bzw. nicht-optimalen Beanspruchung begrenzter Verarbeitungskapazitäten im Arbeitsgedächtnis bestehen.

DeLeeuw und Mayer (2008) bieten hierfür mit ihrer „triarchischen Theorie der kognitiven Belastung“ einen hilfreichen Rahmen, da diese Theorie bereits auf einer Synthese der beiden Theorien der kognitiven Belastung einerseits und der kognitiven Theorie des multimedialen Lernens andererseits beruht. Die triarchische Theorie differenziert drei Arten der kognitiven Belastung:

- a) Belastung, die durch die *lerngegenstandsinhärente oder auch essenzielle Verarbeitung* nötig ist;
- b) Belastung, die durch das *Lernen bzw. durch generative Lernprozesse hervorgerufen* wird;
- c) *zusätzliche, sachfremde Belastung*, welche dem Lernziel nicht dient, sondern eine vermeidbar wirkende Zusatzbelastung darstellt.

Diese drei Grundformen der kognitiven Belastung – die ersten beiden gelten dabei als günstig, die dritte hingegen nicht – bilden den Gegenstand der Gegenüberstellung in Tabelle 33. Was in dieser Gegenüberstellung aus der Warte der Selbstregulation besonders ins Auge sticht, ist die *Kategorie der lernbezogenen Last bzw. der generativen Verarbeitung*, die je nach Theorie dazu dient, eine tiefenorientierte Auseinandersetzung mit Inhalten zu ermöglichen und neue Inhalte mit dem Langzeitgedächtnis systematisch zu verknüpfen (s. Zeile b) in Tabelle 33). Hier lassen sich strategische Lese- und Lernprozesse recht eindeutig lokalisieren, zumal sie eine lernendenseitige Variable betreffen, die aus Sicht der (metakognitiven) Selbstregulationsforschung als zentral gilt (s. Teilkap. [6.1.2](#)).

Tabelle 33: Gegenüberstellung von Arten der kognitiven Belastung und drei Arten kognitiver Verarbeitungen beim multimedialen Lernen

1) Theorie der kognitiven Belastung		2) Kognitive Theorie des multimedialen Lernens	
Art der kognitiven Belastung	Definition, Veränderlichkeit und Auswirkung auf das Lernen	Art der kognitiven Verarbeitung	Definition, Veränderlichkeit und Auswirkung/Bedeutung für das Lernen
a) lerngegenstands-inhärente kognitive Last (intrinsic cognitive load)	• <i>Definition:</i> dem Lerngegenstand innewohnende Komplexität, die sich vor allem aus der Element-Interaktivität des Gegenstands und dem eigenen Vorwissen der lernenden Person ergibt • <i>Veränderlichkeit:</i> nicht ohne weiteres veränderbar, nur durch Veränderung des Lerngegenstands bzw. des Vorwissens der lernenden Person beeinflussbar • <i>Auswirkung auf das Lernen:</i> unabdingbar, um den Gegenstand in seiner Komplexität angemessen zu verstehen	a) essenzielle Prozesse (essential cognitive processing)	• <i>Definition:</i> Prozesse, die dazu nötig sind, essenzielle, eher oberflächenorientierte Informationen im Arbeitsgedächtnis präsent zu halten • <i>Veränderlichkeit:</i> wird von der Komplexität des Lerngegenstands bestimmt, lässt sich durch Optimierung der Gestaltung der multimedialen Informationen reduzieren • <i>Auswirkung/Bedeutung für das Lernen:</i> Auswahl von sprachlichen und piktorialen Informationen für das Lernen, eher oberflächenorientiertes Lernen
b) lernbezogene kognitive Last (ehemals: germane cognitive load, inzwischen: germane processing)	• <i>Definition:</i> Belastung, die sich daraus ergibt, dass Personen ihr Arbeitsgedächtnis möglichst effektiv dazu nutzen, Schemata aufzubauen oder zu verändern bzw. allgemein mit der Element-Interaktivität zu verfahren • <i>Veränderlichkeit:</i> liegt vor allem in der Person des Lerners/der Lernerin begründet, die über Lernstrategien dazu befähigt wird, die kognitive Last zu optimieren • <i>Auswirkung auf das Lernen:</i> lernförderlich, da diese Prozesse sich auf die lerngegenstandsinhärente kognitive Last beziehen und dafür benötigt werden, diese günstig zu verarbeiten und die Arbeitsgedächtniskapazitäten lernstrategisch zu nutzen	b) generative Verarbeitung (generative cognitive processing)	• <i>Definition:</i> Prozesse, die für tieferes Verstehen und für Transfer des Gelernten nötig sind • <i>Veränderlichkeit:</i> beeinflussbar durch positive Motivation der Lernenden und durch Gestaltung der Lerngegenstände zur Optimierung generativer Lernprozesse • <i>Auswirkung/Bedeutung für das Lernen:</i> Organisations- und Integrationsprozesse zum Aufbau sinnvoller kohärenter mentaler Modelle auf der Basis sprachlicher und piktorialer Informationen unter Rückgriff auf Vorwissensbestände

1) Theorie der kognitiven Belastung		2) Kognitive Theorie des multimedialen Lernens	
Art der kognitiven Belastung	Definition, Veränderlichkeit und Auswirkung auf das Lernen	Art der kognitiven Verarbeitung	Definition, Veränderlichkeit und Auswirkung/Bedeutung für das Lernen
c) sachfremde kognitive Last (extraneous cognitive load)	• <i>Definition</i>: Belastung, die sich aus der instruktionalen Darbietung und Vermittlung und der Interaktivität dieser sachfremden Elemente ergibt und mit dem Lerngegenstand als solchem nichts zu tun hat • <i>Veränderlichkeit</i>: optimier- und reduzierbar • <i>Auswirkung auf das Lernen</i>: nicht lernförderlich, sondern mindestens tendenziell lernerschwerend, da zusätzliche Belastung für das Arbeitsgedächtnis generiert wird, welche zulasten der lerngegenstandsinhärenten und lernbezogenen Belastung auftritt	c) sachfremde Verarbeitung (extraneous cognitive processing)	• <i>Definition</i>: Prozesse, die nicht mit dem Ziel des Lernens verbunden sind bzw. ihm dienen • <i>Veränderlichkeit</i>: beeinflussbar durch Prinzipien der Reduktion sachfremder kognitiver Belastung • <i>Auswirkung/Bedeutung für das Lernen</i>: nicht förderlich für lernbezogene Prozesse des Auswählens, Organisierens und Integrierens sprachlicher und piktorialer Informationen

(Quelle der Darstellung: 1) Theorie der kognitiven Belastung; Kalyuga, 2011, S. 13–14; Paas & Sweller, 2014, S. 39; Schnotz & Kirschner, 2007, S. 471–472; Sweller, 2010, S. 124–126; Sweller et al., 2019, S. 263–264; 2) kognitive Theorie des multimedialen Lernens; Mayer, 2014, S. 59–61, 2021, S. 31–53)

Eine Thematik, die seit jüngerer Zeit größere Aufmerksamkeit auf sich zieht, betrifft das Verhältnis von Arbeitsgedächtnis und dessen Beanspruchung durch selbstregulatorische Prozesse (s. bspw. de Bruin et al., 2020; de Bruin & van Merriënboer, 2017; Seufert, 2018, 2020; Sweller et al., 2019). Grob lassen sich die gegenwärtigen Bemühungen um Synthesen zwei Richtungen zuordnen. Die auch in der folgenden Darstellung vorkommende erste Richtung betrifft Bemühungen seitens der Fraktion der Forscherinnen und Forscher mit ihrem Fokus auf die kognitive Belastung. Die Forschungslage ist zusehends komplexer geworden (Kriegelstein et al., 2022) und hat zu Ausdifferenzierungen geführt, welche die Nähe zum Selbstregulationsparadigma deutlicher haben zutage treten lassen (Moreno & Park, 2010). Die zweite Richtung stammt aus der Selbstregulationsforschung. Diese beginnt sich, nicht zuletzt ausgehend von den Anforderungen des Lernens in digitalen Kontexten, damit zu befassen, wie Lernerinnen und Lerner solche Anforderungen erfolgreich meistern können (de Bruin et al., 2020).

Perspektive 1: Unklare Effekte der kognitiven Belastung und die Rolle der Selbstregulation

Zunächst zur ersten Richtung: Zunehmend gewinnen metakognitive Aspekte aus der Warte der Theorie der kognitiven Belastung an Bedeutung, die sich in teils länger anhaltende Debatten zum Konstrukt der kognitiven Belastung einreihen (Jong, 2010). Diese Perspektiven sind deshalb wichtig, weil sie auf theoretische Schärfungsnotwendigkeiten hin- und auf blinde Flecken der Theorie verweisen. Insbesondere die Fragen, welche Belastungsarten lernförderlich sind und welchen Stellenwert Ziele und die durch sie hervorgerufenen Strategien haben, sind veritable Kristallisationspunkte der Fragen danach, wie die Theorie der kognitiven Belastung sich öffnen sollte für Überlegungen aus dem Feld der Selbstregulationsforschung.

So diskutieren Kalyuga und Singh (2016) beispielsweise, dass zum einen die Element-Interaktivität bei der lerngegenstandsinhärenten Belastung zu simplizistisch sei. Sie weisen darauf hin, dass komplexes Lernen – wie etwa das digitale Lesen zum Zweck des sich anschließenden Schreibens – mit multiplen Zielen erfolgt, und generell solche Ziele unterbelichtet seien. Eben solche variablen Ziele seien jedoch als Bestandteil einer allgemeinen Intensität kognitiver Aktivitäten zu werten und modifizieren damit, welche Art von kognitiver Belastung woraus stammt. Hinzu komme zum anderen, dass solche Ziele sowohl lernförderlich als auch lernhinderlich sein können, sie also nicht per se produktive oder unproduktive Last generieren. Diese Kritik richtet sich an allzu rigide und allzu starre Konzeptionen bei den Kategorien kognitiver Belastung und ihren Folgen. Die

kognitive Belastung ist für das digitale Lesen ebenfalls von Belang: Denn dieses Lesen ist ein Prozessverbund verschiedenster Aktivitäten, die mutmaßlich eine hohe Beanspruchung für das Arbeitsgedächtnis mit sich bringen (DeStefano & LeFevre, 2007; Wylie et al., 2018). Ziele und ihre Effekte auf Verstehensleistungen (verbunden mit weiteren relevanten Variablen im Zusammenspiel) verdienen daher eine gesonderte Beachtung, und das wiederum wirkt deutlich anschlussfähig an Theorien und Modelle der Selbstregulation, zumal es durchaus Vorschläge gibt, Metakognition explizit als Bestandteil der lernbezogenen kognitiven Belastung zu berücksichtigen und zu modellieren (Valcke, 2002).

Hinzu kommt, dass digitale Medien und das Lernen mit ihnen offenbar Eigenheiten aufweisen, welche sich schlecht mit den Prämissen der ursprünglichen Theorie der kognitiven Belastung vereinen lassen. Skulmowski und Xu (2022) machen darauf aufmerksam, dass typische Merkmale des Lernens mit digitalen Medien, etwa ihre Interaktivität oder auch ihre nicht immer optimal wirkende Darbietung im Sinne der Erhöhung sachfremder kognitiver Belastung, trotzdem höhere Lerneffekte bei kognitiven bzw. motivationalen Variablen mit sich bringen. Dies ist im Sinne der strengen Auslegung der Theorie der kognitiven Belastung erwartungswidrig, muss aber dennoch theoretisch erklärbar gemacht werden. Der entscheidende Punkt ist hierbei für die Zwecke dieses Buchs: Digitales Lesen und dessen wegen seiner kognitiven Belastung hervorruhenden vermeintlich lernhinderlichen Merkmale von Kontexten, welche die prototypischen Merkmale dieses Lesens ausmachen, können besseres Verstehen nach sich ziehen. Das ist allerdings daran gebunden, dass ebenfalls lernförderliche Prozesse ausgelöst werden, die mutmaßlich wiederum der Selbstregulation bedürfen.

Etwas generischer ist die Perspektive, die Castro-Alonso et al. (2021) einnehmen. Sie gehen darauf ein, dass nicht sämtliche Texte und Materialien, darunter auch solche in digitalen Umgebungen, immer zwingend optimal im Sinne der Annahmen zur Reduktion sachfremder lernbezogener Belastung sind, etwa weil sie redundante Informationen enthalten. Hier helfen unter anderem und speziell leistungsstärkeren Personen selbstregulierte Lernstrategien. Anders gesagt: Personen haben eine Handhabe, sich selbst nicht-optimal dargebotene Inhalte anzueignen. Dies wird unter dem Stichwort des „Selbstmanagements der kognitiven Belastung“ thematisiert (Mirza et al., 2019; Zhang et al., 2022) und als metakognitiver Kontrollprozess mit der Bedingung verknüpft, unter suboptimalen Bedingungen selbstreguliert vorgehen zu können und zu wollen (Eitel et al., 2020).

Diese drei Aspekte – a) Rolle von Zielen, insbesondere in komplexen Settings, b) scheinbar theoretisch nicht zu erwartende positive Effekte von digitalen Umgebungen mit erhöhter sachfremder kognitiver Belastung und c) das Selbstmanagement der kognitiven Belastung über selbstregulatorische Prozesse der Lernenden – bilden mehr oder minder deutliche Schnittstellen zur Selbstregulation. Auffälligerweise betreffen diese drei Aspekte Prozesse des Lernens.

Es gibt neben der eher aus der Theorie der kognitiven Belastung stammenden Tendenz, Selbstregulation und Arbeitsgedächtnisanforderungen systematischer zu verbinden, auch die umgekehrte Zugangsweise: Die Selbstregulationsforschung greift in dieser Hinsicht ebenfalls auf ein Nachbargebiet zu. Auf einen Zusammenhang von kognitiver Belastung und Selbstregulation machen beispielsweise Wirth et al. (2020) aufmerksam. Sie postulieren in ihrem Mehrebenen-Modell, dass bewusste (im Gegensatz zu automatisch ablaufenden) Ausschnitte der Verarbeitung von Informationen seitens der eigentlichen Lerninhalte, der Strategieanwendung und deren metakognitiver Steuerung eine eigene kognitive Belastung generieren. Diese Inhalte, die im Arbeitsgedächtnis einer bewussten Verarbeitung zur Verfügung stehen, sind dann Gegenstand selbstregulativer Prozesse. In etwas vereinfachter Form ist dieser Grundgedanke in einem weiteren Modell enthalten, welches die Logik von Objekt- und Metaebene (s. Abbildung 20 in Teilkap. 6.1.2.2) aufgreift. Im Modell von de Bruin et al. (2020) ist die kognitive Belastung zwischen Objekt- und Metaebene verortet. Die bewusste Wahrnehmung der Prozesse auf der Objektebene führt über Überwachungsprozesse zu einer metakognitiven Repräsentation von verschiedenen Belastungen, die über Kontroll- und Regulationsprozesse dann reguliert werden können. In diesem Sinne fungiert die kognitive Belastung als eine Art Informationssystem, welches vor allem über die Einschätzung des kognitiven Einsatzes beim Lernen Abgleiche mit Zielen und metakognitivem Wissen auf der Metaebene ermöglicht. Zusammengefasst gesagt: Diese Modelle eint die Idee, dass bewusste Anstrengung bei der Zielerreichung eine Schnittmenge beider Theorien ist.

Daneben gibt es andere Zugänge. Seufert (2018, 2020) führt an, dass selbstregulatorische Prozesse zu einer Erhöhung der kognitiven Belastung beim Lernen führen. Seufert (2018) zufolge gibt es mit Rekurs auf Phasenmodelle der Selbstregulation in jeder der drei Phasen (s. Teilkap. 6.1.1.3) Grund zur Annahme, dass Ziele setzen (in der vorbereitenden Phase), das Überwachen und der Einsatz von Lernstrategien (in der Durchführungsphase) und der abschließende metakognitive Vergleich von Prozessen und Produkten mit den ursprünglichen Zielen eine kognitive Zusatzbelastung generiert. Diese kognitive Belastung bezieht sich dezidiert auf alle drei Formen der kognitiven Belastung. In einer Folgepublikation erweitert Seufert (2020) diesen Gedanken. Es geht ihr nunmehr nicht mehr nur darum, dass Selbstregulation die kognitive Belastung erhöht und verursacht, sondern auch darum, dass kognitive Belastung Selbstregulation erfordert. In diesem zweiten Verständnis reduziert das selbstregulatorische Vorgehen die Informationsmenge gezielt, um mit geschickter Handhabung die tendenziell zur Überlastung führenden Anforderungen, die sich aus den Facetten der kognitiven Belastung ergeben, bewältigen zu können.

Resümierend lässt sich also sagen, dass aus Sicht der pädagogisch-psychologischen Selbstregulationsforschung kognitive Belastung und Selbstregulation ein wechselseitiges Verhältnis aufweisen. Wenn Leserinnen und Leser im digitalen Kontext mit verschiedenen Zielen Aufgaben bewältigen, birgt dies die Gefahr, dass sie kognitiv überlastet werden, weil Inhalte, Vorgehensweisen und die Überwachung und Planung ihre Arbeitsgedächtniskapazitäten derart stark überschreiten, dass ihnen die Selbstregulation dabei hilft, die Limitierungen zu umgehen bzw. diese so zu nutzen, dass die kognitive Belastung nicht zu hoch ausfällt. Sie zahlen dafür aber gleichsam den Preis, dass solche metakognitiven Regulationen und die Auswahl und Durchführung von strategischen Operationen Zusatzlast generieren. Es ist also ein wechselseitiges Verhältnis von Arbeitsgedächtnisbeanspruchung und der Regulation dieser Beanspruchung, das als gemeinsamer Nenner der integrativen bzw. synthetisierenden Theoriearbeit benannt werden kann.

6.1.3.3 Inhibitorische Kontrolle

Die inhibitorische Kontrolle ist ein Sammelbegriff für verschiedene Phänomene. Für die Zwecke des Lesens und der Selbstregulation ist eine Prozessgruppe von Belang, die Nigg (2000, S. 238) als „exekutive Inhibition“ bezeichnet und darunter explizit „Prozesse der intentionalen Kontrolle oder Unterdrückung von Reaktionen zum Erreichen übergeordneter oder längerfristiger Ziele“ subsumiert. Damit ist deutlich erkennbar, dass im Sinne der Selbstregulation nicht nur das Auswählen von geeigneten Vorgehensweisen durch metakognitive Planungsprozesse bei der Strategiewahl angesprochen ist (Winne, 2018a), sondern gemäß Nigg (2000) auch die Vermeidung von Interferenzen und Ideen, welche das Primärziel gefährden. Das ist deshalb so betonenswert, weil komplexe Prozessverbünde wie das digitale Lesen diverse Risiken bergen, beispielsweise durch anklickbare Links, zu durchsuchende Suchmaschinenergebnisse, aber auch durch die aufmerksamkeitsökonomische Funktionsweise von digitalen Medien bzw. Social Media, sodass zielführende Impulskontrollen nötig scheinen (Kozyreva et al., 2020). Zudem wird die Fähigkeit zur Unterdrückung von interferierenden Gedanken im Kontext des Lernens komplexer Sachverhalte als nötig erachtet, z. B. bei der Korrektur von naturwissenschaftlichen Fehlvorstellungen (Mason & Zaccoletti, 2021).

In der normativ und teils kulturpessimistisch geführten Debatte zu positiven und negativen Effekten digitaler Medien (s. z. B. Appel et al., 2020; Appel & Schreiner, 2014; Kozyreva et al., 2020; Loh & Kanai, 2015) spielen Aspekte der inhibitorischen Kontrolle anscheinend insofern eine Rolle, wenn es um die *Ablenkung* geht (Dontre, 2021), also den Ausdruck der misslingenden inhibitorischen Kontrolle beim Wechsel zwischen Aufgaben, etwa dem Lesen von Texten und weiteren

Tätigkeiten wie dem Prüfen von Social-Media-Aktivitäten (Multitasking, s. dazu kritisch Kirschner & Bruyckere, 2017, und empirisch: Jeong & Hwang, 2016; Parry & Le Roux, 2021; Wiradhany & Koerts, 2021; Wiradhany & Nieuwenstein, 2017, s. dazu Teilkap. 6.4.1.1). So zeigt sich beispielsweise in korrelativen Studien, dass die Nutzung von Smartphones (Kates et al., 2018) und die Nutzungshäufigkeit von sozialen Netzwerken – speziell im Fall des Multitaskings – (Huang, 2018; Liu et al., 2017; Marker et al., 2018) mit schlechteren (hoch-)schulischen Leistungen einhergeht. Zur Wahrheit gehört aber auch zweierlei: Korrelationen, zumal auf der Basis von rein querschnittlichen Daten, geben wenig Auskunft über Kausalität und über die Art des Nutzungs- und Lernverhaltens. Und: Es ließen sich zudem leicht *positive* Zusammenhänge zwischen Nutzungsfrequenz und dem Leseverstehen ermitteln (Hu & Yu, 2021; Liu et al., 2017). Letztlich ist, wenn man die Nutzungsfrequenz digitaler Medien alleinig als Indikator für den Mangel inhibitorischer Kontrolle als Indikator nutzt, noch recht wenig ge- und erklärt.

Einen indirekten Erklärungsansatz für die Bedeutsamkeit der inhibitorischen Kontrolle bemühen Liu et al. (2017) im Rahmen ihrer Metaanalyse, um die negativen Zusammenhänge zwischen (hoch-)schulischen Leistungen und Social-Media-Nutzung zu erklären. Sie führen die kognitiven Kosten an, die sie als „*Strafe für das Wechseln*“ bezeichnen. Diese ‚Strafe‘ stellt eine Begleiterscheinung insofern dar, als jeder Wechsel von einer Aktivität zur anderen dazu führt, dass die Person mehr Zeit benötigt, um sich wieder in die Aufgabe hineinzudenken (Zeitverlust), und dass die Fehleranfälligkeit sich durch dieses Vorgehen erhöht (Verlust der Lösungsakkuratheit; s. für eine umfassendere Diskussion Koch et al., 2018). Dies lässt sich ex negativo als Schutzmechanismus einer glückenden inhibitorischen Kontrolle interpretieren: Wer sich vor Ablenkungen oder einem Multitasking schützt, spart Zeit und ist effektiver. Dies wurde bereits am Beispiel des Multitaskings für das Leseverstehen jüngst metaanalytisch bestätigt (Clinton-Lisell, 2021).

6.1.3.4 Kognitive Flexibilität

Die kognitive Flexibilität bezeichnet die Fähigkeit, im Prozess von Aufgabenbearbeitungen zwischen einzelnen Zielen bzw. Aufgaben zielführende Wechsel vorzunehmen. Solche Anforderungen sind beispielsweise prototypisch für komplexe, nicht eindeutig in ihren Lösungsschritten vorgegebene Problemlösungssettings, wozu auch das digitale Lesen gezählt werden kann. Die Fähigkeit, eine solche flexible Herangehensweise auszuführen, gilt als besonders bedeutsam, wenn Personen multiple Dokumente, Repräsentationen und Perspektiven für das Lernen nutzen. Hierbei ist es nötig, die Zusammenhänge zwischen einzelnen Elementen über strategische Vorgehensweisen mit dem Ziel der Inferenzbildung aktiv zu konstruieren und dabei auf neue Informationen flexibel zu reagieren (Cartwright, 2009; Follmer & Sperling, 2020).

Indirekte Unterstützung erhält dies aus einem Zweig der integrativen Lese- und Schreibforschung. In diesem Zweig werden Diskurssynthesen analysiert, also individuelle Nutzungen multipler Dokumente, die jemand lesend rezipiert und diese Dokumente dafür nutzt, ein eigenes kommunikatives Ziel in einem eigenen Text zu erreichen (Nelson, 2001). Es hat sich studienübergreifend ein Muster als günstig erwiesen, das Mateos et al. (2014) als „rekursives Vorgehen“ bezeichnen. Dieses Muster zeichnet sich durch einen hohen Wechsel und eine starke Dynamik von verschiedenen Aktivitäten aus, die nicht-linear erfolgen. Ein solches Vorgehen hat sich in verschiedenen Studien als vorteilhaft insofern herausgestellt, als die Texte der Probandinnen und Probanden auf eine stärkere Integration der gelesenen Dokumente hinweisen (Philipp, 2018a, 2022b).

Deshalb ist – auch wenn kognitive Flexibilität, wie die anderen Komponenten der exekutiven Funktionen, bisher kein eigenständiger elaborierter Bestandteil in aktuellen Theorien des Lesens ist (Butterfuss & Kendeou, 2018) – kognitive Flexibilität eine bedeutsame kognitive Basis für erfolgreiches Leseverstehen (Follmer, 2018). Für die Belange des digitalen Lesens kann das nicht ausdrücklich genug betont werden, weil die kognitive Flexibilität im Gegensatz zur abschirmenden Nutzung des Arbeitsgedächtnisses (durch die inhibitorische Kontrolle) als Gegenspieler dafür dient, der Kontingenz im umfassenden Leseprozess gerecht zu werden und Ad-hoc-Prozesse und Strategien nicht nur zuzulassen, sondern auch gewinnbringend zu verwenden.

6.2 Prozesse und Produkte

Selbstregulation ist ein Vorgehen, das Prozesse zielbezogen moduliert. Damit ist eine Prozessperspektive gleichsam das Rückgrat dieses Konstrukts. Dies schlägt sich nieder in der thematischen Entfaltung in diesem Teilkapitel, welches anders als die zuvorigen Teilkapitel die Produktperspektive etwas in den Hintergrund rückt. Das ist dem erkenntnisleitenden Interesse geschuldet, in diesem Teilkapitel für einige zentrale Punkte herauszuarbeiten, wie glückende Selbstregulation zum (digitalen) Lesen bzw. dem Lernen mit Texten bzw. multimedialen Inhalten modelliert wird.

Der erste Zugang ist vor allem eine theoretische Vertiefung in Form einer genaueren Konsultation eines der Selbstregulationsmodelle aus Teilkapitel [6.1.1](#). Im Teilkapitel [6.2.1](#) wird das COPES-Modell nochmals genauer fokussiert. Es handelt sich um ein dezidiert der Tradition der Metakognitionsforschung verhaftetes Modell. Dass es ausgewählt wurde, hat neben seiner Prominenz in der Wissenschaft auch damit zu tun, dass es wie kaum ein anderes dem Einsatz kognitiver Lernstrategien als metakognitiv gesteuertem Vorgehen folgt, worauf dieses Kapitel fokussiert. Damit geht einher, dass affektive und motivationale Variablen aus rein darstellerischen Gründen in diesem auf Kognitionen stark abzielenden

Buch nicht behandelt werden, da sie unvermeidlich komplexitätssteigernde Folgen hätten. Es ist gänzlich unbenommen und unbestritten, dass sie theoretisch und empirisch bedeutsam sind (Camacho-Morles et al., 2021; Credé & Phillips, 2011; Fong et al., 2021; Howard et al., 2021; Li et al., 2018; Loderer et al., 2020; Schneider & Preckel, 2017; Sun et al., 2021; Toste et al., 2020; Tze et al., 2016).

Der zweite thematische Zugang ist ein selektiver Forschungsüberblick, der sich auf handlungsnah erfasste kognitive und metakognitive Selbstregulationsprozesse bezieht. Im Teilkapitel [6.2.2](#) stehen entsprechend Ergebnisse im Vordergrund, welche dabei helfen, das unübersichtliche Forschungsfeld zumindest grob zu kartieren. Dieser Unübersichtlichkeit, die partiell eine Folge reger Forschungsaktivitäten ist, begegnet die Darstellung, indem sie Akzente setzt und Befunde aus einem Segment von Studien in ihr Zentrum stellt. Es handelt sich um solche Untersuchungen, welche mit besonders validen, die Prozessualität und mit ihr in gewissen Teilen ebenso die Kontextualität in den Blick nehmenden Datengewinnungsverfahren gearbeitet haben (Butler & Cartier, 2018; Cleary & Callan, 2018; Veenman et al., 2006; Winne, 2010). Es mehren sich nämlich die Hinweise, dass ein solches Vorgehen Selbstregulation als dynamisches Konstrukt valider zu messen imstande ist. Dies bildet die Entscheidungsgrundlage, gezielt einen Ausschnitt ergiebiger Befunde besonders zu fokussieren.

Der dritte Teil in der Darstellung ([6.2.3](#)) basiert auf Studien mit qualitativen Zugängen zu Verbalprotokollen beim Lesen (Pressley, 2000). Forschungssynopsen haben aus diversen Einzelstudien einen taxonomischen Katalog von lesestrategischen Vorgehensweisen extrahiert, der sich als Idealtypus des Lesens von Expertinnen und Experten begreifen lässt. Dieses „konstruktiv-responsive Lesen“ (Pressley & Afflerbach, 1995) hat durch die Veränderung des Forschungsfeldes inzwischen eine noch stärkere Ausdifferenzierung erfahren, was die Konstituenten der Trias großer Gruppen von Lesestrategien auf der Objekt- und der Metaebene betrifft (Cho & Afflerbach, 2017). Die empirisch gewonnene Taxonomie von Strategien mit abstrahierbaren Funktionen wie a) dem Finden und Kombinieren von Informationen, b) dem Evaluieren von Informationen und Informationsquellen sowie c) der metakognitiven Regulation des Verstehens erlaubt es, einen Zielhorizont kompetenten Lesens zu beschreiben, der auch und gerade für das digitale Lesen Gültigkeit hat.

6.2.1 COPES als Beispiel für ein Selbstregulationsmodell mit stark metakognitiver Prägung

6.2.1.1 Das Modell und seine Logik im Überblick – einige Vorbemerkungen

Das in der Tradition von Informationsverarbeitungsprozess-Modellen stehende, von Winne und Hadwin (1998) vorgelegte COPES-Modell ist eines der einflussreichsten Modelle in der Selbstregulationsforschung (Alexander et al., 2011; Greene & Azevedo, 2007b; Panadero, 2017) und wird auch für typische Anwendungsbereiche des digitalen Lesens mit dem Herstellen von Aufgabenprodukten als Grundlage genutzt (Raković & Winne, 2022). Zugleich sticht es aus den Modellen deshalb hervor, weil es konsequent eine metakognitive Perspektive verfolgt und dies außerdem in einer (lose) verbundenen Abfolge von Phasen – bzw. wie es im Original heißt: Stufen („Stages“) – modelliert. Motivationale Aspekte werden hingegen subsumiert unter den Bedingungen des Modells, selbst wenn spätere Publikationen die Rolle und die Bedeutung von motivationalen Bestandteilen präzisiert haben (Winne & Hadwin, 2008).

Für das namensgebende Akronym COPES sind fünf Merkmale konstitutiv (s. Tabelle 34), welche wiederum in einem Zusammenspiel dynamisch zusammenwirken und dies primär innerhalb des kognitiven Systems einer Person tun (s. Abbildung 21): Bedingungen (Conditions), Prozesse (Operations), Produkte (Products), Beurteilungen (Evaluations) und Standards. Dieses Zusammenspiel der fünf Merkmale erfolgt in den vier Phasen 1) Aufgabendefinition, 2) Ziele und Plan/Pläne, 3) Durchführung sowie 4) Adaptionen, die in der Abbildung 21 bei den Produkten verortet sind. Für jede der Phasen postuliert das Modell dieselbe Architektur, also ein Zusammenspiel aus den modellimmanenten Komponenten, und Selbiges gilt für die Prozesse, von denen im COPES-Modell angenommen wird, dass es sich um fünf Grundarten handelt, die zusammenwirken (Winne, 2018a). Das bedeutet, es gibt eine sich durch die kognitiven Prozesse und ihre Dynamiken und Rückkopplungen vollziehende Veränderung, die sich idealtypisch als ein Durchlaufen von Phasen beschreiben lässt. Dennoch weisen Winne und Hadwin (1998) explizit darauf hin, dass dies eine Vereinfachung ist und ihr Modell erheblich mehr Kontingenz im Geschehen erklärbar macht, worin strikte Phasenmodelle – selbst solche mit Rückkopplungen im Sinne von Entscheidungsbäumen mit dichotomen metakognitiven Beurteilungen und den aus ihnen resultierenden Handlungsschritten – ihre Grenzen finden.

Tabelle 34: Phasen und definierte COPES-Merkmale

COPES-Merk- male	Definitionen	Phase 1: Aufgabendefi- nition	Phase 2: Ziele und Plan/ Pläne	Phase 3: Durchführung	(Fakultative) Phase 4: Adaptionen
Bedingungen (Conditions)	Bedingungen, die beein- flussen, wie die Aufgabe in Angriff genommen wird, einschließlich konditio- nalem Wissen („Wenn“ in „Wenn-dann“-Regeln, die filtern, welche Operationen ausgeführt werden)	• Interesse • Zielorientierung • Lernstrategieportoire • zeitliche Beschränkungen • verfügbare Ressourcen • Kenntnisse über Taktiken • Aufgabenwissen • Fachwissen	Bedingungen aus Phase 1 + Produkte aus Phase 1	Bedingungen aus Phasen 1 und 2 + Produkte aus Phase 2	Bedingungen aus Phasen 1–3 + Produkte aus Phasen 1–3
Operationen (Operations)*	kognitive Prozesse, Tak- tiken, und Strategien, die eine lernende Person zur Bewältigung der Aufgabe einsetzt	• Recherchieren • Überwachen (einschließlich Taktiken wie Selbstbefragung, Aufzeichnung von Ähnlichkeiten und Unterschieden) • Zusammenfügen (einschließlich Taktiken wie Ausarbeiten, Integrieren) • Wiederholen (einschließlich Taktiken wie Kopieren von Notizen, erneutes Lesen) • Übersetzen (einschließlich Taktiken wie das Anfertigen von Diagrammen und die Verwendung von Anfangsbuch- staben-Mnemotechniken)			
Produkte (Products)	Informationen, die mittels Operationen hergestellt wurden und sowohl interne Produkte als auch externe Produkte sein können	Wahrnehmungen der Aufgabenerfordernisse (COPES-Merkmale)	• Aufgabenstandards aus Phase 1 oder individuell verändert aufgrund indi- vidueller Merkmale • Plan, die Lerntaktiken zu koordinieren (im Sinne von Strategien) • Standards für die Anwen- dung von Taktiken (z. B. Geschwindigkeit)	• Spuren der angewendeten Lerntaktiken (z. B. Rand- notizen, Concept-Maps, Antworten auf Fragen) • neues/restrukturiertes Fachwissen	Aktualisierungen von • COPES-Merkmalen, die Aufgaben unterscheiden, • Motivation und Überzeu- gungen, • konditionalem Wissen und Lerntaktiken, • Standards für die Beurtei- lung von Qualitäten von Taktiken

COPES-Merkmale	Definitionen	Phase 1: Aufgabedefinition	Phase 2: Ziele und Plan/pläne	Phase 3: Durchführung	(Fakultative) Phase 4: Adaptionen
Evaluationen (Evaluations)	Feedback zu Produkten, die entweder intern von den Lernenden oder von einer (oder mehreren) externen Quelle(n) stammen	Einschätzung von • Aufgabenverständnis, • COPES-Merkmale der Aufgabe; Vergleiche mit anderen Informationsquellen	Urteile über • Komplexität/Schwierigkeit, • Anreiz des Ziels, • Fähigkeit, den Plan auszuführen, • erforderlichen Aufwand	Urteile über • Lernen (Kalibrierung), • Nutzen der Taktiken, • Effektivität, • Attribuierungen	Urteile über • Abstand zwischen früherer Version und Anpassung, • vorhergesagte Effekte der Anpassung; • weitere Einschätzungen, die auf zukünftige Aufgaben warten müssen
Standards (Standards)	Kriterien, anhand derer die Produkte überprüft werden	• Beurteilungskriterien • bisherige Leistungen	• Standards und Produkte aus Phase 1 • Aufwand/Nutzen-Schwellenwerte • motivationale Orientierung	• Standards und Produkte aus Phase 2 • Standards sowohl auf der Objekt- als auch auf der Metaebene	Standards und Produkte aus den Phasen 1–3

(Quelle: leicht modifizierte Übersetzung und Ergänzung von Winne & Hadwin, 1998, S. 280; * Winne, 2001, geht davon aus, dass die Prozesse in jeder Phase als übergeordnete Gruppen von Kognitionen prinzipiell zum Tragen kommen und benennt unter dem Akronym SMART fünf Prozesse: Searching (Recherchieren), Monitoring (Überwachen), Assembling (Zusammenfügen), Rehearsing (Wiederholen), and Translating (Übersetzen) – aus diesem Grund sind die Prozesse auch über alle Phasen hinweg gleichermaßen aufgeführt, sie unterscheiden sich in ihrer konkreten Ausprägung und in ihrem Verhältnis in den spezifischen Aufgaben)

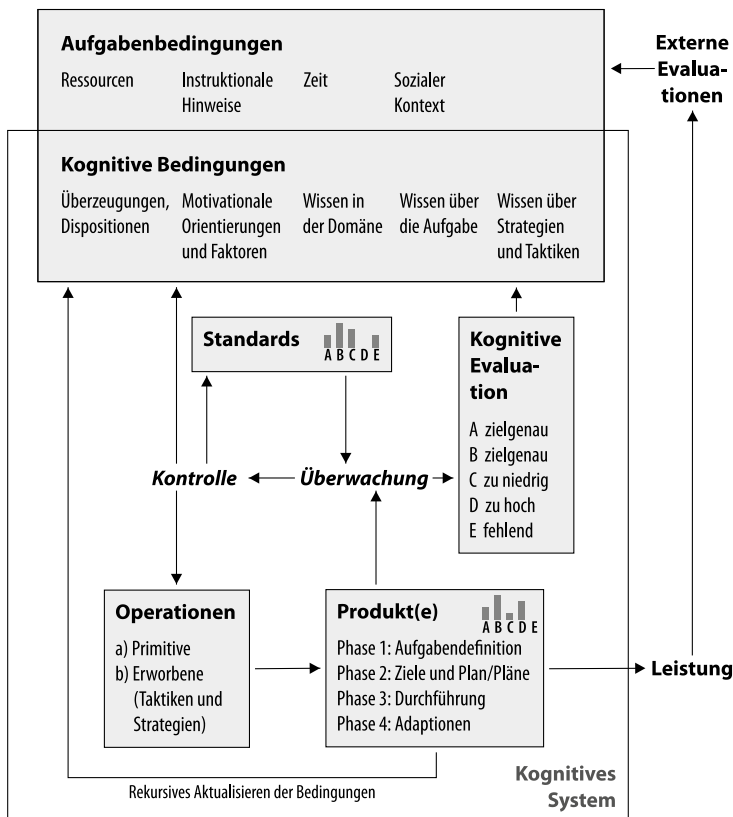


Abbildung 21: Das COPES-Modell selbstregulierten Lernens mit seinem Fokus auf metakognitive Kontrolle und Überwachung in vier Phasen (Quelle: leicht modifizierte Übersetzung von Winne & Hadwin, 1998, S. 282)

Wegen seiner dezidierten Verortung in Modellen der Informationsverarbeitung (Winne, 2001) fügen sich die Merkmale von COPES relativ nahtlos ein in die Logik von Meta- und Objektebene nebst verbindenden Prozessen. Deshalb ist es auch nicht verwunderlich, dass im Modell als zentrale metakognitive Prozesse das *Kontrollieren* und das *Überwachen* auftauchen. Mittels dieser beiden Prozesse erfolgt ein steter Abgleich (im Modell als kognitive Evaluation bezeichnet) zwischen den Produkten auf der Objektebene und den Standards auf der Metaebene als Bezugsebenen und -punkten. Diese Evaluation dient als internes Feedback für die Person. Je nach Verhältnis des wahrgenommenen Ist-Zustandes im Produkt und des angestrebten Soll-Zustandes erfolgt qua *Überwachung* eine Modifikation bzw. eine Weiterführung der Operationen qua *Kontrolle*. In Abhängigkeit von der Komplexität des Standards wird zudem angenommen, dass entsprechend komplexere Strategien zum Einsatz gelangen (Winne, 2018b).

Die Phasenhaftigkeit ist im Modell in Abbildung 21 wie erwähnt lediglich im Bestandteil der Produkte erkennbar und damit in der grafischen Darstellung eher implizit angelegt; das Modell selbst ist in dieser Hinsicht wesentlich deutlicher in seiner Beschreibung. Wenn zudem von vier Phasen die Rede ist, dann ist dies nicht als jederzeit strikt zu durchlaufende Sequenz zu denken, weil beispielsweise Adaptionen (Phase 4) prinzipiell jederzeit möglich sind, z. B. bei der Formulierung von Zielen und dem Planen (in Phase 2), ohne dass es bereits zu sichtbaren Manifestationen im Aufgabenkontext kommen muss. Zusätzlich sind persönliche Merkmale von Belang: Wer mit einem Aufgabentypus umfassend vertraut ist, also eine breite Wissensbasis in puncto *Aufgabenwissen* mit sich bringt, oder wer ein starkes Fundament an *Fachwissen* aufweist, muss beispielsweise nicht jederzeit eine detaillierte Aufgabenanalyse und Planung (Phasen 1 und 2) absolvieren, um trotzdem erfolgreich zu sein. Damit sind partiell auch Aspekte angesprochen, nach denen nicht alle Prozesse bei Aufgabenbewältigungen zwingend einer bewussten metakognitiven Überwachung und regulierenden Kontrolle unterliegen müssen, etwa weil sie stark automatisiert sind (Panadero, 2017; Wirth et al., 2020).

Nach all diesen Vorbemerkungen zum COPES-Modell werden die idealtypisch vereinfachte Phasenhaftigkeit mitsamt ihren Komponenten und ihrem Zusammenspiel nachfolgend gemäß Winne (2001) sowie Winne und Hadwin (1998) entfaltet.

6.2.1.2 Phase 1: Aufgabendefinition

Diese erste Phase dient dazu, dass eine Leserin bzw. ein Leser die Merkmale der Aufgabe und ihrer ex- und internen *Bedingungen* möglichst umfassend repräsentiert. Dabei geht es nicht nur um die Verarbeitung der explizit in der Aufgabenstellung formulierten Anforderungen, sondern auch um die (eigensinnige) Interpretation aufgrund verschiedener im Langzeitgedächtnis verfügbarer Wissensbestände (Aufgabenwissen, Fachwissen, konditionales Strategiewissen (im Modell: Wissen über Taktiken und Lernstrategierepertoire) und – von Winne und Hadwin (1998) eher en passant bei den empirischen Befunden eingeführt – die eigenen epistemischen Überzeugungen), die Klärung der eigenen Motivation und der externen Rahmenbedingungen. Eine solch umfassende Klärung durch planerische *Prozesse* dient dazu, metakognitiv die Aufgabe sowie die in- und externen Informationsquellen zu analysieren, und ist damit den Kontrollprozessen zuordenbar, die von der Metaebene ihren Ausgangspunkt nehmen. Das *Produkt*, in diesem Fall das interne Produkt, ist gemäß dem Modell eine umfassende Repräsentation der Aufgabenerfordernisse, mithin ein Teil des Aufgabenmodells (s. Teilkap. [2.1.1](#)). Als Bemessungsgrundlage für dieses Produkt fungieren im Sinne der *Standards* externe Kriterien (etwa die Kriterien für Leistungsbeurteilungen)

oder interne metakognitiv repräsentierte Erfahrungen mit einem Aufgabentypus, und beurteilt wird im Sinne der *Evaluation* die (Un-)Vollständigkeit der Parameter im Produkt.

Da die Aufgaben in dieser ersten Phase als eine zentrale Komponente angenommen werden, ist eine Präzisierung in dem Modell hervorzuheben, die Winne (2001) vorgenommen hat. Er differenziert zwischen zwei Arten von Aufgaben, nämlich typischen und untypischen. Das ist für das selbstregulierte Lernen insofern bedeutsam, als bei typischen Aufgaben angenommen werden darf, dass sie bekannte Standards und auch metakognitive Wissensbestände betreffen, während unbekannte Aufgaben aufgrund ihrer Klärungsbedürftigkeit in ihren Anforderungen und Erfolgskriterien eine höhere Unschärfe aufweisen, welche den Schwierigkeitsgrad erhöht. Anders gesagt: Unbekannte bzw. untypische Aufgaben erfordern mehr aufgabenanalytische Vorgänge. Das ist für das digitale Lesen von Belang, da ihm eine gewisse Offenheit wesenseigen ist, die sich vor allem aus dem fortlaufenden Prozess der anzusteuern und lesend zu verwertenden Texte ergibt, welche über Rückkopplungen auf die Aufgabendefinition dynamisch Einfluss nehmen dürften und dadurch zu besser kalibrierten Repräsentationen der Aufgabenerfordernisse führen (Schoor et al., 2021).

Ein fiktives Beispiel für diese erste Phase besteht im Arbeitsauftrag, für einen eigenen Text Informationen zur Sicherheit von Impfstoffen gegen schwere Ausbrüche von Covid-19 zu finden, um auf dieser Basis einen kurzen schriftlichen Bericht zu erstellen. Der Arbeitsauftrag soll lauten: „Welche Vor- und Nachteile haben die zur Verfügung stehenden Impfstoffe für welche Personen?“ In diesem Fall muss eine Person mit ihrem Gedächtnis abgleichen, ob sie mit Aufgaben wie der aktuellen bereits Erfahrungen hat, was sie über das Impfen im Allgemeinen und über das Impfen im Falle von Covid-19 im Besonderen bereits weiß und welche Merkmale die Textsorte Bericht aufweist. Die Person muss zudem prüfen, ob sie das Thema interessant findet, welches Ziel sie mit der Erledigung dieses Auftrags verbindet etc. Diese Klärung der Bedingungen wird flankiert vom Prüfen, nach welchen Kriterien eine solche Leistung als geglückt oder misslungen gilt, indem die Person über kognitive Analysen der Aufgabenstellung beispielsweise die Berücksichtigung von verschiedenen Perspektiven im eigenen schriftlichen Bericht als Maß für Ausgewogenheit erkennt – und damit die Notwendigkeit, ausreichend umfassende Informationen beizusteuern. Prozesse, die aus ihnen resultierenden Produkte, die Standards und die Klärung der Parameter als Grundlage für die Evaluation interagieren damit stark. Hierbei gilt bereits die grundsätzliche Möglichkeit, dass es zu Adaptionen kommt, indem etwa die Lückenhaftigkeit des Verstehens der Aufgabenerfordernisse der Person auffällt, die zu weiteren Aufgabenanalysen führt, damit die Aufgabendefinition erfolgreich ist.

6.2.1.3 Phase 2: Ziele und Plan/Pläne

Die zweite Phase, in der eine lesende Person Ziele und Pläne generiert, wird in dem Modell von der Aufgabenanalyse und -definition getrennt. Dies ist deshalb bedeutend, weil dies den Personen eine Autonomie zugesteht, ihre eigenen Ziele auf der Basis ihres Aufgabenverständnisses und ihrer eigenen Ressourcen festzulegen und für sie adäquat wirkende Vorgehensweisen (*Prozesse*) im Sinne von Kontrollprozessen zu planen. Dies steht in Einklang mit dem, wie das Aufgabenmodell in Teilkapitel [2.1.1](#) dargestellt wurde (Rouet et al., 2017). Die *Standards* im Modell fungieren dabei als die metakognitive Repräsentation der Ziele (die Produkte aus Phase 1, die jetzt als *Bedingungen* fungieren), die über die später auszuführenden *Prozesse* erreicht werden und in definierte Produkte münden sollen. Daneben speisen sich die *Evaluationen* auch aus den motivationalen Überlegungen einer Person, die zusätzlich den Aufwand, den Nutzen, den Anreiz und die Schwierigkeit der Aufgabe abschätzt – also damit ganz klassische motivationale Komponenten der Selbstregulation in Anschlag bringt. Am Ende dieser Phase steht dann als *Produkt* ein handlungsleitender Plan, der in der nächsten Phase der Durchführung als Ressource fungiert.

Im Fall des Beispiels eines Berichts über die Vor- und Nachteile der Covid-19-Impfstoffe besteht in dieser Phase das Vorgehen darin, das Produkt – in diesem Fall vor allem das interne Produkt im Sinne eines Plans für das Vorgehen – erstellt zu haben und dafür die metakognitiven Repräsentationen als Basis dafür zu nutzen, zu diesem Plan zu gelangen. Hierbei ist idealerweise dieses Produkt davon gekennzeichnet, dass die lesende Person vorab definiert, wie sie die Informationen über die Impfstoffe aufbereitet, also etwa eine Tabelle mit den Dimensionen des Vergleichs einzelner Impfstoffe oder Impfstoffgruppen herzustellen gedenkt, wie viel Zeit sie sich nehmen will, wie viele Vor- und Nachteile sie pro Impfstoff(-kategorie) zu recherchieren plant. Da es bei dem Planen im Sinne von Strategien als übergeordnete Handlungspläne darum geht, die Sequenz von eigenen Handlungen sowohl prospektiv als auch adaptiv-situativ festzulegen, gehört zum Planen, dass man das übergeordnete Projekt des Bericht-Verfassens in Teilschritte bzw. Arbeitspakete zergliedert. Dazu zählt, Recherche-, Aufbereitungs- und Schreibphasen vorgängig zu definieren, die Planungen können selbstredend kleinteiliger sein, etwa indem man sich vornimmt, die genauere Darstellung davon abhängig zu machen, als wie disparat oder konvergent sich die Inhalte gestalten, die man sich erst lesend und recherchierend erschließt. Selbst darüber hinaus stattfindende kleinteiligere Planungen, dass man sich auf den Seiten des Paul-Ehrlich-Instituts oder der Weltgesundheitsorganisation informieren will und sich nur auf die aktuellsten wissenschaftlichen Erkenntnisse zu stützen gedenkt, sind planerische Aktivitäten. Solche verästelten und konkretisierenden Überlegungen zur individuellen Bewerkstellung des Hauptziels des Bericht-Schreibens bilden

den Kern der Kontrollprozesse von der Metaebene auf die Objektebene, welche dann in der darauffolgenden Phase der Umsetzung harren.

6.2.1.4 Phase 3: Durchführung

In der Phase der Durchführung mit den akkumulierten *Bedingungen* aus den vorherigen beiden Phasen spielen die *Prozesse* eine zentrale Rolle, da sie – vor allem die umgesetzten Strategien – in interne kognitive oder externe *Produkte* münden, die mit den *Standards* verglichen werden und damit zum Gegenstand von *Evaluationen* und also Feedbackschleifen avancieren. Als Standards fungieren dabei sowohl die Ziele auf der Objektebene des entstehenden Produkts als auch jene auf der Metaebene zum zielführenden Vorgehen.

Dieser fluide Gesamtprozess von strategischen Prozessen und Produkten, die über das angestrebte Soll in den Standards definiert sind, führt zu einer Aktualisierung und einem Fortschritt bei der Aufgabenbearbeitung, wobei neben wenig Selbstkorrekturen auch Fehlanwendungen und deren Korrektur erfolgen können und sogar ein Abbruch einer als in der Durchführung zu schwierigen Aufgabe Szenarien sein können. In dieser Phase mit ihrer insgesamt hohen Dynamiken von „Events“ genannten Abfolgen von Prozessen (Winne, 2010) ist das Zusammenspiel von Kontroll- und Überwachungsprozessen besonders zu betonen. Denn es gilt, die gefassten Pläne auf der Metaebene auf der Objektebene exekutiv umzusetzen und den Fortschritt zu überwachen.

Im Falle des Schreibens des Berichts spiegelt sich dies darin wider, dass entstehende Notizen, Textteile, aber auch das sich allmählich anreichernde und verändernde Wissen von der lesenden Person mit einem Soll an als nötig erachtetem Verstehen abgeglichen wird. Dabei ist im Prozess des Recherchierens, Lesens und Schreibens an praktisch jeder Stelle eine Wenn-dann-Sonst-Logik möglich, z. B. dass unbekannte medizinische Fachausdrücke wie „mRNA“ das Verstehen erschweren und zu Recherchen nach eben diesen Begriffen führen, ehe das Lesen weitergehen kann. Diverse Prüfungsvorgänge – Hilft mir diese Information, um das übergeordnete Ziel des Berichts zu erreichen? Ist das ein vernachlässigenswertes Detail? Habe ich den Text verstanden oder muss ich unverstandene Texte erneut lesen? – sind Ausdruck von Überwachungsprozessen. In jenen werden Wissen, Strategieinsatz und entstehendes Text(zwischen)produkt evaluiert, um sich dem erwünschten Zielzustand anzunähern. Das kann das Verstehen von Konzepten wie mRNA sein, das Anfertigen von Notizen für den Text etc. Das bedeutet: In- und externe Produkte in ihrer Genese während der Durchführungsphase sind Gegenstand kognitiver Prozesse und deren metakognitiver Regulation. Damit geht bereits einher, dass Adaptionen nicht nur denkbar, sondern sogar wahrscheinlich sind. Dies leitet über zur vierten Phase.

6.2.1.5 (Fakultative) Phase 4: Adaptionen

Die Adaptionen sind in der Logik des COPES-Modell eine optionale Phase. Winne und Hadwin (1998) skizzieren diese Phase nur knapp als eine von zwei Varianten, nämlich dass es zu Adaptionen in allen Komponenten kommen kann – bis hin zu umfassenden Rekonfigurationen –, die sich auf die aktuelle Aufgabe beziehen. Eine zweite Variante kann zu einem Lerneffekt insofern führen, dass die Erfahrungen mit der Aufgabenbearbeitung die kognitiven Bedingungen verändern, also das Wissen modifizieren, z. B. indem neue, erfolgreiche Strategien hinzukommen. Winne (2001) ergänzt, dass Lernende im Sinne der Adaptionen eigenständig Taktiken zu Strategien kombinieren können. Diese letztgenannten Adaptionen sind als Lernprozesse zu werten, da sie das metakognitive Wissen und damit das lernstrategische Repertoire erweitern.

Wer auf der Metaebene bei der Aufgabenbearbeitung wie dem Verfassen eines Berichts über Covid-19-Impfungen umfassende metakognitive Erfahrungen gesammelt hat, der kann auf dieser Basis seine Wissensbestände verändern. So ist es denkbar, dass die Person erkannt hat, dass sich bislang bewährte Vorgehensweisen beim Recherchieren, beim Lesen oder beim Schreiben nicht als angemessen erwiesen haben oder dass andere Strategien von Erfolg gekrönt waren. Auf dieser Grundlage erweitert sich ein Strategierepertoire, welches dann bei weiteren Aufgaben ähnlicher Art wieder zum Einsatz gelangen kann. Auch die Bemessungsgrundlage für Evaluationen im Sinne der Standards kann durch die Erfahrungen (re-)kalibriert werden. Insofern besteht durch die Adaption die Chance, die Fähigkeiten zur metakognitiven Selbstregulation in Zukunft auszubauen.

6.2.1.6 Ein COPES-kompatibles Beispiel

Die Forschung zu Prozessen, wie sie im COPES-Modell modelliert werden, gestaltet sich komplex, weil COPES eine hohe Interaktion von Prozessen bei Offenheit und Kontingenz im Gesamtprozess ausdrücklich zulässt. Deshalb ist eine der großen Herausforderungen in der Empirie, wie man Selbstregulation definiert, operationalisiert und analysiert (Azevedo, 2009). Eine Möglichkeit, Selbstregulation als Gesamtprozess in einer Abfolge von günstigen oder weniger günstigen Verbindungen von (meta-)kognitiven Prozessen zu erforschen, besteht im Data-Mining-Ansatz, bei dem es darum geht, in Datensätzen durch statistische Verfahren Zusammenhänge und Muster zu erkennen (Biswas et al., 2018). Dieses Verfahren bzw. ein Verfahren aus dieser Gruppe verschiedener Ansätze liegt dem folgenden Beispiel zugrunde (s. Kasten „Illustrierendes Beispiel 24“). Dieses Beispiel verdeutlicht, dass Leistungsgruppen sich in ihren Vorgehensweisen vor

allem darin unterscheiden, wie tiefen- oder oberflächenorientierte Leseprozesse entweder stark metakognitiv kontrolliert und überwacht werden oder aber eher lose mit metakognitiver Selbstregulation korrespondieren.

Illustrierendes Beispiel 24

Auf die Sequenz kommt es an – wenn Metakognition und tiefes Leseverstehen bessere Leseleistungen erklärbar machen

A) Was war das Erkenntnisinteresse?

Das Experiment widmete sich der Frage, welche prozessnah erfassten kognitiven und metakognitiven Prozesse in zwei Leistungsgruppen sich in ihrer Auftretenshäufigkeit und vor allem in ihrer Abfolge unterscheiden. Das heißt: Die zeitliche Reihenfolge von Strategien spielte hier eine große Rolle (Bannert et al., 2014).

B) Wer wurde untersucht?

Das Sample bildeten 38 deutsche Studierende bildungswissenschaftlicher Fächer. Diese Personen waren im Durchschnitt knapp 24 Jahre alt, die Frauenquote lag bei 84 Prozent.

C) Welche Instrumente kamen zum Einsatz?

Die Testpersonen lasen einen auf 44 Einzelwebsites verteilten Text in einem Browser, dieser Hypertext bestand aus fast 12.500 Wörtern, 19 Diagrammen bzw. Bildern, vier Tabellen und enthielt 234 Links. Der Hypertext konnte auf dreierlei Art navigiert werden: über die erwähnten Links, über Vorwärts- und Rückwärts-Schaltflächen zu Beginn und am Ende jeder einzelnen Website und zu guter Letzt über eine hierarchisch organisierte Darbietung in einem Navigationsmenü. Der interessierende Textteil, für den die Testpersonen lesend tätig werden mussten, um den Inhalt zu lernen, bestand aus neun Einzelwebsites, welche 2.300 Wörter, zwei Bilder, drei Tabellen und 56 Hyperlinks enthielten. Es war Aufgabe der Studierenden, innerhalb von 30 Minuten spezifische Konzepte und Prinzipien zum Thema operante Konditionierung so zu lernen, dass sie danach anderen Studierenden die Inhalte erklären konnten.

Während der Aufgabenbearbeitung gaben die Testpersonen ihre Prozesse per Lautdenk-Protokoll kund. Diese wurden aufgenommen, transkribiert und prioritär die metakognitiven und kognitiven Strategien kodiert. Zusätzlich wurden die Personen nach der Aufgabenbearbeitung getestet und anhand des anspruchsvollsten Tests, welcher der Aufgabenstellung für das Lesen am ehesten entsprach, wurden zwei Extremgruppen gebildet: die leistungsstärksten fünf und die leistungsschwächsten sechs Personen. Deren kodierte Prozesse wurden mittels algorithmischer Datenauswertung danach reduktiv analysiert, ob es Differenzen in der Häufigkeit und in der Abfolge von den empirisch wichtigsten Prozessgruppen gab. Ganz besonders interessierte man sich also für die temporale Abfolge und Auftretensmuster von Prozessen.

D) Was waren die Hauptbefunde?

Ein erstes wichtiges Resultat der Studie bestand darin, dass vor allem die metakognitiven Prozesse während der Aufgabenbearbeitung mit den später in drei unterschiedlich schwierigen Tests positiv korrelierten. Daneben zeigten die Extremgruppenvergleiche distinkte Vorgänge in der allgemeinen Aufgabenbearbeitung. Diese manifestieren sich weniger in der durchschnittlichen Auftretenshäufigkeit der einzelnen Prozesse, sondern vielmehr in deren Sequenz und der per Korrelationskoeffizient ausgewiesenen Nähe zweier Prozesse (s. Abbildung 22).

Generell waren die fünf empirisch wichtigsten Prozesskategorien in den beiden Gruppen davon gekennzeichnet, dass es sehr wahrscheinlich war, dass eine Episode aus der Kategorie von einer weiteren derselben Kategorie gefolgt wurde, was sich in den hohen Korrelationskoeffizienten in den die Wiederholung anzeigenden Rückwärtspfeilen jeweils außen manifestiert.

Einzelne Prozesse spielten für das gesamte Muster keine statistische Rolle, doch gab es in den beiden Gruppen Unterschiede: Leistungsstarke Personen wiederholten im Gegensatz zu den leistungsschwachen Personen beim Lesen weniger die Inhalte (s. mittlerer Kasten rechts in Abbildung 22). Dafür nutzen sie wiederum stärker die Fortschrittsüberwachung (s. unterer Kasten links in Abbildung 22). Doch nicht nur hierin bestehen Differenzen zwischen den Gruppen:

- Generell zeichnet sich das Vorgehen der *leistungsstärksten Personen* dadurch aus, dass die metakognitiven Aktivitäten – das Analysieren als ein Konglomerat von drei Prozessen, das Überwachen sowie die Überwachung – miteinander zusammenhängen, teils über Rückkopplungen, also den Wechsel von der einen in die andere Prozessgruppe. Vor allem aber bilden Metakognitionen die Klammer einer auf Tiefenverarbeitung, d. h. auf organisierende und elaborierende Leseprozessen basierenden Leseaktivität, deren Fortschritt mit der Zielprüfung korrespondiert.
- Die *schwächeren Leserinnen und Leser* hatten ein stärker unidirektionales Vorgehen mit weniger Rückkopplungswahrscheinlichkeiten. Metakognitives Überwachen bezieht sich in der Gruppe leistungsschwächerer Personen zuvorderst auf die Leseprozesse, welche im Gesamt eher oberflächenorientiert (wiederholendes Lesen) sind bzw. sich im Falle tieferer Leseverstehensprozesse nur dann (unidirektional) zeigen, wenn zuvor Textteile mündlich repetiert worden waren. Die metakognitive Überwachung erstreckt sich in dieser Gruppe primär auf das Lesen selbst, war also eher unverbunden mit metakognitiven Zielen bzw. bezog sich mittelbar auf die Leseprozesse.

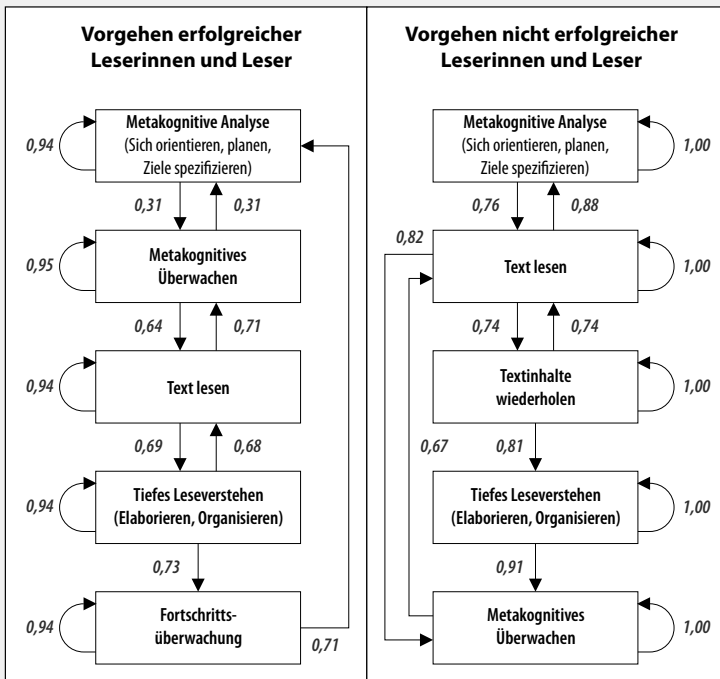


Abbildung 22: Zwei Muster bei der Aufgabenbearbeitung beim Lesen digitaler Texte (Quelle: leicht modifizierte Darstellung von Bannert et al., 2014, S. 175; Dezimalzahlen geben Korrelationskoeffizienten des Zusammenhangs von Auftretensmustern an; Pfeile stehen für entweder wiederholtes Auftreten von distinkten Prozessen derselben Gruppe (jeweils mit Halbkreis versehen) bzw. zwischen den einzelnen Gruppen bei vertikalen Verbindungen zwischen den Prozessen in den Kästen)

E) Diskussion

Die Studie verdeutlicht, dass die Zusammenhänge zwischen den Prozessen auf der Objekt- und auf der Metaebene sich je nach Gruppe unterscheiden. Metakognitive Steuerungsprozesse sind stärker und durchaus bidirektional mit dem tiefen Leseverstehen verknüpft, wenn bessere Leserinnen und Leser aktiv sind. Demgegenüber neigten leistungsschwächere Personen zu einem oberflächlichen Lesen, das zudem wenig metakognitiv gesteuert wirkt. Dieses Ergebnis ist deshalb so wichtig, weil die Quantität der einzelnen kognitiven Prozesse in den Gruppen weniger zu trennen scheint als vielmehr die Verbindung zwischen diesen Prozessen. Leistungsstärkere Individuen haben im Licht der Ergebnisse nämlich einen fluiden Übergang zwischen Kognition und Metakognition.

6.2.2 Empirische Bedeutsamkeit selbstregulatorischer Fähigkeiten für das Lernen und das verstehende Lesen

6.2.2.1 Schlaglichtartige produktbezogene Befunde zur Wichtigkeit kognitiver und metakognitiver Strategien für das Lesen und das Lernen mit Texten

Die empirische Wichtigkeit von Strategien ist in diversen, in ihrer Gesamtheit kaum noch überblickbaren Einzelstudien erforscht worden. Gleich mehrere, diese Einzelstudien quantitativ bündelnden Metaanalysen haben auf breiter empirischer Basis unterstrichen, dass lernstrategische Vorgehensweisen leistungsförderlich sind, darunter im Leseverstehen:

- Diese Metaanalysen aus der *Grundlagenforschung* zum Zusammenhang von lernstrategischen Vorgehensweisen und Leistungen zeigen, dass verschiedene Arten von Strategien positiv mit Leistungen zusammenhängen (Broadbent & Poon, 2015; Burnette et al., 2013; Credé & Phillips, 2011; Danişman, 2017; Dent & Koenka, 2016; Dinsmore et al., 2018; Fong et al., 2021; Hattie & Donoghue, 2016; Li et al., 2018; Ohtani & Hisasaka, 2018; Richardson et al., 2012; Schneider & Preckel, 2017; Sun et al., 2021; Tise et al., 2019).
- Die *angewandte Forschung* zur Förderung selbstregulatorischer Fähigkeiten hat in den vergangenen Jahren ebenfalls vielfältige Ergebnisse zu den Auswirkungen gezielter Vermittlung und Nutzung der Selbstregulation erbracht und die Effektivität in variierendem Ausmaß demonstriert (Boer et al., 2018; Chen, 2022; Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014; Elhusseini et al., 2022; St. Graham & Hebert, 2011; Guo, 2022b; Mentel et al., 2022; Okkinga et al., 2018; Perry et al., 2012; Plonsky, 2011; Popham et al., 2018; Sanders et al., 2019; Sencibaugh, 2007; Theobald, 2021; Yapp et al., 2021; Zheng, 2016).

Beide Forschungszweige weisen darauf hin, dass selbstregulatorische Fähigkeiten in einer Vielzahl von Situationen und Domänen leistungsförderlich sind, was sich in den Produkten des Lernens niederschlägt. Das Entscheidende hieran ist der Plural: Lernen bzw. lernbasierte Leistungen korrespondieren mit *verschiedenen Strategien*, über deren wechselseitiges Zusammenwirken die bisherigen Metaanalysen in ihrer eher lineare Zusammenhänge fokussierenden Zugangsweise freilich nichts aussagen bzw. aussagen können.

Die Theorien zur Selbstregulation, die im Teilkapitel [6.1.1](#) synoptisch dargestellt wurden und welche eine Trennung von Objekt- und Metaebene nebst dynamischen Prozessen der Top-down-Kontrolle und der Bottom-up-Überwachung modellieren ([6.1.2](#)), implizieren eines deutlich. *Selbstregulation ist mehr als die Summe und die Anzahl einzelner Strategien, vielmehr bildet die metakognitiv*

konzertierte Koordination von diversen zielführenden Aktivitäten den Gradmesser für Erfolg und Misserfolg. Vor allem gestaltet sich das strategische Handeln, welches Winne (2001) für eine Vorgehensweise reserviert, die sich durch umfassende metakognitive prozedurale und konditionale Strategiewissensbestände in ihrer Anwendung in situ auszeichnet, in der Umsetzung qualitativ anspruchsvoll. Ein solches Vorgehen bemisst sich damit weniger an der Zahl der gezeigten Strategien denn an zwei weiteren Merkmalen. Das sind auf der einen Seite Wechsel zwischen einzelnen Strategien (Azevedo & Witherspoon, 2008; Witherspoon et al., 2008) und auf der anderen Seite die Güte und Qualität der adaptiven Strategiauswahl und -nutzung (Dinsmore, 2017b).

Um diese Adaptivität einmal plastischer herauszuarbeiten, stellt der Kasten „Illustrierendes Beispiel 25“ eine Studie vor. In ihr wurde die situative Reaktion auf Verstehensprobleme zum Gegenstand der Analyse gemacht und dadurch das Zusammenspiel von Meta- und Objektebene bei den Prozessen und die Effekte auf die Produkte der Wissensaneignung durch das Lesen konturiert.

Illustrierendes Beispiel 25

Statik schadet – vom Vorteil der Strategiewechsel für den Wissenszuwachs

A) Was war das Erkenntnisinteresse?

Die Studie widmete sich der Frage, ob Personen beim Lesen digitaler Texte ihr Verstehen überwachen und ob metakognitive Überwachungsprozesse mit strategischen Verhaltensänderungen korrespondieren (Binbasaran Tüysüzoğlu & Greene, 2015). Konkret wollte man wissen, ob Wechsel im Vorgehen nach einem erkannten Verstehensproblem eine Verhaltensänderung nach sich zogen und ob die Häufigkeit des Vorkommens mit besseren Verstehensleistungen korrespondiert.

B) Wer wurde untersucht?

Insgesamt 170 Studierende aus den USA bildungswissenschaftlicher Fächer nahmen teil. Davon waren 103 Frauen, das Durchschnittsalter betrug knapp 20 Jahre. Von diesen Personen wurden absichtsvoll nur diejenigen fokussiert, die in ihren Lautdenk-Protokollen mindestens einmal ein negatives Urteil über ihr Verstehen abgegeben hatten. Ausgewertet wurden darum 81 Lautdenk-Protokolle, also etwas weniger als die Hälfte der Personendatensätze.

C) Welche Instrumente kamen zum Einsatz?

Die Personen schrieben zweimal einen kurzen Text, in dem sie über die Funktionsweise des menschlichen Blutkreislaufs berichten sollten. Diesen Text verfassten sie sowohl vor als auch nach dem Bearbeiten der Aufgabe. Die so entstandenen Texte wurden auf das Vorhandensein von zentralen Informationen kodiert. Es entstanden damit analytisch eingeschätzte Produktmaße, welche die Tiefe des konzeptuellen Verstehens erfassten.

Bei den Prozessdaten handelt es sich um Verbaldaten aus den Lautdenk-Protokollen. Diese Daten wurden feinteilig nach einem Schema kodiert, das verschiedene selbstregulatorische Strategien zum Gegenstand hatte. In der Analyse interessierte man sich für zwei Maße, nämlich erstens die negativen Einschätzungen beim aktuellen Leseverstehen, also Aussagen, welche darauf verwiesen, dass die Person ein aktuelles Verstehensproblem hatte. Zweitens wurde kodiert, ob die Strategie vor dem Äußern des Verstehensproblems die gleiche war wie danach. Gab es einen Wechsel in der Vorgehensweise, wurde das als „adaptives metakognitives Verhalten“ kodiert, weil anzunehmen war, dass der Wechsel eine sinnvolle Anpassung darstellte. Wurde stattdessen die gleiche Strategie weiterhin ausgeführt, galt diese Episode als „statisches metakognitives Verhalten“. Die Häufigkeit beider Verhaltensweisen wurde gezählt und in Regressionsanalysen zur Vorhersage des Posttest-Ergebnisses genutzt, wobei das Pretest-Ergebnis berücksichtigt wurde.

Die eigentliche Lernaufgabe, aus der die parallel erhobenen Verbaldaten stammten, bestand aus dem Lesen eines Online-Lexikons mit drei Artikeln (Herz, But und Blutkreislaufsystem). Die Lexikon-Einträge umfassten neben Text auch 40 Illustrationen und ein Video.

D) Was waren die Hauptbefunde?

Adaptives und statisches metakognitives Verhalten kamen selten vor und waren zudem unkorreliert. Das heißt, die Häufigkeit beider Vorgehensweisen bestand rechnerisch unabhängig voneinander. In Regressionsanalysen ergab sich ein spezifischer Effekt: Je häufiger die Personen adaptiv vorgegangen waren, desto besser war ihr Verständnis im Posttest. Ein häufigeres statisches Vorgehen ging zulasten des Verstehens. Eine metakognitive Überwachung mit einer Modifikation des Verhaltens war also für das konzeptuelle Lernen günstig. Dieser Effekt bestand unter statistischer Kontrolle des ursprünglichen Wissens zum Thema.

E) Diskussion

Die Studie zeigt, dass Metakognition beim selbstregulierten Lernen wichtig für ein tieferes konzeptuelles Verständnis ist. Denn die bottom-up-geleiteten Überwachungsprozesse, die eine günstige Vorgehensweise nach sich ziehen (und damit als ein Indiz für Top-down-Prozesse der Planung wertbar sind), münden in besseres Verstehen. Die kontingenzbasierte Operationalisierung der metakognitiven Regulation – hier als Kombination von erkanntem Verstehensproblem und einer Reaktion im Sinne einer möglichen Verhaltensmodifikation – lässt sich zudem als Teil der Wenn-dann-sonst-Regel gemäß Winnes Konzept von Lernstrategien begreifen: Wenn jemand ein Verstehensproblem hat, dann sollte er/sie eine neue Strategie bemühen. Selbst mit dieser groben Operationalisierung von Strategiewechsels ließ sich empirische Differenzen in Leistungen vorhersagen, was für die weitere Erforschung genauerer prozessualer Zusammenhänge ein positives Signal ist.

6.2.2.2 Befundmuster zu selbstregulatorischen Prozessen (und Produkten)

Wenn man sich dem weiten Forschungsfeld Selbstregulation bzw. selbstreguliertes Lernen in der pädagogischen Psychologie zuwendet, um einen Überblick zu gewinnen, dann kann das nicht geschehen, ohne die Wellen und Entwicklungen in der Frage zur Kenntnis zu nehmen, wie Selbstregulation zu messen ist (Dent & Hoyle, 2015). In der Forschungsgeschichte zum selbstregulierten Lernen hat sich tatsächlich eine Kontroverse entsponnen, wie das Konstrukt angemessen zu operationalisieren ist. Verschiedene Operationalisierungen wie Selbstauskünfte als Indikatoren für eher stabile Personenmerkmale und in Aufgaben selbst parallel zum Vorgehen erhobene Prozessmerkmale haben sich als komplementäre Zugänge erwiesen (Panadero et al., 2016), allerdings besteht in jüngerer Vergangenheit die Tendenz, sich nicht alleinig auf verzerrungsanfällige Selbstauskünfte zu beschränken, sondern multimethodisch und triangulativ vorzugehen und daher verschiedene Indikatoren zu verwenden.

Eine für die Zwecke der folgenden Darstellung wichtige Implikation besteht darin, aus Validitätsgründen synchron zur Aufgabenbearbeitung prozessnahe Daten zu gewinnen (Veenman, 2011), und hierbei gibt es zwei Arten, wie mit denselben Daten umgegangen werden kann, wenn die Auftretenshäufigkeit von Strategien erfasst wird. Beide Arten setzen voraus, dass Verhalten oder Verbaldaten kodiert und klassifiziert werden, also beispielsweise als Vertreter distinkter Strategien oder Strategiegruppen höherer Ordnung Kategorien zugewiesen wurden. Die Auswertung derartig gewonnener Daten kann dann in zwei Varianten erfolgen:

- Erstens lässt sich auszählen und vergleichen, wie viele Strategien Individuen verwendet haben. In diesem Fall ist die *Häufigkeit bzw. das Spektrum von Strategien* von Belang, das Maß ist statisch, denn es berücksichtigt nicht die Kontingenz des Auftretens (Meijer et al., 2006). Dafür ist es jedoch dafür geeignet, allgemeinere Zusammenhänge zu betrachten und Personengruppen miteinander zu vergleichen (Dent & Hoyle, 2015). Studien, die diesen Zugang gewählt haben, werden in diesem Teilkapitel konsultiert und ihre Essenz herausgearbeitet.
- Zweitens kann man auch *Wechsel* zwischen einzelnen Strategien fokussieren, um so Muster in der Flexibilität zu eruieren und damit festzustellen, ob es ein musterhaftes Auftreten von Strategien in Sequenzen gibt und wie es sich gestaltet (Meijer et al., 2006). Dieses zweite Vorgehen, welches auf die Rekursivität und deren Zusammenhang mit Leistungen und Leistungsgruppen abzielt, ist neben seiner Behandlung im Kasten „Illustrierendes Beispiel 25“ im Teilkapitel zuvor bereits Gegenstand des ausführlicheren Beispiels im

Teilkapitel [6.2.1.6](#) gewesen (s. Kasten „Illustrierendes Beispiel 24“) und steht an anderer Stelle nochmals im Vordergrund (s. Teilkap. [6.3.1.2](#)).

Prozessperspektiven: das Beispiel Adaptivität

Eingangs im Teilkapitel wurde postuliert, dass die metakognitiv konzentrierte Interaktion von Strategien für den Erfolg beim selbstregulierten Vorgehen zentral ist. Dieser Gedanke speist sich daraus, dass es eine doppelte Sensitivität der Lernenden braucht, die sich in einem adaptiven Vorgehen manifestiert, also nicht in einem rigiden akontextuell anmutenden Absolvieren einer Sequenz von Strategien. Zum einen ist hierfür das im Teilkapitel [6.1.2.3](#) bereits angesprochene konditionale Strategiewissen nötig, also das Wissen darüber, unter welchen Bedingungen welche Strategien geboten erscheinen, was wiederum die Aufgabe und diverse kontextuelle Merkmale in ihrer Eminenz für das selbstregulatorische Geschehen betont. Zum anderen ist die Kontingenz des komplexen Geschehens bei der Aufgabenbearbeitung im Zusammenspiel von Objekt- und Metaebene ein wichtiger Faktor, da sich dynamische Interaktionen von Planen, Ausführen und Überwachen ergeben, welche die lesende Person steuert. Dieser zweitgenannte Aspekt wird an anderer Stelle anhand des rekursiven Vorgehens vertiefend betrachtet (s. Teilkap. [6.3.1.2](#)). Hier jedoch steht der erstgenannte Aspekt im Zentrum: *Personen passen ihre Strategien an die Spezifika von Aufgaben an*. Diese Adaptivität kann sich in Phasen bei einzelnen Aufgaben manifestieren, über verschiedene Aufgaben hinweg (Intrapersonenvergleich) bzw. im Vergleich verschiedener Personengruppen auftreten, die unterschiedliche Aufgaben lösen und teils auch im Vergleich von Medien vorkommen. Diese Aufzählung verdeutlicht schon, dass das Konzept Adaptivität hier sehr weit gefasst werden muss, um die sehr unterschiedlichen Studien unter sich zu vereinen. Daher ist die nachfolgende Darstellung kein systematischer Überblick, sondern eher ein Streiflicht in einem noch lückenhaften, emergierenden Forschungsfeld:

- *Merkmale von Aufgaben:* Bei *schwierigeren Aufgaben* schätzten Personen die Notwendigkeit komplexerer Strategien höher ein (Bromme et al., 2014; Pieschl et al., 2014; Stahl et al., 2006). Entsprechend nutzten Personen beispielsweise mehr metakognitive und tiefenbezogene kognitive Strategien als bei leichteren Aufgaben (Chevrier et al., 2020; Malmberg et al., 2013, 2014; Moos, 2013; Taub et al., 2014; Veenman & Beishuizen, 2004) und verbrachten mehr Zeit mit dem Lösen solcher Aufgaben (Pieschl et al., 2012). Sie verwendeten in für sie *motivierenden Aufgaben* mehr Strategien als in weniger motivational stimulierenden Aufgaben (Moos & Miller, 2015). Teils bestanden ebenfalls unterschiedliche Nutzungsfrequenzen von metakognitiven Strategien *bei verschiedenen Lern- bzw. Leistungszielen*, welche in der Aufgabe vorgegeben waren (Moos & Azevedo, 2006). Zudem gibt es Muster insofern, dass bei zu

niedrigen Ergebnissen Personen auf der Basis ihrer eigenen Überwachung beim hochschulischen Lernen zu einem adaptiveren Vorgehen neigen, dieses dann über Handlungen realisieren und dadurch bessere Leistungen erbringen (Raković, Bernacki et al., 2022).

- *Mediumsbezogene Eigenheiten* ließen sich zudem feststellen, wenn Personen Inhalte in Texten statt in Videos dargeboten bekommen; im Fall der textuellen Darbietung wendeten sie im Vergleich zur Verarbeitung audiovisueller Inhalte mehr kognitive Strategien an, die dem Auffinden wichtiger Informationen und der vorwissensbasierten Inferenzenbildung dienten (Lee & List, 2019).
- Ferner sind Unterschiede in verschiedenen *Domänen* zu beobachten. In einer Studie war beispielsweise feststellbar, dass Personen, die naturwissenschaftliche Aufgaben bearbeiteten, mehr metakognitive Strategien nutzten als Gleichaltrige, die einen Leseauftrag aus der Domäne Geschichte absolvierten (Greene et al., 2015, Studie 1). In derselben Studie erwiesen sich zudem in den beiden Domänen die Zusammenhänge zwischen dem Einsatz von Strategien partiell unterschiedlich. Das bedeutete hier, dass sich Makrogruppen von positiv oder negativ kommentierten Passagen in den Vorgehensweisen zum einen aus mitunter deutlich unterschiedlichen einzelnen Strategien zusammensetzten. Zum anderen erwies sich nur in der Domäne Naturwissenschaft die Gesamtheit positiv kommentierter Vorgehensweisen als prädiktiv für den Lernzuwachs, nicht aber in der Domäne Geschichte.
- Neben Untersuchungen zu solchen allgemeinen Effekten gibt es ebenso Experimente, die sich mehr der *Phasenlogik des selbstregulierten Lernens und Lesens* verschrieben haben. Follmer und Sperling (2019) konnten beispielsweise zeigen, dass Personen in der Planungsphase weniger und in ihrer Informationstransformation weniger restrukturierende und den Inhalt mit Vorwissen vernetzende Strategien nutzten als in der Phase der Durchführung und teils in der Selbstreflexionsphase. Ähnliche Befunde stammen aus anderen Studien, die differenzielle statistische Zusammenhänge zwischen der Zahl von Strategien in einzelnen Phasen vorfanden (DiBenedetto & Zimmerman, 2013) oder auch markante Verschiebungen bei längerfristigen Selbstregulationszyklen auf der Basis des Lesens feststellten (Bråten & Strømsø, 2003; Cho, 2014). In Studien, die mit dem lauten Denken die Prozesse kontinuierlich erfassten und die Gesamtzeit in gleich große Abschnitte zerteilten, um dann die Auftretenshäufigkeiten statistisch auf Differenzen zu überprüfen, sind solche Befunde allerdings nicht durchgängig ermittelt worden (Johnson et al., 2011; Moos & Azevedo, 2008a; Paans et al., 2019).

Mögen die in den Studien berichteten, oben gelisteten Befunde zunächst disparat erscheinen, so eint sie auf den allgemeinsten Nenner gebracht der Umstand, dass *Selbstregulation nicht alleinig eine stabile Fähigkeit einer Person ist, sondern*

sich situativ und in ihrer Passung zum Kontext und im Zusammenspiel von Aufgaben- und Individualmerkmalen manifestiert. Exakt das ist mit *Adaptivität* gemeint (Pieschl et al., 2013).

Perspektiven auf Prozesse im Verbund mit Produkten

Die Forschung zu den Bei- und Erträgen des selbstregulatorischen Lernens versucht häufig über prozessnahe Indikatoren das individuelle Vorgehen zu erfassen und dieses dann überindividuell auszuwerten. Dabei spielt es offenkundig nicht nur eine Rolle, welche Strategien wie ermittelt und kodiert werden, sondern auch wie sie aggregiert mit welchen Verfahren statistisch ausgewertet werden. Aus den Studien lassen sich einige komplexe Muster filtern.

- *Teils sind für die rechnerischen Vorhersagen von Leistungen metakognitive Indikatoren bei den Strategien bedeutsamer als die kognitiven Strategien.* Das Konstrukt Selbstregulation beinhaltet es, dass Personen verschiedenartige Strategien auf der Meta- und Objektebene koordinieren, um das Leseziel zu erreichen. Selbst wenn damit impliziert ist, dass günstigenfalls ein Verbund verschiedener und als Wenn-dann-sonst-Sequenz verketteter Strategien zum Einsatz gelangt (Winne & Marzouk, 2019), stellt sich aus empirischer Sicht die Frage, welche dieser Strategien besonders bedeutsam ist. Um es vorwegzunehmen: Die Forschungslage gestaltet sich als ausgesprochen disparat, weil wegen der Verschiedenheit in den Operationalisierungen kaum echte Vergleiche möglich sind oder gar Replikationen ausreichend häufig vorgelegt worden sind. Daher lässt sich annäherungsweise die Thematik etwas auffächern, indem einige Tendenzen aus der Forschung extrahiert werden. Einen allgemeinen Hinweis auf die besonders hohe Bedeutsamkeit metakognitiver Strategien lieferte die *Metaanalyse* von Dent und Koenka (2016), in der sich statistisch signifikante Differenzen in den Zusammenhängen von kognitiven und metakognitiven Strategien und Leistungen zeigten. Die metakognitiven Strategien als gesamte Gruppe hängen stärker mit verschiedenen Leistungsmaßen zusammen, als es bei den kognitiven Strategien der Fall ist. Diese Analyse, die gezielt die eine Gruppe von Strategien mit einer anderen Gruppe verglich, homogenisierte (notwendigerweise) die vielen Einzelstrategien und ließ außer Acht, dass Strategien zusammenwirken. Um dies besser abzubilden, lohnt sich ein Blick in die Primärstudien, in denen synchron beide Gruppen von Strategien erfasst und ihre Zusammenhänge mit Leistungen statistisch analysiert wurden, z. B. in *Studien mit Regressions- und pfadanalytischen Zugängen*. Metakognitive und kognitive Strategien leisten in derartigen Experimentalstudien in varianzanalytischen Auswertungen separate Beiträge, d. h. sie haben jeweils eigene positive erklärende Anteile, wenn es darum geht, Unterschiede in Leistungen zu prädictieren (Bernacki et al., 2012; Greene &

Azevedo, 2009; Greene, Bolick & Robertson, 2010; Greene et al., 2011, 2014). Studien, die z. B. über Strukturgleichungsanalysen Zusammenhänge prüften, konnten zeigen, dass metakognitive Strategien indirekt wirken, indem sie mit einem erhöhten Einsatz von tiefenorientierten Lernstrategien einhergehen, welche ihrerseits verbesserte Leistungen in Produktmaßen nach sich zogen (Deekens et al., 2018, Studien 1 und 2; Greene et al., 2020).

- *Trotz theoretischer Bedeutsamkeit sind die direkten Zusammenhänge von einzelnen Strategien und Gesamtleistung teils nicht immer leicht erkennbar, sondern werden dann deutlicher, wenn Daten aggregiert und phasensensibel ausgewertet werden.* Dieses Muster lässt sich in den Studien besonders deutlich ermitteln, in denen aus Lautdenk-Protokollen die einzelnen Strategien einzeln ausgezählt und dann als Summenwerte und mit Leistungsdaten in Bezug gesetzt werden (Bernacki et al., 2012; Greene et al., 2011).

Dieses Muster mutet auf den ersten Blick etwas paradox an, lässt sich aber teils darüber erklären, dass nicht die Anzahl ausgeführter Strategien allein für den Lernerfolg in einem linearen Verhältnis steht. Strategien nutzen Personen nämlich hochgradig individuell. Bei der Auswertung besteht zudem das Problem, dass viele Datensätze bestimmte statistische Voraussetzungen nicht ausreichend erfüllen (Greene et al., 2011) oder dass verschiedene Vorgehensweisen bei der statistischen Auswertung zu divergierenden Resultaten führen (Greene, Copeland et al., 2018; Greene et al., 2014). Um diesem Umstand dennoch gerecht zu werden, gibt es zwei Verfahren in der Forschung mit Lautdenk-Protokolldaten: die volle und die datengetriebene Aggregation (Greene, Deekens et al., 2018). Beide Varianten eint, dass sie sich von der Betrachtung singulärer Strategien eher lösen, um stattdessen Indizes zu bilden, welche eine höhere prognostische Validität aufweisen.

Erstens lassen sich die vielen einzelnen Strategien theoretischen übergeordneten Kategorien zuschlagen, was man als „*volle Aggregation*“ bezeichnet: Alle Vertreter von Strategien auf der konkret operationalisierten Ebene einzelner Strategien werden als Summenwert für eine theoriebasierte Makrokategorie gezählt, was impliziert, dass alle einzelnen Strategien als günstig für das Lernen zum einen und als gleichwertig zum anderen angenommen werden (Greene, Deekens et al., 2018). Mittels dieses theoriegeleiteten Vorgehens hat die Arbeitsgruppe rund um Azevedo und Greene die bis zu neun Makrostrategien (darunter fünf als durchgängige Makrokategorien: 1) Planen, 2) Überwachen, 3) Strategienutzung, 4) Aufgabenschwierigkeit und -erfordernisse; 5) Interesse) auf der Basis mehrerer Dutzend einzelner Strategien gebündelt (Greene & Azevedo, 2009; Greene, Copeland et al., 2018; Greene et al., 2014; Johnson et al., 2011). Andere Autorinnen und Autoren haben ein ähnliches Vorgehen genutzt (Schellings et al., 2013) bzw. andere Taxonomien für die Bündelung von einzelnen Strategien wie oberflächen- und tiefenorientierte Strategien (Chevrier et al., 2020; Deekens et al., 2018; Dinsmore & Alexander, 2016; Greene et al., 2020;

Parkinson & Dinsmore, 2018), die Untergliederung in verschiedene metakognitive Strategien (Meijer et al., 2006) oder allgemein gehaltene Funktionen von Lesestrategien (Anmarkrud et al., 2014) verwendet. Teils ließen sich mittels diesen Makrokategorien als Prozessmerkmale Unterschiede in den Produkten vorhersagen (Greene & Azevedo, 2009; Greene et al., 2020; Moos, 2010).

Ein anderes, zweites Vorgehen derselben Arbeitsgruppe wird als „*datengetriebene Aggregation*“ deklariert und basiert auf dem Gedanken, dass nicht theoretische, sondern empirische Zusammenhänge von einzelnen Strategien den Ausschlag geben (Greene, Deekens et al., 2018; Greene et al., 2014; Raković, Bernacki et al., 2022). Konkret werden dann mit Leistungsdaten korrelierte Strategien abhängig von einer Mindestkorrelation ebenfalls aufsummiert, und zwar sowohl für positive als auch für negative Korrelationen. Dieses empirisch induzierte Vorgehen dient mithin dazu, negative und positive Prädiktoren zu einem Gesamtkonstrukt zu addieren. Solcherart konstruierte Konstrukte haben ebenfalls ihren Wert als unabhängige Variablen unter Beweis gestellt (Greene et al., 2014).

Neben der theoriebasierten bzw. datengetriebenen Bündelung von kodierten Lautdenk-Protokollen beginnt sich eine weitere Form zu etablieren, das selbstregulierte Vorgehen zum einen handlungsnah zu erfassen und zum anderen multimethodisch und zu kontrollierten Zeitpunkten Daten zu gewinnen. Dieses als „*Mikroanalyse*“ bezeichnete Vorgehen erlaubt es, zu definierten Zeitpunkten interindividuell vergleichbare Indikatoren zu erheben und auszuwerten. Selbstreguliertes Lernen wird über strukturierte, an die Prozessphasen des selbstregulierten Lernens angepasste teilstandardisierte Interviews systematisch gemessen (Cleary & Callan, 2018; Cleary et al., 2012, 2023). Das im Kasten „Illustrierendes Beispiel 26“ enthaltene Exempel einer Studie steht für eine mikroanalytische Vorgehensweise in der Erforschung des selbstregulierten Lernens bzw. des Lesens. Das Beispiel stammt aus dem Lernen mit Fachtexten in der Naturwissenschaft und gibt Zeugnis darüber, dass Selbstregulation einerseits leistungsgruppenspezifisch ist und andererseits Leistungen vorherzusagen imstande ist. Mit anderen Worten: In diesem Exempel manifestieren sich Zusammenhänge von Prozessen und Produkten allgemein und in Bezug auf verschiedene Leistungsniveaus im Besonderen.

Illustrierendes Beispiel 26

Lern- und Lesestrategien und metakognitive Kontrolle und ihre positiven Effekte auf das naturwissenschaftliche Lernen mit Texten

A) Was war das Erkenntnisinteresse?

Die in zwei Publikationen beschriebene Studie ging der Frage nach, wie sich verschiedene prozessnah erfasste Merkmale selbstregulativen Lernens von Inhalten aus Texten in verschiedenen Leistungsgruppen unterscheiden und wie sie untereinander zusammenhängen sowie Lernleistungen vorhersagen (DiBenedetto & Zimmerman, 2010, 2013).

B) Wer wurde untersucht?

Das Sample bestand aus 51 Elftklässlern aus den USA, davon waren zwei Drittel Mädchen, etwas mehr als die Hälfte waren europäischer Herkunft, die restlichen Jugendlichen waren afroamerikanisch, hispanisch bzw. asiatisch-pazifisch. Der sozioökonomische Hintergrund war mittel, die Gegend der Schulen vorstädtisch, die Schulen waren in konfessioneller Trägerschaft. Das Sample war gezielt zusammengestellt worden, wobei die Leistungen im Fach Biologie aus der zehnten Klasse herbeigezogen worden waren: Je ein Drittel repräsentierte schwache, mittlere und leistungsstarke Personen.

C) Welche Instrumente kamen zum Einsatz?

Die Studie setzte verschiedene Instrumente ein, wobei seitens der Produktmaße entscheidend war, wie gut die Personen eine drei Seiten lange Passage aus einem Lehrbuchtext über Tornados verstanden hatten. Zwei Tests kamen dafür zum Einsatz: ein Multiple-Choice-Test mit 10 Fragen zu Tornados (Faktenlernen) und ein Test, bei dem drei Phasen der Entstehung von Tornados als Zeichnung anzufertigen waren (konzeptuelles Lernen). Beide Tests waren hoch miteinander korreliert.

Die Tests zu den *Prozessen* bzw. zu den theoriebasierten Phasen-Merkmalen orientierten sich strikt am zyklischen Modell der Selbstregulation von Zimmermann (s. Teilkap. [6.1.1.2](#)) und kamen in den verschiedenen Phasen zum Einsatz. Ganz im Sinne des mikroanalytischen Analyseansatzes wurden Fragen gezielt über die Aufgabenbearbeitung verteilt. Fünf Maße wurden in der Planungsphase hinsichtlich der Motivation mittels geschlossener Fragen fokussiert (Selbstwirksamkeit einerseits für das Lernen, andererseits für das Testergebnis, Interesse, Ergebniserwartung, Ziel für das angestrebte Testergebnis), drei weitere (Kausalattributionen, Zufriedenheit, adaptive/defensive Reaktionen) wurden ganz am Ende in der Selbstreflexionsphase mit offenen und geschlossenen Fragen erhoben, wobei den Jugendlichen ihr Ergebnis im Test mit der Zeichnung zuvor mitgeteilt worden war. Über alle drei Phasen hinweg wurden zudem (*meta-*)*kognitive Indikatoren* erfasst. In der *Planungsphase* wurden die Personen gebeten, mögliche Pläne zum Vorgehen für das Bearbeiten des Textes zu benennen, den sie vorliegen hatten, aber noch nicht lesen durften. Aus der Anzahl verschiedener Nennungen wurde ein Summenwert von Strategien gebildet. In der *Durchführungsphase* wurden zwei Arten von Strategien mit separaten offenen Fragen gesammelt, nämlich zum einen welche Strategien die Personen für das Lesen des Textes

genutzt hatten und zum anderen jene zur Vorbereitung auf den bevorstehenden Test. Für beide Arten Fragen wurden die Strategien als Summenwert gezählt. Danach kamen die Verstehenstests zum Einsatz, nach deren Bearbeitung die Jugendlichen in jeweils zwei Maßen angeben sollten, wie zuversichtlich sie waren, bei den Tests erfolgreich gewesen zu sein, und welches Ergebnis sie erwarteten. In der *Selbstreflexionsphase* wurde schließlich noch die Einschätzung erbeten, wie die Jugendlichen ihren Lernzuwachs zu den Tornado-phasen einschätzten. Summa summarum wurden also das Strategierepertoire in Planung und Durchführung gemessen und die Einschätzung der Urteilsgenauigkeit im Lernen sowie die Selbsteinschätzung des eigenen Lernens umfassend und prozessnah ermittelt. Die Daten wurden in Zusammenhang mit der Zugehörigkeit zu den drei Leistungsgruppen gebracht und auf statistische Gruppenunterschiede überprüft. Außerdem wurden diese Daten auch untereinander ins Verhältnis gesetzt, nämlich in Korrelationsanalysen.

D) Was waren die Hauptbefunde?

Die Ergebnisse der Studie sind reichhaltig und lassen sich unter zwei Gesichtspunkten bündeln, nämlich einerseits dem Vergleich der Leistungsgruppen und andererseits dem Zusammenhang von Prozess- und Produktmaßen. Die *Leistungsgruppenvergleiche* sind insgesamt konsistent:

- In den Produktmaßen spiegelte sich die Gruppenzugehörigkeit wider: Je leistungsstärker die Personen laut Gruppe waren, desto höher waren ihre Verstehensleistungen auch in beiden Tests.
- Die motivationalen Maße in der Planungsphase waren in allen Leistungsgruppen ähnlich, signifikante Differenzen traten erst in der Selbstreflexionsphase auf. Dort waren die Muster bei den leistungsstärksten Jugendlichen am eindeutigsten zugunsten einer selbstregulativ günstigen Einschätzung.
- In sämtlichen kognitiven und metakognitiven Maßen waren die leistungsstärksten Schülerinnen und Schüler den leistungsschwächsten überlegen. Differenzen zwischen allen drei Leistungsgruppen kamen vor allem in den metakognitiven Urteilen der Personen während und nach der Durchführung der Tests zum Tragen, denn hierin unterschieden sich die Gruppen in vier von fünf Variablen durchgängig.

Betrachtet man die *Zusammenhänge zwischen den Variablen*, ergeben sich diverse Muster:

- Es gab Zusammenhänge zwischen der Dauer der Aufgabenbearbeitung und Maßen der Selbstregulation: Je länger die Personen lasen und sich auf den Test vorbereiteten, desto mehr Lese- und Lernstrategien nannten sie und desto höher schätzten sie ihren Lernzuwachs ein.
- Motivationale Maße der Vorbereitungsphase korrelierten überwiegend nicht mit den kognitiven und metakognitiven Prozess- und Produktmaßen.
- Die Menge an möglichen Strategien in der Vorbereitungsphase, die Anzahl während der Aufgabenbearbeitung zu Protokoll gegebenen Strategien sowie die Einschätzungen zur Aufgabenbearbeitung korrelierten mit wenigen Ausnahmen recht konsistent und substantiell.

- In Regressionsanalysen wurden anhand der kognitiven und metakognitiven Indikatoren aus den drei Phasen geprüft, ob diese die Leistungen in den Tests vorhersagen. Das war bei allen Maßen der Fall, vor allem die Maße aus der Durchführungsphase waren besonders stark in der Prädiktion späterer Verstehensleistungen. Beim Multiple-Choice-Faktenlernen war eine Varianzaufklärung von 50 Prozent festzustellen, im konzeptuellen Lernen waren es sogar 81 Prozent.

E) Diskussion

Die mikroanalytische, einen gesamten Selbstregulationszyklus (bzw. richtiger: zwei Zyklen, denn die Studie fokussierte das erschließende Lesen und das Lernen für einen Test) fokussierende Untersuchung erbrachte reichhaltige Befunde aus Prozess- und Produktsicht. Die konsistenten Gruppenunterschiede in den metakognitiven Maßen und den Maßen in der Selbstreflexionsphase weisen darauf hin, dass leistungsstärkere Schülerinnen und Schüler dazu in der Lage sind, ihr breiteres Strategierepertoire besser zu nutzen und ihren Strategieeinsatz und dessen Folgen adäquater zu beurteilen. Hinzu kommt, dass die faktisch in der konkreten Aufgabenbearbeitung genannten Strategien und die eigenen metakognitiven Leistungseinschätzungen substanziell zusammenhängen, was auf die hohe Bedeutung von metakognitiven Kontrollprozessen als Bindeglied von Objekt- und Metaebene verweist. In der Leistungsprädiktion zeigten sich hohe Varianzaufklärungen mithilfe metakognitiver Wissensbestände sowie Einschätzungen und der Anwendung von Strategien, welche in der Studie dazu geeignet waren, interindividuelle Unterschiede im faktenbasierten und ganz besonders im konzeptuellen Lernen vorherzusagen. Dies unterstreicht ebenfalls die Wichtigkeit selbstregulativer Komponenten in ihrem Verbund, in diesem Fall beim textbasierten Lernen naturwissenschaftlicher Inhalte.

6.2.3 Konstruktiv-responsives Lesen als eine empirisch fundierte Idealbeschreibung selbstregulierten (digitalen) Lesens

Digitales Lesen erfordert ein selbstreguliertes Vorgehen, in dem Personen metakognitiv gesteuert ihre Prozesse auf ein Ziel ausrichten und entsprechend planen, überwachen und modifizieren. Da Strategien sich in ihrem Nutzen in konkreten Situationen und im Wechselspiel mit verschiedenen in- und externalen Faktoren in ihrer Eignung beweisen müssen, stellt sich die Frage danach, welche allgemeinen strategischen Vorgehensweisen als günstig gelten bzw. gelten sollen. Aus dieser Frage ergibt sich eine Anschlussfähigkeit an die empirische Bildungsforschung, wenn es darum geht, jenseits von einzelnen Studien übergeordnete Antworten zu finden. Genauer gesagt helfen Synopsen von Studien, um einem Panorama gleich ein vollständigeres Bild zu zeichnen, was überindividuell selbstreguliertes Lesen ausmacht, welches sich idealiter in tatsächlich gesteigerten Ergebnissen des Lesens niederschlägt.

Um diese Gesamtschau zu erstellen, ist es nötig, sorgfältig in Erwägung zu ziehen, auf welcher Datenbasis valide Aussagen überhaupt möglich sind. Hier erweisen sich solche Studien als besonders überlegen, in denen das Verhalten beobachtet bzw. über Lautdenk-Protokolle erfasst wird. Nicht nur hängen solche Maße nachweislich enger mit Leistungen zusammen (Dent & Koenka, 2016; Ohtani & Hisasaka, 2018). Es haben sich zudem in diversen Studien Muster gezeigt, nach denen beispielsweise Selbstauskünfte als wenig valide erweisen, wenn sie mit beobachtbarem Verhalten abgeglichen wurden (Bråten & Samuelstuen, 2007; Cromley & Azevedo, 2006; Dörrenbächer-Ulrich et al., 2021; List & Alexander, 2018b; Rogiers et al., 2020b; Winne & Jamieson-Noel, 2002); dies ließ sich ebenfalls studienübergreifend anhand geringer Korrelationen metaanalytisch belegen (Craig et al., 2020). Dabei scheinen weniger die Strategien auf der Objektenebene problematisch zu sein als metakognitive Strategien der Metaebene, in denen Selbstauskünfte und Verhalten teils deutlich auseinanderklaffen (Merchie & van Keer, 2014a), was sich über Alterseffekte zuungunsten jüngerer Personen erklären lässt (Craig et al., 2020). Außerdem wirken Selbstauskünfte zu einzelnen Strategien eher unzuverlässig als Globalurteile, denen freilich die Genauigkeit fehlt (Rovers et al., 2019). Vor diesem Hintergrund, dass handlungsnah, aber methodisch aufwändige Verfahren validere Aussagen über das Vorgehen zulassen, lässt sich legitimieren, diesen Verfahren den Vorzug für einen Überblick zu selbstregulatorischem Lesen zu geben.

Eine einflussreiche Systematisierung aus Studien mit Lautdenk-Protokollen in Lesestudien haben Pressley und Afflerbach (1995) vorgelegt. Aus der Gesamtschau von diversen Studien extrahierten sie eine idealtypische, das Lesen und Vorgehen von Expertinnen und Experten widerspiegelnde Art (Ertmer & Newby, 1996), die sie als *konstruktiv-responsives Lesen* bezeichnet haben. Die namensgebenden Bestandteile unterstreichen eine aktive Bedeutungskonstruktion in einem Prozess des absichtsvollen Lesens, welcher auf Eigenheiten des Textes responsiv reagiert. Ein solches Lesen zeichnet sich durch ein Bündel von mehreren Merkmalen aus. Konstruktiv-responsives Lesen

- ist geprägt von einem übergeordneten Bemühen, Texte tief zu verstehen und dabei Hauptinhalte zu suchen, sie zu konstruieren und über sie zu reflektieren;
- zeichnet sich durch das eine hohe Bedeutung des Vorwissens aus, welches ebenso Gegenstand von Hypothesen zur Bedeutung des Gelesenen und Indikator der Verarbeitung von Textinhalten ist, wie es sich veränderlich erweist und Gegenstand von Reflexionen auf der Basis des Gelesenen ist;
- ist schließlich ein Lesen, das emotional-affektive und motivationale Bestandteile aufweist und damit mehr als reine Kognitionen inkludiert (Pressley & Afflerbach, 1995).

Diese Extraktion von hier bereits stark aggregiert dargestellten Merkmalen war Ausgangspunkt für eine Reihe von Folgeentwicklungen. Es gibt im Zusammenhang mit dem konstruktiv-responsiven Lesen verschiedene, auf Forschungssynopsen mit Lautdenk-Protokollen fußende Vorarbeiten, die sich auf das Lesen von einzelnen (Sach-)Texten (Afflerbach & Cho, 2009; List, 2020; Pressley & Afflerbach, 1995), multiplen Texten (Afflerbach & Cho, 2009; Cho & Afflerbach, 2017; Cho, Afflerbach & Han, 2018; List, 2020) und Hypertexten (Afflerbach & Cho, 2009, 2010; Cho & Afflerbach, 2017; Cho, Afflerbach & Han, 2018) konzentriert haben. Teils hat sich die Perspektive zu ändern begonnen, indem nach den analogen Texten die digitalen Texte ebenso explizit in den Blick geraten sind (Afflerbach & Cho, 2009) wie nicht-sprachliche bzw. alleinig sprachliche Bestandteile wie Grafiken, Bilder und Videos (List, 2020). Die Ausweitung der Perspektiven ist zuvorderst und erstens dem Umstand geschuldet, dass die Forschungssynopsen sich aus einem größeren und breiteren Pool an Studien speisen. Dabei ist es zweitens zu einer Binnendifferenzierung gekommen, weil beispielsweise Eigenheiten wie strategisches Navigieren im Fall der Hypertextlektüren oder das textvergleichende Lesen zur Klärung der Widerspruchsarmut (Corroboration) bei dem Verstehen multipler Dokumente den spezifischen Anforderungscharakter dieser Lesesituationen widerspiegeln. Drittens ergeben sich in dieser Evolution der Forschungsstände und -gegenstände erhebliche Schnittmengen, sodass generisch wirkende Strategien mit breiter Einsatzmöglichkeit und solche mit spezifischen Erfordernissen, welche also einen enger gefassten Anwendungsbereich haben, deutlicher konturiert wirken. Dies ist generell ein Unterscheidungskriterium von Lernstrategietaxonomien, nämlich in der Differenzierung von Domänen- und Aufgabenspezifika (Dumas, 2020). Diese Ausdifferenzierung findet also ihre Entsprechung und ihren Niederschlag im Bereich selbstregulatorischer Lesestrategien.

Die Taxonomievorschlge, die sich aus der sehr reichhaltigen Forschungsaktivitt ergeben, variieren, was sich aus ihren Bezugspunkten ableitet und aus dem Abstrahieren der Funktionen einzelner Strategien erklrbar ist, welche verschiedene Zuteilungsmglichkeiten mit sich bringen. Zwei solcher aktuellen Taxonomien enthlt Tabelle 35. Wie aus der Spalte ganz links hervorgeht, eint die beiden Taxonomien, dass sie sich auf drei Gruppen von Strategien bzw. deren Funktionen erstrecken. Die A) *Strategien des konstruktiv-integrierenden Verarbeitens* dienen dazu, Informationen zu finden und in ihrem Zusammenhang zu erkennen, worunter auch das Navigieren fllt. Strategien des B) *kritisch-analytischen Verarbeitens* kommt die Aufgabe zu, Informationen aus verschiedenen Perspektiven heraus zu beurteilen. Diese beiden Gruppen von Strategien lassen sich auf der Objektebene der inhaltlichen Auseinandersetzung mit dem Lesen verorten (s. Teilkap. [6.1.2.2](#)). Die letzte Gruppe betrifft hingegen eher die Metaebene und

die Informationsflüsse zwischen Objekt- und Metaebene. Demgemäß sind die C) *Strategien des metakognitiv-reflexiven Verarbeitens* zuvorderst dazu da, sowohl die Prozesse als auch die Produkte des verstehenden Lesens inklusiver etwaiger Störquellen zu fokussieren und zu optimieren.

Die beiden Taxonomien sind hierbei unterschiedlich konzipiert worden. Jene von Cho, Afflerbach und Han (2018) – außen rechts in der Tabelle – geht deziert von den Funktionen aus, die sich aus dem Vorgehen in den Primärstudien rekonstruieren lassen. Die Taxonomie von List (2020), welche auf einer vorherigen basiert (List & Alexander, 2019) und bereits einen Versuch einer Modellintegration auf der Basis der Taxonomie von Cho, Afflerbach und Han (2018) darstellt, ist in der mittleren Spalte der Tabelle 35 enthalten. Wie aus dem Tabellenkopf bei den Bezugspunkten der Taxonomie hervorgeht, sind drei Referenzpunkte leitend: einzelne Texte, mehrere Texte oder die eigene Person. Diese werden dann für die drei Funktionsgruppen spezifiziert, sodass Lists (2020) Taxonomie einen größeren Geltungsbereich (inkl. Lesen von nicht-digitalen und singulären Texten) beansprucht und dabei mehrere Subkategorien erlaubt, nämlich in Form einer 3-x-3-Logik. Es handelt sich mithin um eine umfassendere Taxonomie. Gleichwohl gibt es deutliche Schnittmengen zwischen beiden in der Tabelle skizzierten Taxonomien, auf welche die Querverweise in der Spalte rechts hinweisen.

Tabelle 35: Gegenüberstellung der Binnendifferenzierung von drei Gruppen von Lesestrategien

	Taxonomie von List (2020)	Taxonomie von Cho, Afflerbach und Han (2018)
Bezugspunkte in der Taxonomie	1 = intratextuelle Referenz; 2 = intertextuelle Referenz; 3 = individuumsbezogene Referenz	ausschließlich multiple, durchaus multimodale Texte
A) Konstruktiv-integrierendes Verarbeiten (Strategien, die zur Bedeutungskonstruktion innerhalb von Texten und über Texte hinweg nötig sind)	1 = Zusammenfassen, Paraphrasieren 2 = Inhalte aus mehreren Texten miteinander elaborierend verknüpfen 3 = Informationen mit eigenem Vorwissen anreichern, Assoziationen zum Gelesenen	• Informationen finden und lokalisieren (Recherchieren und Navigieren) • Identifizieren und Lernen von wichtigen Ideen aus mehreren Quellen (s. auch Nr. 2 bei List, 2020, in der Spalte links) • neue Perspektiven und Wissen entwickeln (durch Analyse konfligierender und zunächst nicht kohärent wirkender Aussagen)
B) Kritisch-analytisches Verarbeiten (Strategien, welche die Quelle, Qualität, Glaubwürdigkeit und Wahrhaftigkeit fokussieren)	1 = Sourcing und Beurteilen von Texten 2 = Corroboration von Inhalten, Quellen von Aussagen untereinander vergleichen 3 = Kontextualisieren und Interpretieren	• Sourcing (intra- und intertextuell; s. auch Nr. 1 bei List, 2020, in der Spalte links) • Textanalysen (kritische Evaluation der Inhalte und ihrer Darstellung auf Angemessenheit in Bezug auf Gegenstand und kommunikative Absicht) • Quellenbeurteilung (Einschätzung der Dokumente aufgrund von verschiedenen Analysen, insbesondere intertextuellem Verstehen; s. auch Nr. 2 bei List, 2020, in der Spalte links)
C) Metakognitiv-reflexives Verarbeiten (Strategien, welche der Überwachung und Regulation des Verstehens dienen)	1 = Verstehensüberwachung von einzelnen Texten, Bestimmen des Verstehenserfolgs 2 = Verstehensprobleme bei einem Text dadurch lösen, Informationen in anderen Texten zu suchen 3 = Bewusstsein für eigene verzerrte Perspektive erkennen und überwachen	• Überwachen und Konstruieren von Verarbeitungswegen (einerseits durch Planen und damit Kontrollprozesse, andererseits durch Überwachungsprozesse in Selbstregulationszyklen) • Verstehensüberwachung und Veränderung des Vorgehens anhand eigener Strategieeinsätze und der Wahrnehmung der eigenen Produktivität (s. auch Nrn. 1 und 3 bei List, 2020, in der Spalte links)

(modifizierte Dar- und Gegenüberstellung auf der Basis von Cho, Afflerbach & Han, 2018, S. 137 f., und List, 2020, S. 122)

Die Evolution dessen, was unter dem Konzept des konstruktiv-responsiven Lesens oder abgeleiteten Konstrukten gefasst wird, weist auf Konstanten ebenso hin wie auf neuere Phänomene, die lesebezogenen Anforderungen entsprechen. Diese neuen Anforderungen bestehen etwa beim Lesen multipler Dokumente und Repräsentationen darin, mit Ambiguität und Kohärenzproblemen umzugehen

und intertextuelle Kohärenz aufzubauen, wobei sich dies durchaus durch kritisch-evaluative Leseprozesse erst leisten lässt (s. dazu den Kasten „Illustrierendes Beispiel 27“). Beim Lesen und Finden von Hypertexten kommt hinzu, dass Navigationspfade zu den eigentlichen Dokumenten sich als zusätzliche Erfordernisse darstellen.

Illustrierendes Beispiel 27

Wenn Laien lesebasierte Ratschläge erteilen sollen – der Nutzen von Lesestrategien im Verbund für eigene angemessene argumentative Texte

A) Was war das Erkenntnisinteresse?

In einer Studie mit norwegischen Studierenden interessierte sich das Forschungsteam für den Zusammenhang verschiedener strategischer Prozesse des Lesens multipler Dokumente und Produktmaßen in Form von argumentativen Essays (Anmarkrud et al., 2014).

B) Wer wurde untersucht?

Insgesamt 51 Studierende, davon 43 Frauen, mit einem durchschnittlichen Alter von 22 Jahren nahmen an dem Experiment teil. Sie waren überwiegend Personen mit Norwegisch als Muttersprache, wiesen einen mittleren sozioökonomischen Status auf und studierten im Bachelor bildungswissenschaftliche Studiengänge. Da im Experiment ein naturwissenschaftliches Thema behandelt wurde, sind die Personen als Laien einzuschätzen, was sich in einem zusätzlich eingesetzten Wissenstest verifizieren ließ.

C) Welche Instrumente kamen zum Einsatz?

Das Experiment sah es vor, dass die Testpersonen sechs digitale Dokumente in einer suchmaschinenartigen Umgebung präsentiert bekamen. Diese sechs Dokumente behandelten verschiedene Facetten von Gesundheitsfolgen durch die Nutzung von Mobiltelefonen, und das intertextuelle Verhältnis war konfligierend, eröffnete also unterschiedliche, einander widersprechende Perspektiven auf die Thematik. Die Studierenden sollten diese Dokumente lesen, um auf dieser Basis einer fiktiven Freundin einen begründeten schriftlichen Rat zu den Gesundheitsrisiken bei der Nutzung von Mobiltelefonen zu geben. Diese Textproduktion erfolgte per Schreiben mit der Hand, und die Testpersonen hatten beim Schreiben keinen Zugriff mehr auf die Texte.

Während die Personen lasen, sollten sie ihre Gedanken verbalisieren. Die so entstandenen Lautdenk-Protokolle wurden auf vier Variablen kodiert und auf dieser Grundlage die Häufigkeit pro Person ausgezählt:

- *Verstehensüberwachung*: Hierunter fielen jene Aussagen, die auf das Überwachen des eigenen Verstehens und Verstehensprobleme hinwiesen.
- *Inhalte identifizieren und lernen*: Hierunter fielen Prozesse des Integrierens innerhalb und zwischen den Dokumenten sowie Elaborationen von Inhalten.

- *Informationen evaluieren*: Dies wurde dann kodiert, wenn die Personen erkennbar Informationen dazu nutzten, Inhalte oder ganze Dokumente zu interpretieren oder einzuschätzen.
- *Intertextuelles Verknüpfen*: Dies ist eine Kodierung, welche zu den anderen Kodierungen querliegt. Hier zählte, dass Personen Inhalte über die Dokumentengrenzen hinweg miteinander kombinierten und schlussfolgernd aufeinander bezogen.

Die Texte als Verstehensprodukte wurden auf mehrere Arten ausgewertet; primär interessierten zwei Maße: Erstens zählte man pro Text aus, wie viele *Verbindungen von Quelle und Inhalten* angeführt wurde, d. h. in wie vielen Fällen inhaltliche Aussagen mit ihrer Quelle benannt wurden. Dies ist ein Indikator für das Sourcing. Zweitens wurde die *argumentative Ausgewogenheit* geprüft, wozu die Texte in einem Ratingverfahren holistisch danach klassifiziert wurden, ob sie ausgewogen und textstrukturell angemessen die Strittigkeit des Themas wiedergaben. Dies entspricht eher einem Maß des intertextuellen Integrierens. Diese vier Prozess- und die beiden Produktmaße wurden per Korrelationsanalysen miteinander in Bezug gesetzt.

D) Was waren die Hauptbefunde?

Die Hauptergebnisse in Form von korrelativen Zusammenhängen stellt Abbildung 23 vor. Zunächst zu den *Prozessen*: Die drei kognitiven Strategien Inhalte identifizieren und lernen, intertextuelles Verknüpfen sowie Informationen evaluieren hingen auf der Objektebene der Kognitionen substantiell zusammen. Wer die eine Strategie bemühte, tat es auch bei den beiden anderen. Auf der Metaebene zu verortende Verstehensüberwachungen waren jedoch nur mit dem intertextuellen Verknüpfen statistisch signifikant verbunden: Wer mehr metakognitiv überwachte, stellte mehr intertextuelle Bezüge her.

Die Produkt- und Prozessrelationen zeigten ebenfalls Muster: Explizite Verbindungen von Inhalten und Quellen zum einen und argumentativ ausgewogenere Texte hingen mit zwei der drei kognitiven Strategien zusammen. Mehr intertextuelles Verknüpfen und häufigeres Evaluieren von Informationen schlugen sich in beiden Produktmaßen positiv nieder. Außerdem gab es noch einen Zusammenhang von Verstehensüberwachung und argumentativer Ausgewogenheit: Texte wurden dann qualitativ balancierter, wenn die Studierenden ihr Verstehen zuvor in höherem Maß überwacht hatten.

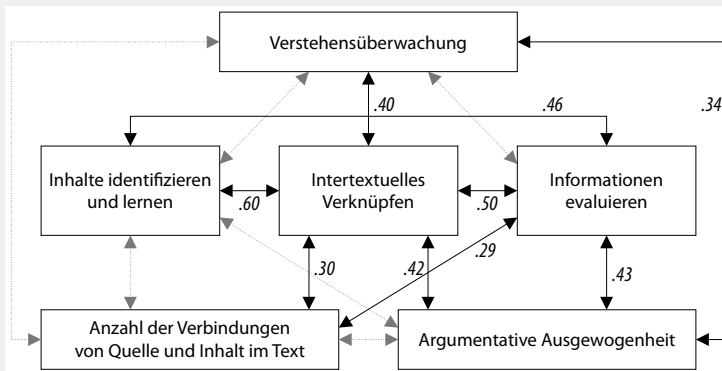


Abbildung 23: Zusammenhänge von strategischen Prozess- und Produktmaßen (oberer Kasten bezieht sich auf Metakognitionen, mittlere drei Kästen auf kognitive Lesestrategien, untere beide Kästen auf zwei ausgewählte Produktmaße; Quelle der eigenen Darstellung: Anmarkrud et al., 2014, S. 72, Koeffizienten beziehen sich auf Korrelationen; gestrichelte graue Linien geben nicht-signifikante Zusammenhänge an)

E) Diskussion

Die Studie verdeutlicht, dass Strategien im Verbund nachweisbar sind, dass sie aber auf der einen Seite untereinander nicht in gleichem Maß miteinander zusammenhängen. Auf der anderen Seite sind sie auch nicht zwingend mit Produktmaßen, in diesem Fall: Textprodukten, linear korreliert. Im Licht dieser Studie mit ihren beiden Produktmaßen sind kritische und intertextuelle Verstehensprozesse bei der Aneignung des Sachverhalts beide wichtig dafür, dass thematisch balancierte Wiedergaben von kontroversen Inhalten mit klarer Zuordenbarkeit der Aussagen zustande kommen. Zusätzlich erwies es sich für die prozess- und produktbezogenen Maße der Integrationsleistungen als günstig, wenn die Personen ihr Verstehen überwachten. Natürlich bedurfte es weiterer Prozesse wie jene des Schreibens, auf welche diese Studie nicht eingegangen ist.

Insgesamt differenziert sich damit das konstruktiv-responsive Lesen aus, es wird also als Konstrukt breiter. Das dokumentiert sich im Konkreten darin, dass sich hinter den einzelnen Gruppen von Lesestrategien teils Dutzende unterscheidbarer, für das Lesen von analogen und digitalen, singulären und multiplen Dokumenten geltende Vorgehensweisen verbergen (Afflerbach & Cho, 2009; Cho & Afflerbach, 2017), auf die hier aus Platzgründen nicht eingegangen werden konnte (s. dazu Philipp, 2015). Literarisches Lesen hingegen ist in diesen Strategien ebenso unterbelichtet wie multimodales Lesen, was durch zukünftige Forschungsarbeiten gelöst werden kann, um so blinde Flecken zu beseitigen.

6.3 Determinanten und interindividuelle Unterschiede

Aus dem weiten Feld der Forschung zum selbstregulierten Lesen und Lernen präsentiert dieses Teilkapitel drei (meta-)kognitive Merkmale, die systematisch mit differenziell ausgeprägtem Vorgehen und Leistungen in Verbindung stehen. Es handelt sich erstens um Leistungsgruppen (z. B. im Leseverstehen, aber auch in ihrer Leistungsfähigkeit oder in ihrem Zuwachs an Wissen in Experimentalstudien), also um Personen, die sich in ihren kognitiven Fähigkeiten unterscheiden. Entsprechend prüft das Teilkapitel [6.3.1](#), ob und in welchen Prozessmerkmalen sich Personengruppen mit divergierenden kognitiven Leistungsmerkmalen trennen lassen. Der zweite Zugang erfolgt über subjektive Theorien zur Wissenskonstruktion: die epistemischen Überzeugungen ([6.3.2](#)). Der dritte Teil der Darstellung ist einer Variable gewidmet, die in der Selbstregulation bis dato eher randständig behandelt wurde, doch zunehmende Beachtung erfährt, nämlich dem inhaltlichen Vorwissen ([6.3.3](#)).

6.3.1 Leistungsgruppen in kognitiven Fähigkeiten

In der empirischen Forschung zum selbstregulierten Lesen gibt es neben dem Ermitteln von Beiträgen selbstregulatorischer Fähigkeiten für die statistische Erklärung von interindividuellen Unterschieden zwei weitere Zugänge. Zum einen interessiert in manchen Studien, wie sich a priori oder a posteriori in ihren (Lese-/Lern-)Leistungen unterscheidende Gruppen darin unterscheiden oder darin gleichen, wie sie in Bezug auf die *häufigkeitsbezogene Ausprägung einzelner Strategien bzw. Strategiegruppen* selbstregulatorisch vorgehen. Diese Perspektive ist leitend in der Darstellung des Teilkapitels [6.3.1.1](#). Zum anderen gibt es eine komplementäre Perspektive, die sich daraus ergibt, von den Vorgehensweisen als *temporale Prozessverbünde* auszugehen, die prozessnah erhoben werden, und dann zu prüfen, in welchem Zusammenhang sie zu den später gezeigten Leistungen stehen. Ergebnisse dieser Forschung berichtet das Teilkapitel [6.3.1.2](#).

6.3.1.1 Unterschiede zwischen Leistungsgruppen

Mit einem differentiell-psychologischem Zugang operieren jene Studien, in denen interindividuelle Differenzen herausgearbeitet werden, in diesem Fall: im selbstregulatorischen Vorgehen bei der Bearbeitung von komplexen Aufgaben, in denen textbasiertes Lernen erfolgen soll. Kennzeichnendes Merkmal dieser Studien ist, dass mindestens zwei Gruppen miteinander kontrastiert werden und in aller Regel die Nutzungsfrequenz handlungsnah erfasster Strategien zwischen den Gruppen verglichen wird. Die Resultate aus solchen Studien, welche Differenzen vorfanden, lauten wie folgt:

- *Leistungsstärkere Personengruppen neigen zu einem stärker metakognitiv gesteuerten Vorgehen in der Koordination von Strategien auf der Metaebene der Selbstregulation.* Leistungsstärkere Personen zeichnen sich in zum Teil erheblichen Vorsprüngen dadurch aus, wie sie ihre Strategien metakognitiv regulieren. Solche Vorteile bestehen sowohl beim Planen von Strategien (Azevedo, Guthrie & Seibert, 2004; DiBenedetto & Zimmerman, 2010; Greene et al., 2011; Heirweg et al., 2019; Hilbert & Renkl, 2008) als auch bei der metakognitiven Überwachung von Strategieeinsatz und Verstehen (Azevedo, Guthrie & Seibert, 2004; Denton et al., 2015; Goldman et al., 2012; Greene & Azevedo, 2007a; Greene et al., 2011; Heirweg et al., 2019; Hilbert & Renkl, 2008; Lim et al., 2021; Rott & Gavin, 2015). Teils ergeben sich Merkmale des Zeitpunkts, wobei dies nur selten erforscht wurde: Erfolgreichere Personen neigen gerade zu Beginn zu mehr metakognitiven Strategien und zu weniger kognitiven Strategien (Paans et al., 2019).
- *Leistungsstärkere Personen wenden auf der Objektebene mehr kognitive Lern- und Lesestrategien an.* Differenzen schlagen sich zugunsten der leistungsstärkeren Personengruppen ebenso in jenen Strategien nieder, welche dazu dienen, Textinhalte zu verstehen und zu nutzen. Studien fanden mehr Anwendungen von a) Wiederholungen von Inhalten und Lektüren, also *Wiederholungsstrategien* (Heirweg et al., 2019; Lim et al., 2021); b) *Organisationsstrategien* wie dem Concept-Mapping, Zusammenfassen, Notizenanfertigen oder Unterstreichen (Greene et al., 2011; Greene, Moos et al., 2008; Heirweg et al., 2019; Slotte & Lonka, 1999); c) *Elaborationsstrategien*, also dem Kommentieren, vorwissensbasierten Anreichern und korrekten Verknüpfen von textuellen Inhalten und eigenem Wissen (Denton et al., 2015; Goldman et al., 2012; Greene & Azevedo, 2007a; Greene et al., 2011; Heirweg et al., 2019; Laing & Kamhi, 2002; Rott & Gavin, 2015; Seipel et al., 2017). Auch in anderen Kategorisierungen bzw. Bündelungen gab es Vorsprünge, z. B. bei der Sammelkategorie „effektive Strategien“ in der Studie von Azevedo, Guthrie und Seibert (2004), einem Konglomerat von Organisations- und Elaborationsstrategien, im Vergleich von „oberflächlichen“ bzw. auf Integration abzielenden Strategien (Ferguson-Hessler & Jong, 1990) sowie bei der Zählung von Strategien als Summenwert (DiBenedetto & Zimmerman, 2010). Eine Ausnahme von reinen Über- und Unterlegenheiten bildete die Studie von Rogiers et al. (2020a), in welcher vier aus retrospektiven Selbstauskünften ermittelte Gruppen in ihren Verbaldaten verglichen wurden. In dieser Studie zeigten sich weniger Über- und Unterlegenheiten denn spezifische Profile im Vorgehen – hinzu kommen ebenfalls nur geringe Differenzen in Verstehensleistungen (Rogiers et al., 2019). Solche Befunde sind bereits andernorts in methodisch ähnlichen Studien ermittelt worden (Wade et al., 1990).

Bei der Suche nach Differenzen, welche erklärtes Ziel dieses Teilkapitels ist, darf nicht vergessen werden, dass (statistisch) auffällige Häufigkeiten in der Strategienutzung nur einen Teil der empirischen Ergebnisse ausmachen. Ein häufiges Befundmuster besteht entsprechend darin, dass *Gruppen sich nicht überzufällig unterscheiden*. Das ist zum Teil auf geringe Samplegrößen zurückzuführen, da wegen des hohen Auswertungsaufwands häufig nur kleine Studienpopulationen dafür genutzt werden, selbstregulatorische Prozesse handlungsnah zu erfassen. In solchen Studien müssen Differenzen stark ausgeprägt sein, um statistisch signifikant zu sein. So mögen sich zum Teil die Ergebnisse aus Untersuchungen erklären lassen, in denen nicht die Differenz zwischen Gruppen vorgefunden wurde, sondern die Absenz von Unterschieden bzw. die breite Streuung von Vorgehensweisen markantes Ergebnis waren. Zwei Auffälligkeiten lassen sich benennen:

- *Es gibt trotz Leistungsgruppenunterschieden immer auch Ähnlichkeiten zwischen Leistungsgruppen.* So zeigen auch leistungsschwächere Personen ein zeitlich ähnliches Vorgehen in selbstregulatorischen Bearbeitungen von Aufgaben, dessen Grundbestandteile jenen anderer Leistungsgruppen ähnelte (Heirweg et al., 2020). Hinzu kommen die immer wieder in verschiedenen Experimenten beobachtbaren Befunde, dass sich in einzelnen Strategienutzungen im Sinne der alleinigen Auftretenshäufigkeit keine statistisch auffälligen Gruppendifferenzen vorfinden lassen (Greene, Moos et al., 2008; Lim et al., 2021; Linderholm & van den Broek, 2002; Merchie & van Keer, 2014b; Seipel et al., 2017).
- *Leistungsgruppen sind keine homogenen Gruppen, sondern es zeichnen sich in ihnen interindividuelle Unterschiede und eine Variabilität ab.* Besonders deutlich wird dies in Studien, in denen mittels diverser Instrumente einzelne Fälle herausgearbeitet werden (Hadwin et al., 2004; Hadwin et al., 2007; Malmberg et al., 2014; Malmberg et al., 2013). Selbst in leistungshomogenen Gruppen gibt es Unterschiede im Vorgehen von Personen (Neuenhaus et al., 2023). Doch auch in den Studien, in welchen Personengruppen miteinander verglichen wurden, ist ein stabiler Befund deskriptiver Statistik, dass es teilweise erhebliche Streuungen in den Maßen gab, die sich entweder in großen Standardabweichungen oder beträchtlichen Spannen zwischen Minimal- und Maximalwerten manifestieren (s. z. B. Denton et al., 2015; Lim et al., 2021; Rott & Gavin, 2015). Offenkundig variiert demnach das individuelle Vorgehen substanziell.

Ob und in welchem Maße sich Personengruppen in der Nutzungshäufigkeit von Strategien im selbstregulierten Vorgehen unterscheiden, ist im Licht der Empirie weniger eindeutig, da eine vergleichsweise große Zahl von Studien diesbezüglich keine durchgängigen Vorteile zugunsten leistungsstärkerer Personen nachweisen

konnten – teils gibt es dieses Muster sogar in ein und derselben Studie. Damit stellt sich – nicht nur hier, sondern im Grunde für die Strategieforschung allgemein (Dinsmore, 2017b) – die Frage danach, ob die Frequenz von Strategien allein entscheidend ist für den Erfolg beim Lesen und Lernen mit Texten (Fox, 2009). Das gilt umso mehr, als teils nur sehr geringe Korrelationen kognitiver Strategien mit Leistungsmaßen ermittelt wurden (Dent & Koenka, 2016). Daher lohnt es sich, Studien zu konsultieren, welche Muster in der Nutzung von Strategien durch Datenaggregationen ermittelt und diese mit Leistungsdaten verknüpft haben.

6.3.1.2 Rekursive, flexible Vorgehensweisen als Indikator glückenden, selbstregulatorischen Lesens

Ein Teilzweig der empirischen Leseforschung hat sich damit beschäftigt, aus synchron erfassten Verhaltensweisen wie Augenbewegungen, aber auch Verbaldaten in Lautdenk-Protokollen Muster im Vorgehen bei Leseaufgaben zu extrahieren. Es handelt sich also um aggregierende Datenanalysen, die zum Beispiel den temporalen Aspekt beim Vorgehen fokussieren, indem die Forscherinnen und Forscher ermittelten, wie linear oder nicht-linear die Experimentalpersonen vorgehen, wenn sie verschiedene kognitive Aktivitäten absolvieren. Dem liegt die Annahme zugrunde, dass solche Wechsel zwischen einzelnen Prozessen, wie sie typisch sind als Ausdruck der kognitiven Flexibilität, eine strategische Dimension aufweisen: Solche Wechsel bei Blickbewegungen gelten als Ausdruck selbstregulatorischer Verstehensprozesse (Catrysse et al., 2018; Rayner et al., 2006) bzw. als Indikator kognitiven Engagements (Miller, 2015).

In Studien wurden meist mittels Clusteranalysen Personengruppen danach gebündelt, dass ihre Verhaltensweisen sich pro Gruppe (Cluster) möglichst ähneln, sich die Gruppen allerdings untereinander unterscheiden. Anzahl und Merkmale der Gruppen fielen in den Studien anders aus, doch gab es studienübergreifend ein Kontinuum von Vorgehensweisen. Auf der einen Seite ist ein *strikt lineares, wenig flexibles Vorgehen* anzuführen, bei welchem ein sequenzielles Vorgehen ohne Wechsel beobachtbar ist. Auf dem anderen Ende des Kontinuums ist ein *durch häufige Wechsel und non-lineare Abfolgen von Aktivitäten gekennzeichnetes, ergo flexibles Vorgehen* typisch. Der Kontinuumslogik, die hier vereinfacht als eine einzige Dimension wiedergegeben ist, liegt zugrunde, dass sich die vorgefundenen Gruppen entlang dieser Dimension verorten lassen.

Unterschiedliche Vorgehensweisen wurden in verschiedenen Personengruppen und dort wiederum in Leistungsgruppen vorgefunden, z. B. Primarschulkindern (Coté et al., 1998; Jian & Ko, 2017; Mason et al., 2013), Sekundarschuljugendlichen (Bråten, Ferguson, Anmarkrud & Strømsø, 2013; Mason et al., 2015; Minguela et al., 2015; Rogiers et al., 2020a) und Studierenden (Bannert et al., 2014; Cromley & Wills, 2016; Goldman & Saul, 1990; Hyönä et al., 2002; Hyönä & Nurminen, 2006; Lim et al., 2021; Schmidt et al., 2020; Zhang et al., 2019). In

solchen Studien weisen die Befunde in aller Regel auf eine hohe Dynamik hin, die allgemeiner Natur ist. So ließ sich zeigen, dass wiederholtes Lesen und ein Zurückkehren zu einzelnen Textteilen keineswegs selten sind. Coté et al. (1998) ermittelten beispielsweise, dass ein mehrfaches Lesen von Textteilen schon bei Kindern der vierten und sechsten Klassenstufen eher die Regel als die Ausnahme ist. Dies steht in Einklang mit Eye-Tracking-Studien zum relevanzbezogenen Lesen, nach denen wiederholte Fixationen zielkongruenter Textteile typisch für ein erfolgreiches aufgabenbezogenes Lesen sind (s. Teilkap. [2.2.2.2](#), [2.2.2.3](#) und [2.2.2.5](#)). Hinzu kommt, dass Studien, die etwa mit Data-Mining-basierten Analysen arbeiten, in aller Regel hochkomplexe Vorgehensweisen extrahieren. In solchen Auswertungen dienen die Auftretenswahrscheinlichkeiten von Wechseln zwischen Aktivitäten als Basis für die Mustererkennung (Biswas et al., 2018). Diese Muster kennzeichnet zweierlei: Erstens wiederholen sich diverse strategische Prozesse einer Art häufig, und zweitens wechseln sich kognitive und metakognitive Prozesse ab, darunter oftmals in einer bidirektionalen Richtung (Azevedo & Witherspoon, 2008; Bannert et al., 2014; Engelmann & Bannert, 2021; Johnson et al., 2011; Lim et al., 2021; Sonnenberg & Bannert, 2015).

Doch nicht nur der Umstand, dass Personen unterschiedliche Vorgehensweisen bemühen, ist aus der selbstregulatorischen Perspektive von Belang, sondern auch die Frage, ob sich verschiedene Vorgehensweisen eher als hinderlich oder förderlich für Leseleistungen erweisen. Hier ist die Ergebnislage zu den Zusammenhängen relativ eindeutig und lässt sich in zwei Arten von Ergebnislinien bündeln:

- *Ein rekursives Vorgehen ist mit besseren, vor allem kognitiv anspruchsvollen Leseverstehensleistungen assoziiert.* Dies sind zum einen Haupteffekte von Gruppenvergleichen (Cromley & Wills, 2016; Hannus & Hyönä, 1999; Hyönä et al., 2002; Hyönä & Nurminen, 2006; Jian & Ko, 2017; Mason et al., 2013; Rogiers et al., 2020a; Schmidt et al., 2020; Zhang et al., 2019) bzw. Korrelationsanalysen (Wiley et al., 2009) und teils auch rein deskriptiven Studien (Bannert et al., 2014; Lim et al., 2021; Mateos & Solé, 2009; Solé et al., 2013). Dieses Ergebnismuster besteht zum anderen auffälligerweise dann noch stabil übergreifend über Studien hinweg, wenn man weitere Einflussfaktoren in den Analysen rechnerisch berücksichtigt (Bråten, Ferguson, Anmarkrud & Strømsø, 2013; Mason et al., 2015). Um dieses Muster darzustellen, wird im Kasten „Illustrierendes Beispiel 28“ ein ausführlicheres Exempel präsentiert.
- *Rekursive Vorgehensweisen hängen in der Regel nicht mit dem Erwerb bzw. dem Verstehen von singulären Fakten zusammen.* Mehrere Studien, die Leseleistungen differenzierten, machten darauf aufmerksam, dass faktuales Leseverstehen von meist explizit in Texten enthaltenen Informationen weder günstig noch ungünstig von einem rekursiven Vorgehen tangiert wird (Goldman & Saul, 1990; Mason et al., 2015; Minguela et al., 2015; Zhang et al.,

2019), wobei dies jedoch bei jüngeren Personen im Grundschulalter durchaus festgestellt wurde (Mason et al., 2013). Möglicherweise gibt es hier also Altersgruppendifferenzen, bei denen differenzielle Effekte auf verschiedene Leistungen in einzelnen Altersgruppen bestehen. Die ausbleibenden positiven Effekte des metakognitiv aufwändigen rekursiven Vorgehens auf oberflächenorientierte Lernleistungen sprechen dafür, dass Selbstregulation im Prozess Entsprechungen in der Verarbeitungstiefe hat, nämlich dergestalt, dass sich dieses Vorgehen erst dann bezahlt macht, wenn tiefere Verstehensleistungen erforderlich sind.

Diese Ergebnismuster machen auf eine gewisse Sensitivität aufmerksam: Rekursive Vorgehensweisen mit erneuter Konsultation von Textteilen als Ausdruck von Selbstregulation beim Lesen sind offenbar dann günstig, wenn für die jeweilige Personengruppe die zu erbringenden Leseleistungen anspruchsvoll sind. Das ist deshalb von Belang, weil diese Ergebnisse sich aus ganz unterschiedlichen Lese- studien speisen, die etwa das Integrieren von Bild und Text (Hannus & Hyönä, 1999; Mason et al., 2013, 2015) ebenso fokussierten wie das Lesen multipler Texte (Bräten, Ferguson, Anmarkrud & Strømsø, 2013; Schmidt et al., 2020).

Illustrierendes Beispiel 28

Der Nutzen der durchbrochenen Linearität für das tiefe Leseverstehen

A) Was war das Erkenntnisinteresse?

Die Studie widmete sich der Frage, inwieweit als selbstregulatorisch geltende rekursive Vorgehensweisen beim Lesen eines digitalen Sachtexts zwei verschiedene Verstehensleistungen unter Jugendlichen vorhersagen (Minguela et al., 2015).

B) Wer wurde untersucht?

In dieser Studie wurden 55 spanische Jugendliche am Ende der Pflichtschulzeit im Alter von 15 Jahren getestet, die aus drei verschiedenen Schulen stammten. Diese Jugendlichen wurden aufgrund geringen thematischen Vorwissens ausgewählt und repräsentierten zwei Gruppen im Leseverstehen, solche mit geringem Leseverstehen und solche mit hohem Leseverstehen. Diese Klassifizierung erfolgte auf der Basis eines vorab durchgeführten standardisierten Lesetests.

C) Welche Instrumente kamen zum Einsatz?

Die Jugendlichen lasen einen 680 Wörter langen Schulbuchtext zu den Folgen des Zweiten Weltkriegs. Zu diesem Text bearbeiteten sie nach dem Lesen zwei Tests, die unterschiedliche Lesefähigkeiten erfassten, ohne dass der Text noch zur Verfügung stand. Einige Fragen betrafen das Verstehen explizit gegebener Informationen (oberflächliches Leseverstehen), einige weitere das tiefe Leseverstehen auf der Basis von Inferenzen, also dem Verknüpfen von Informationen aus weiter entfernten Textteilen oder mit dem Vorwissen.

Durch die digitale Leseumgebung war es möglich, das Leseverhalten des in Textblöcke aufgeteilten Textes zu erfassen und die Linearität und Rekursivität als Ausgangspunkt für das Bestimmen von musterhaften Leseprozessen zu machen. Dabei interessierten rekursive Vorgehensweisen besonders, weil sie als Indikator metakognitiv überwachter Selbstregulation gelten. Damit wurde ergo prozedurale Selbstregulation erfasst.

Ein weiteres, eher produktbezogenes Maß war die Urteilsgenauigkeit im Sinne des metakognitiven Überwachens, die über Einschätzungsfragen nach jeder Testfrage erbeten wurde. Die Jugendlichen sollten einschätzen, wie gut sie glaubten, bei der just bearbeiteten Frage abgeschnitten zu haben. Über diese Urteile und die tatsächlich erbrachte Leistung wurde ein Maß gebildet, welches die Kalibration erfasste, nämlich ob die Testpersonen ihre Leistungen realistisch oder zu pessimistisch oder zu optimistisch einschätzten. Dazu wurden Lesetestergebnis und Einschätzung der eigenen Leistung ins Verhältnis gesetzt.

D) Was waren die Hauptbefunde?

Ausgehend vom Leseverhalten in einer speziellen Leseumgebung, in der das Leseverhalten per Tracking erfasst wurde, wurden vier Arten des Leseverhaltens ermittelt: 1) lineares Lesen (einmaliges Lesen des gesamten Textes von Anfang bis Ende; 22 Prozent der Fälle), 2) lineares wiederholtes Lesen (mehrere Durchgänge des linearen Lesens; 20 Prozent der Fälle), 3) lineares Lesen mit selektiver Relektüre (zunächst lineares Lesen mit anschließender nicht-linearer Relektüre einzelner Textteile; 25 Prozent der Fälle), 4) Relektüren während des Lesens (Relektüren bereits vor dem Erreichen des Textes, teils auch noch Relektüren nach dem Erreichen des Textendes; 33 Prozent der Fälle). Die beiden Gruppen von Personen mit hohem und niedrigem Leseverstehen waren ungleich auf diese Gruppen verteilt. Überproportional viele Jugendliche mit geringem Leseverstehen lasen rein linear, besonders wenige mit hohem Leseverstehen demonstrierten hingegen dieses Muster.

Bei dem Maß der metakognitiven Urteilsgenauigkeit zur erbrachten Leistung ergaben sich Differenzen bei den Leseleistungen. Die Jugendlichen schätzten im oberflächlichen Leseverstehen unabhängig von ihrer Gruppenzugehörigkeit zu besseren oder schlechteren Leserinnen und Leser akkurat ein. Beim tieferen Leseverstehen kam es hingegen zu Überschätzungen.

Die Studie prüfte in separaten Regressionsanalysen, ob oberflächliches und tieferes Leseverstehen durch das Vorgehen (prozessbezogenes Maß der Selbstregulation) oder durch die Urteilsgenauigkeit (produktbezogenes Maß) vorhergesagt wurde, und zwar unabhängig davon, welches Leseverstehen im standardisierten Lesetest die Personen erbracht hatten. Die Befunde weisen ein komplementäres Muster auf:

- Für das oberflächliche Leseverstehen war es rechnerisch unerheblich, ob die Personen linear oder rekursiv vorgegangen waren, hier zählte nur die Urteilsgenauigkeit. Je genauer die Leistungseinschätzung war, desto mehr Fragen hatten die Personen korrekt gelöst.

- Beim tieferen Leseverstehen war der Fall genau andersherum gelagert: Hier war die Urteilsgenauigkeit rechnerisch irrelevant, entscheidend war, dass die Personen das am stärksten rekursive Procedere verwendet hatten (Relektüren während des Lesens). Alle anderen Vorgehensweisen waren hingegen mit weniger korrekten Lösungen assoziiert.

E) Diskussion

Die Studie nutzte prozess- und produktbezogene Maße, um kognitiv divergierende Leseleistungen zu präzisieren, wenn Personen wenig Vorwissen aufweisen. Das Ergebnis: Produktbezogene Maße wie die metakognitiv überwachte eigene Leistung genühten für weniger anspruchsvollere Leistungen. Bei den anspruchsvolleren Leseleistungen war es hingegen günstiger, wenn die Personen sich im Prozess von der Linearität des Textes lösten und stattdessen gezielt Textteile erneut lasen. Anders gesagt: Eine gezielte Modifikation des eigenen Vorgehens im Sinne des rekursiven Vorgehens wirkt im Licht dieser Studie geradezu geboten für ein tieferes Leseverstehen. Vor allem macht die Studie aber darauf aufmerksam, dass verschiedene Aufgabentypen mit unterschiedlichen Effekten der Selbstregulation zusammenhängen.

6.3.1.3 Abschlussbemerkung

Studien mit lautem Denken oder anderen handlungsnahen Zugängen, selbstregulatorisches Vorgehen beim Lernen aus Texten zu erfassen, haben sich mehr und mehr der Forschungsfrage verschrieben, ob sich Personengruppen mit verschiedenen kognitiven Leistungen im Vorgehen unterscheiden. Ausgehend von einer ähnlichen Datenbasis, nämlich einer Sequenz von Strategien lassen sich verschiedene Variablen generieren. Ein vergleichsweise simpler Zugang besteht in der Auszählung von Häufigkeiten und statistischen Vergleichen der mittleren Ausprägungshäufigkeit. Hier ergeben sich in Vergleichen von leistungsstärkeren und -schwächeren Personengruppen zwei Auffälligkeiten: Leistungsstärkere Personen geben häufiger metakognitive Überwachungen zu Protokoll und nutzen mehr vorwissensbasierte Elaborationsstrategien wie das Schlussfolgern. Dieses Muster ist deshalb so hervorhebenswert, weil es bei einer anderen individuellen Determinante, nämlich dem Vorwissen, ganz ähnliche Auffälligkeiten gibt (s. Teilkap. 6.3.3). Es ist auch deshalb hervorzuheben, weil Inferenzen zu den Tiefenverarbeitungsstrategien gezählt werden (Alexander et al., 2018; Winne, 2018b), die für das Leseverstehen bedeutend, um nicht zu sagen: konstitutiv sind (Kendeou, 2015; van den Broek et al., 2015, s. Teilkap. 4.1.2.1).

Ein zweiter Zugang, mit der auf Sequenzen von distinkten Strategien basierenden Häufigkeitsauszählung umzugehen, besteht in der Betrachtung von Wechseln zwischen den Strategien, aus denen sich komplexere Verhaltensmuster extrahieren lassen. Solche Analysen geben Hinweise auf rekursive Vorgehensweisen von lesenden Personen. Zielgerichtete dynamische Prozesse des Verstehens auf der Objektebene mit Rückkopplungen zur Metaebene, die sich in metakognitiven

Kontroll- und Überwachungsstrategien manifestieren, sind mit besseren Leistungen verknüpft. Darin zeigt sich auch die Kontingenz der metakognitiven Selbstregulation, die aus theoretischer Sicht dann glückt, wenn Personen aufgrund ihres Strategiewissens flexibel, aber nicht beliebig auf Aufgaben reagieren und einen für sich sinnvollen Bearbeitungsprozess absolvieren und steuern (Winne & Marzouk, 2019). Offenbar glückt es leistungstärkeren Personen besser, durch häufigere Wechsel von kognitiven und metakognitiven Strategien ihren Erfolg beim Lesen voranzutreiben.

6.3.2 Epistemische Überzeugungen

6.3.2.1 Subjektive Theorien und Selbstregulation – einige theoretische Klärungen

Epistemische Überzeugungen sind bereits für die Objektebene des Verstehens als Einflussgrößen erkannt (s. Teilkap. 4.3.3). Im Feld der Selbstregulationsforschung werden epistemische Überzeugungen ebenfalls theoretisch als Determinanten modelliert und empirisch erforscht. Der Diskurs ist hierbei leicht anders als beispielsweise bei der Multiple-Documents-Literacy. Zum einen sind teils andere theoretische Konstrukte von Belang, zum anderen wird vor allem die Frage fokussiert, wie die Mechanismen der epistemischen Überzeugungen als metakognitive Standards für die Zielerreichung auf der Metaebene beschrieben und geprüft werden können.

Subjektiven Theorien, die in der Forschung auf sehr unterschiedliche Weise modelliert werden, indem sie sich auf implizite Überzeugungen zur Intelligenz (Dweck & Leggett, 1988) bzw. inzwischen auf die Formbarkeit kognitiver Fähigkeiten allgemein (Yeager & Dweck, 2020), Merkmale guten Lesens (Schraw & Bruning, 1996) oder zum Wesen und zur Herstellung von Wissen (Hofer & Pintrich, 1997) erstrecken, wohnt eine Gemeinsamkeit inne. Sie besteht darin, dass in wissenschaftlichen Modellen naive, eher starre und rigide Überzeugungen mit fortgeschrittenen und flexiblen, aber regelgeleiteten Überzeugungen kontrastiert werden. Solche Überzeugungen kann man als Kontinuum beschreiben, auf dem sich Personen mehrdimensional verorten lassen, und diese Überzeugungen bestehen selbst bei Erwachsenen noch. Je nachdem, wie elaboriert die Überzeugungen bzw. subjektiven Theorien sind, desto eher neigen Personen dazu, spezifische Strategien beim Lernen anzuwenden, weil sie dafür nötig wirken, die über die Überzeugungen präformierten Standards zu erreichen (Muis, 2007; Muis & Singh, 2018). Wenn jedoch von naiven oder elaborierten Überzeugungen die Rede ist und sich damit unweigerlich eine Konnotation in den etablierten Begrifflichkeiten ergibt, dann soll dies nicht negieren, dass sich Überzeugungen durchaus situativ adaptiv darstellen und im Falle „naiver“ Überzeugungen als funktionale

subjektive Theorien erweisen können (Sandoval et al., 2016). Wenn Personen sich beispielsweise Wissen zu einem für sie neuen Gegenstand anzueignen beginnen und sich dabei auf ihre Annahme stützen, Wissen werde von Autoritäten vermittelt, dann ist es aus selbstregulatorischer Sicht sinnvoll, sich gezielt solchen Autoritäten der Wissensherstellung und -vermittlung zuzuwenden, woraus sich beispielsweise große Teile des Sourcings legitimieren (s. Teilkap. 3.1.3.1). Als Heuristik im Sinne eines Kontrasts werden naive und elaborierte Überzeugungen im Folgenden jedoch verwendet, um damit Muster besser darstellen zu können.

Was mit naiven und elaborierten Überzeugungen gemeint ist, lässt sich domänenspezifisch exemplarisch beschreiben. Im Falle des Lesens unterscheiden beispielsweise Schraw und Bruning (1996) transmissive von transaktionalen Überzeugungen. Die erstgenannte Variante zeichnet sich durch eine vonseiten des Textes klar vorgegebene Bedeutungsübertragung aus, also durch eine vermeintlich objektive Transmission von Bedeutung, in welcher der lesenden Person lediglich die Bedeutungsrekonstruktion zugewiesen wird. Demgegenüber besteht die transaktionale Überzeugung darin, dass auch die lesende Person ko-konstruktive Eigenleistungen bei der Bedeutungsherstellung zu erbringen hat. Transmissive Überzeugungen sind demnach als naiv und transaktionale Überzeugungen als fortgeschritten insofern zu werten, weil die transaktionalen Überzeugungen eine erhebliche größere Schnittmenge mit elaborierten Modellen des Leseverstehens aufweisen und die darin postulierte Komplexität von lesendenseitigen, textuellen und kontextuellen Merkmalen adäquater abbilden (McNamara & Magliano, 2009). Zudem korrespondieren solche Überzeugungen mit besseren Leseleistungen (Mason et al., 2006; Schraw & Bruning, 1996), allerdings hängen auch transmissive Überzeugungen durchaus mit Leistungsindikatoren positiv zusammen (Mateos et al., 2011).

Insgesamt bestehen auf breiter empirischer Basis metaanalytische Hinweise darauf, dass subjektive Theorien mit diversen schulischen und hochschulischen Leistungen schwach positiv korrelieren, seien es epistemische Überzeugungen (Greene, Cartiff & Duke, 2018), seien es die subjektiven Theorien zur Intelligenz als fixierte vs. als formbare Größe (Burnette et al., 2013; Costa & Faria, 2018; Sisk et al., 2018). Es ist auch im Lichte solcher relativ geringen Zusammenhänge zwischen Überzeugungen als Dispositionen und Leistungsmaßen anzunehmen, dass subjektive Theorien eher distale Prädiktoren von Leistungen wie im Lesen sind, da sie nicht direkte Einflüsse aufweisen dürften, sondern als stabile Merkmale eher vermittelt über weitere individuelle Merkmale wie den Umgang mit schwierigen Situationen und Fehlschlägen (Molden & Dweck, 2006), negativen Emotionen (Burnette et al., 2013) oder das Verfolgen von verschiedenartigen Zielen (Burnette et al., 2013; Dweck & Leggett, 1988) auf selbstregulatorische Verhaltensweisen in konkreten Situationen wirken (s. Teilkap. 6.1.1).

Diese mediierenden Effekte sind in Theorie und Empirie verstärkt bearbeitet worden, und dabei wurden zwei Muster beschrieben und erforscht:

- *Fortgeschrittene Überzeugungen* werden als *adaptive und günstige Varianten* aufgefasst, weil sie *eine höhere Anstrengungsbereitschaft und eine intensivere Auseinandersetzung mit dem Gegenstand* erwarten lassen.
- Demgegenüber sind Personen mit eher *naiven Überzeugungen* eher auf die *Vermeidung von potenziell selbstwertgefährdenden, die eigenen Defizite zutage tretenden Situationen bedacht* (Dweck & Leggett, 1988; Schwinger et al., 2022).

Diese theoretischen Annahmen spiegeln sich in empirischen Studien wider. Die Metaanalyse von Burnette et al. (2013) hat deutlich empirisch herausgearbeitet, dass erstens normativ ungünstige subjektive Theorien mit Tendenzen der Aufwandsreduktion und ungünstigen Zielen beim Lernen einhergehen, welche mit geringeren Leistungen korrespondieren. Zweitens sind die Annahmen zum Erfolg (Erwartungskognitionen) beim Lernen ebenfalls selbst ungünstig, wenn Personen unvorteilhafte, naive Annahmen zur Formbarkeit kognitiver Fähigkeiten hegen. Drittens sind zudem die Emotionen beim Lernen umso ungünstiger, je rigider die subjektiven Theorien sind, und auch dies schlägt sich in schlechteren Leistungen nieder.

Subjektive Theorien und Überzeugungen haben nunmehr das genuine Erkenntnisinteresse der Selbstregulationsforschung auf sich gezogen. Schommer-Aikins (2004) betont beispielsweise die allgemeine Notwendigkeit, subjektive Theorien als Ausschnitt bzw. als ein eigenes System innerhalb eines umfassenderen, auch kontextsensitiven Systemverbunds der Selbstregulation zu konzeptualisieren, um Wechselwirkungen zwischen kontextuellen und individuellen (Prozess-)Merkmalen präziser zu fassen. Diese Forderung von Schommer-Aikins (2004) ist breit gefasst, doch es gibt auch spezifischere Überlegungen. Aus den diversen subjektiven Überzeugungen haben *epistemische Überzeugungen* als ein Ausschnitt der subjektiven Theorien viel Zuwendung aus der Wissenschaft erfahren. Epistemische Überzeugungen sind subjektive Theorien dazu, was Wissen ist bzw. wie es hergestellt wird (s. dazu Teilkap. 4.3.3), und sie werden als einflussreiche individuelle Merkmale für das Leseverstehen untersucht (Bråten & Strømsø, 2020; Bråten, Strømsø & Ferguson, 2016).

Solche subjektiven wissensbezogenen Annahmen werden in jüngerer Vergangenheit in der Modellierung von selbstreguliertem Lernen berücksichtigt und in variierendem Maß Gegenstand der Modelle. Diese Berücksichtigung erfolgt auf zweierlei Weise: Entweder werden epistemische Überzeugungen in eigenen (in ihrer Grundanlage bereits synoptischen) Modellen aufgenommen (Muis, 2007; Muis et al., 2018) oder sie werden als dezidierte Erweiterungen des COPES-Modells ergänzt (Bromme, Pieschl & Stahl, 2010; Greene, Muis & Pieschl, 2010). Diese Zugänge eint ihre Prämisse, dass sie epistemische Überzeugungen als kognitive Bedingungen beschreiben und von einem wechselseitigen Zusammenwirken von sowohl kognitiven als auch metakognitiven Prozessen ausgehen

(Barzilai & Zohar, 2014). Entscheidend ist hierbei, dass die Art, wie Personen die Eigenschaften von Wissen (statisch vs. veränderlich; komplex vs. einfach) bzw. zur Wissensherstellung (internale vs. externale Wissensherstellung; nach Regeln des „gesunden Menschenverstands“ vs. nach komplexen Regeln der Wissenschaft) für sich wahrnehmen, die Formulierung von zielbezogenen Standards beeinflusst. Angenommen wird, dass epistemische Überzeugungen einen antizipierenden Charakter insofern haben, als sie den Aufwand für die Aneignung des zu lernenden Sachverhalts implizieren, welcher wiederum die Wahl von aus Sicht der Person geeignet wirkenden selbstregulatorischen Vorgehensweisen lenkt (Bromme, Pieschl & Stahl, 2010). Wer beispielsweise die Überzeugung hegt, Wissen sei grundsätzlich einfach und statisch, wird nach dieser Grundannahme wenig Aufwand betreiben, Widersprüchlichkeiten und deren Zustandekommen zu beleuchten. Es würden also oberflächliche Strategien genügen. Bei einer Person, die Wissen als komplex und veränderlich begreift, sind andere, aufwändiger anmutende und stärker metakognitiv überwachte Strategien wahrscheinlicher.

Der Effekt der Überzeugungen auf Zielsetzungsprozesse ist der Punkt, in dem sich die Modelle je nach Geltungsbereich der epistemischen Überzeugungen einig sind. Bestehende theoretische Zugänge unterscheiden sich allerdings darin, ob der Einfluss der epistemischen Überzeugungen primär lediglich die Zielsetzungsphase bei der Selbstregulation betrifft oder sämtliche Phasen (Bromme, Pieschl & Stahl, 2010). Die Arbeitsgruppe um Bromme postuliert, dass epistemische Überzeugungen ganze Selbstregulationszyklen beeinflussen, weil sie durchgängig als flexibel zu nutzende Standards im Hintergrund wirken und insbesondere bei komplexen Aufgaben ihre Wirkung entfalten, indem sie eine angemessene, kontinuierliche Kalibration erlauben (Bromme, Pieschl & Stahl, 2010). Wer fortgeschrittene epistemische Überzeugungen hat, die dem Charakter solcher Aufgaben zudem mehr entsprechen (etwa indem die Lerninhalte komplex sind oder weil mehrere, durchaus widersprüchliche Aussagen zu Themen zu verarbeiten sind), und ferner metakognitives Strategiewissen aufweist, dürfte zu einem adaptiveren Gesamtverfahren neigen, welches sich nicht im Setzen angemessener Ziele zu Beginn eines Selbstregulationszyklus erschöpft.

Zusammengefasst: Subjektive Theorien fungieren also als Filter, wie Personen Aufgabenerfordernisse wahrnehmen, weil sie allgemeine persönliche Annahmen zu generischen Merkmalen von kognitiven Prozessen und Produkten sowie zur Funktionsweise von Kognitionen beinhalten. Auf dieser (impliziten) Basis entstehen dann (planerische) Einschätzungen der Person, welche Arten von korrekten Vorgehensweisen zielführend wirken, weil solche Überzeugungen die Zielsetzungen betreffen und beeinflussen. Je nach Konfiguration der mehrdimensionalen Überzeugungen sind individuelle Interpretationen von Zielsetzungen erwartbar. Hierbei besteht bei aller noch zu attestierenden Offenheit, wie die Wirkmechanismen genau zu modellieren sind, Grund zur Annahme, dass die metakognitiv

repräsentierten Ziele auch Ausdruck der epistemischen Überzeugungen sind, welche dadurch indirekt auf die Selbstregulation einwirken.

6.3.2.2 Empirische Perspektiven auf Prozesse

Obwohl inzwischen die Forschungsaktivitäten zum Zusammenhang von epistemischen Überzeugungen reichhaltigere Befunde hervorzubringen begonnen haben (s. Muis & Singh, 2018, für einen Überblick), so sind prozessual ausgerichtete Studien noch immer vergleichsweise rar. Dennoch lassen sich zwei Muster extrahieren, die als Gegenüberstellung von Personen mit naiven und fortgeschrittenen Überzeugungen darstellbar sind:

- *Personen mit fortgeschrittenen epistemischen Überzeugungen zeigen eine tiefere Auseinandersetzung im Prozess mit dem Lerngegenstand.* Ein Hinweis darauf kommt aus Studien, in denen dem Wissen widersprechende Texte zum Einsatz kommen, um Alltagsvorstellungen inhaltlich zu korrigieren (sog. „Conceptual Change“). Solche Studien sind deshalb besondere Kristallisationspunkte der Forschung zu Überzeugungen, weil sie sich nachweislich für die Korrektur von (falschem) Wissen und weiteren kognitiven Merkmalen anbieten (Aleksavičiūtė et al., 2023; Schroeder & Kucera, 2022; Zengilowski et al., 2021). Hier engagierten sich Personen mit günstigeren epistemischen Überzeugungen stärker in adaptiven Prozessen. Dies zeigte sich generell in einer häufigeren Nutzung von Lesestrategien im Allgemeinen (Kardash & Howell, 2000) und mit einzelnen Strategien im Besonderen. Dazu zählen: mehr metakognitive Überwachung des Fortschritts beim Verstehen (Kardash & Howell, 2000) sowie im Falle kognitiver Strategien mehr Klärungen von Bedeutung auf lokaler Ebene (Kardash & Howell, 2000), mehr globale Inferenzen (Kardash & Howell, 2000) und eine stärkere Beachtung von Widersprüchlichkeiten in den Texten (Kendeou et al., 2011, Experiment 1). Andere Studien außerhalb des Conceptual-Change-Forschungsparadigmas ergaben, dass Personen mit fortgeschrittenen epistemischen Überzeugungen ebenfalls grosso modo zu mehr selbstregulatorischen Strategien neigen (Greene, Copeland et al., 2018; Greene, Costa et al., 2010; Greene et al., 2014). Elaborierte Überzeugungen korrespondieren z. B. damit, dass Personen Aufgabenschwierigkeiten und geeignete Strategien angemessener einschätzen (Bromme et al., 2014; Stahl et al., 2006), ihr Lesen stärker metakognitiv überwachen (Muis et al., 2011), planen (Stahl et al., 2006) und regulieren (Muis & Franco, 2010, Studie 2). Außerdem ist das Vorgehen von Personen mit fortgeschrittenen Überzeugungen davon gekennzeichnet, dass sie Lerngegenstände zu einem größeren Abdeckungsgrad konsultieren (Pieschl et al., 2008), mehr kognitive Organisationsstrategien wie Unterstreichen sowie Zusammenfassen nutzen (Ioannou & Iordanou, 2020, Studie 2) und die

logischen Zusammenhänge zwischen textuellen Aussagen stärker evaluierend kommentieren (Iordanou et al., 2019). Bei Internetsuchen ist zudem auffällig, dass Personen mit fortgeschrittenen Überzeugungen rekursiv vorgehen und die Informationen von Websites und der Suchmaschine stärker abgleichen (Hsu et al., 2014).

- *Personen mit naiven epistemischen Überzeugungen überwachen ihr Verstehen auf eher oberflächliche Art und neigen zu eher oberflächlichen kognitiven Strategien.* Einen zunächst positiv wirkenden Befund erbrachte die Studie von Trevors und Muis (2015). In ihr artikulierten Personen mit einem Profil an weniger elaborierten epistemischen Überzeugungen mehr metakognitive Einschätzungen zum Verstehen in Lautdenk-Protokollen. Seinen prekären Charakter erhält dieses Ergebnis jedoch dadurch, dass die Verstehensüberwachung sich auf eher oberflächliche einzelne Fakten bezog und nicht auf ein Gesamtverständnis. Das Forschungsteam sah entsprechend darin ein unkritisches Akzeptieren von einzelnen Informationen denn eine tatsächliche tiefergehende Überprüfung eines sich entwickelnden mentalen Modells bei einem dem Vorwissen widersprechenden Gegenstand. In einer weiteren Studie zeigte sich, dass für die Domäne unangemessene epistemische Überzeugungen sogar mit einer verringerten metakognitiven Überwachung einhergingen (Muis et al., 2011). Betrachtet man die kognitiven Strategien, so ergeben sich ebenfalls Auffälligkeiten. Beispielsweise lesen Personen mit naiven Überzeugungen Texte schneller und wiederholen die Inhalte verbatim häufiger (Ioannou & Iordanou, 2020, Studie 2).

Die bisherige Darstellung suggeriert, dass es deutliche Zusammenhänge zwischen epistemischen Überzeugungen und handlungsnah erfassten Indikatoren des selbstregulierten Lesens gibt. Zur Wahrheit gehört auch, dass es Interaktionen zwischen epistemischen Überzeugungen und der Anwendung von Strategien gibt (Greene, Costa et al., 2010), die Zusammenhänge also nicht so linear sind, wie Korrelationsanalysen und Gruppenvergleiche es andeuten, die z. T. auch nicht-signifikante Ergebnisse als Hauptbefunde erbrachten (Kardash & Howell, 2000). Ähnliches gilt für die Einschätzung von Aufgabenmerkmalen, die teils nicht mit epistemischen Überzeugungen zusammenhängen (Pieschl et al., 2014). Diese Widersprüchlichkeit und Uneindeutigkeit, die dem Forschungsfeld inhärent ist, spiegelt der Kasten „Illustrierendes Beispiel 29“ wider.

Illustrierendes Beispiel 29

Hängen epistemische Überzeugungen mit dem selbstregulierten Lernen in Prozessen und Produkten zusammen?

A) Was war das Erkenntnisinteresse?

Die kanadische Studie (Chevriert et al., 2020) war einem doppelten Erkenntnisinteresse zu den Effekten von epistemischen Überzeugungen gewidmet: der Korrespondenz erstens mit geplanten und durchgeführten Prozessen und zweitens den Produkten, wenn Studierende zwei Texte lesen und im Anschluss gedächtnisbasiert Fragebatterien mit variierender Schwierigkeit beantworten.

B) Wer wurde untersucht?

66 kanadische Studierende, ca. 70 Prozent weiblich, im Mittel knapp 21 Jahre alt, die Mehrheit mit Englisch als Muttersprache, nahmen teil. Diese Studierenden stammten aus diversen Studiengängen, die überwiegende Mehrheit aus sozialwissenschaftlichen Studiengängen.

C) Welche Instrumente kamen zum Einsatz?

Für die *epistemischen Überzeugungen* kam eine Batterie von Skalen zum Einsatz, welche mehrere klassische Dimensionen umfasste, nämlich die Veränderlichkeit und die Komplexität des Wissens sowie die Konstruktion und Begründung des Wissens über regelgeleitete Prozesse. Die Antworten bei den Items wurden zu einem Gesamtwert aggregiert, bei dem höhere Werte stärker ausgeprägte konstruktivistische Überzeugungen implizieren. Höhere Werte weisen demnach auf fortgeschrittene Überzeugungen hin.

Vor dem Bearbeiten der eigentlichen Aufgaben wurde mit einem Fragebogen mit drei offenen Fragen erfasst, wie die Personen die bevorstehende *Aufgabe definieren*, *welche Ziele sie haben und wie sie sich vorzubereiten gedenken (Planen des Vorgehens)*. Teil der Fragestellung war, welcher Aufgabentypus (s. u., Produktdaten) unmittelbar bevorstand. Jede der Antworten wurde inhaltsanalytisch kodiert, um entweder einer oberflächenorientierten bzw. simplen resp. einer tiefenorientierten, komplexen Kategorie zugewiesen zu werden. Diese offenen Fragen wurden sowohl für Multiple-Choice-Fragen als auch für die Inferenzverifikationsaufgabe gestellt.

Die *Prozesse beim Lesen* wurden parallel zum Lesen des aktuellen Texts als Verbalprotokoll erfasst und in zweierlei Hinsicht kodiert. Erstens interessierte man sich dafür, ob und wie viele *epistemische Kognitionen* in den Lautdenk-Protokollen vorkamen. Zweitens kodierte man die Strategien auf zweierlei Weisen: *einfache Strategien* und *komplexe Strategien*. Einfache Strategien waren das Paraphrasieren oder erneute Lesen, wobei maßgeblich war, dass singuläre Fakten erkannt und memoriert wurden. Komplexe Strategien beinhalteten metakognitive Überwachungen und die vorwissensbasierte Integration von Informationen, etwa durch Evaluation, Formulieren von Fragen, Inferenzbildungen und Zusammenfassen.

Die *Produkte* des Verstehens wurden mit zwei Maßen bestimmt, welche zwei Ebenen bzw. Verarbeitungstiefen des Verstehens maßen. Die *Multiple-Choice-Fragen* zielten auf das Wiedererkennen von paraphrasierten Inhalten, die *Inferenzverifikationsaufgaben* auf die Verknüpfung von Informationen aus dem jeweiligen Text. Die Inferenzverifikationsaufgaben waren anspruchsvoller. Für jeden Test lasen die Personen jeweils einen kurzen Sachtext, also zwei Texte pro Person.

Die Auswertung erfolgte bei den epistemischen Überzeugungen und den Kognitionen während der Aufgabenbearbeitung über Gruppenvergleiche. Zusätzlich wurden bei Analysen mit metrischen Daten Korrelationen berechnet.

D) Was waren die Hauptbefunde?

Hinsichtlich der Ausprägung konstruktiver epistemischer Überzeugungen wurden über einen Mediansplit zwei Gruppen gebildet, um Gruppenvergleiche vorzunehmen, welche die *Aufgabendefinition, Ziele und das geplante Vorgehen* betrafen. Unterschieden wurden Personen mit geringen und hohen Überzeugungen. Es ließen sich jedoch keine Effekte auf die Elaboriertheit von vorgängigen Plänen, Zielen und Vorgehensweisen feststellen. Stattdessen ergaben sich jedoch positive Zusammenhänge der höher ausgeprägten epistemischen Überzeugungen und *Prozessdaten*, und zwar in drei Prozessmaßen: der Anzahl von epistemischen Kognitionen und der Tendenz, in den beiden Leseaufgaben jeweils komplexe Strategien zu bemühen. Korrelationen mit simplen Strategien ließen sich nicht bestätigen.

Auffällig war zusätzlich der Befund, dass die Personengruppen, welche entweder epistemische Kognitionen während der Aufgabenbearbeitungen geäußert hatten oder nicht, sich in den Strategien bei den anspruchsvolleren Aufgaben unterschieden. Konkret divergierten sie darin, dass sie in den anspruchsvolleren Aufgaben, also jenen mit der später zu bearbeitenden Inferenzverifikationsaufgabe, mehr komplexe und weniger simple Strategien bemühten. In der Gruppe von Personen, die keine epistemischen Kognitionen äußerten, verhielt es sich genau andersherum. Bei den einfacheren Aufgaben gab es ein solches Muster hingegen nicht.

Hinsichtlich der *Produktdaten* zeigten sich Zusammenhänge nur bei einem Maß, und zwar bei den Multiple-Choice-Aufgaben. Von denen hatten die Studierenden jeweils mehr korrekt beantwortet, wenn sie vor der Gesamttestung mehr konstruktivistische epistemische Überzeugungen im Fragebogen geäußert hatten. Keine nennenswerten Korrelationen bestanden allerdings beim zweiten, kognitiv anspruchsvolleren Verstehenstest.

E) Diskussion

Die Bilanz der Studie zu den Effekten der epistemischen Überzeugungen fällt gemischt aus. Die Wahrnehmung von Aufgabenerfordernissen und ihren Implikationen für die bevorstehende Aufgabenbearbeitung war in dieser Experimentalstudie rechnerisch nicht mit der Ausprägung der epistemischen Überzeugungen verknüpft. Die Produkte waren lediglich partiell von den epistemischen Überzeugungen betroffen.

Markanter und konsistenter sind die Prozessdaten. Dispositionale Überzeugungen äußern sich tatsächlich in mehr während der Aufgabenbearbeitung geäußerten epistemischen Kognitionen und in komplexeren Strategien. Zudem treten komplexere Strategien dann häufiger auf, wenn Personen komplexe Aufgaben lösen, aber nur dann, wenn sie davon flankiert worden sind, dass mindestens einmal auch eine epistemische Kognition in den Lautdenk-Protokollen vorkam. Es wirkt also so, als wären Überzeugungen und Kognitionen tatsächlich im Sinne von Disposition und Teil des Verhaltens vereint, aber doch zum Teil daran gebunden, dass die Aufgaben komplexere Strategien erfordern. In dieser Hinsicht markiert die Studie einen ersten wichtigen Schritt für die an Wichtigkeit zunehmende empirische Frage, wie epistemische Überzeugungen als prozessbezogene Ressource fungieren. Dabei wäre es zukünftigen Studien überlassen, spezifische Beiträge einzelner epistemischer Überzeugungen in Verbindung mit individuellen und Kontextmerkmalen auszuleuchten.

6.3.2.3 Abschlussbemerkung

Subjektive Theorien von Leserinnen und Lesern, darunter solche dazu, welche Merkmale Wissen und Wissenskonstruktion aufweisen (epistemische Überzeugungen), gelten als entfernte („distale“) Einflussgrößen auf das selbstregulierte Lesen. Das hat mit ihrer Eigenheit zu tun, dass sie als implizite Filter die Wahrnehmung von Aufgaben und Textinhalten beeinflussen. Entscheidend ist, in welcher Ausprägung und in welcher Konfiguration die mehrdimensionalen epistemischen Überzeugungen innerhalb der lesenden Person vorhanden sind. Diese Überzeugungen weisen unterschiedliche Granularitätsgrößen auf und variieren in ihren Bezugspunkten und in ihrer entwicklungsbedingt erwartbaren Ausprägung (Chinn et al., 2011; Greene, Azevedo & Torney-Purta, 2008; Hammer & Elby, 2002) – kurzum: Sie sind als lose verbundenes, jedoch komplexes Set subjektiver Theorien zu denken.

Trotz dieser Komplexität hat sich in der Forschungsliteratur eine heuristische Differenzierung von naiven und fortgeschrittenen epistemischen Überzeugungen etabliert. Diese nimmt auf, inwieweit die subjektiven Laientheorien sich dem Standard wissenschaftlichen Konsens' über die Eigenschaften von Wissen und Wissensproduktion annähern. Wenn sich Personen darin unterscheiden, für wie komplex, dynamisch, regelgeleitet hergestellt etc. sie Wissen halten, dann wird das in der Selbstregulationsforschung als eine Disposition bzw. als (meta-)kognitive Bedingung angenommen. Gerade das Modell von Muis (2007) hat anhand verschiedener Prämissen dargelegt (s. auch Muis & Singh, 2018, für eine Aktualisierung dieser Prämissen), dass epistemische Überzeugungen in der frühen Phase der Aufgabendefinition speziell bei komplexeren Aufgaben zum Tragen kommen. Aufgrund verschiedener Wissensbestände und weiterer kognitiver Ressourcen wie eben den epistemischen Überzeugungen entwickeln Personen ihre Standards für das zu erreichende Ziel. Solche epistemischen Standards werden dann im

Verlauf als (veränderlicher) internaler Benchmark für Feedbackschleifen genutzt, um die Zielerreichung zu überwachen und dafür nötige Vorgehensweisen dynamisch zu planen und zu aktivieren. Dies sind Funktionsweisen, die kompatibel mit dem COPES-Modell sind.

Empirische Studien mit Fokussierungen auf handlungsnah erfasste Prozesse unterstützen die Annahmen. In diesem Teilkapitel ließen sich allgemeine Zusammenhänge von mehr auf Tiefendurchdringung abzielenden metakognitiven und kognitiven Strategien feststellen, wenn Personen fortgeschrittene epistemische Überzeugungen besitzen. Sie scheinen damit der Komplexität der Aufgaben eher gerecht zu werden, denn es handelte sich in den Studien ausnahmslos um kognitiv anspruchsvolle Aufträge. Wie unter einem Brennglas zeigte sich, dass selbstregulatorisches Vorgehen und epistemische Überzeugungen sich in einem Passungsverhältnis zu befinden scheinen: Anspruchsvolle Strategien korrespondierten teils mit fortgeschrittenen epistemischen Überzeugungen. Dies illustriert, dass epistemische Überzeugungen und Selbstregulation eine Beziehung zueinander aufweisen, die allerdings bislang theoretisch und empirisch noch unterbelichtet ist.

6.3.3 Thematisches Vorwissen

6.3.3.1 Zur Verortung des Vorwissens in Modellen des selbstregulierten Lernens

Das Wissen einer Person zu einem Gegenstand innerhalb einer Domäne wird in verschiedenen Modellen selbstregulierten Lernens als expliziter Bestandteil aufgeführt und den individuellen kognitiven Ressourcen zugeordnet. Die Rolle des Vorwissens war im Vergleich der acht Modelle im Teilkapitel [6.1.1](#) kein eigenes Thema, auch nicht in den Vergleichsdimensionen. Daher erfolgt an dieser Stelle noch einmal ein kurzer Gang durch die Modelle dahingehend, ob und, falls ja, wie das Vorwissen in den Modellen vorkommt und wie es mit anderen Komponenten und Prozessen interagiert. Es lassen sich dabei vier ganz unterschiedliche Weisen rekonstruieren, ob und in welchem Maße das inhaltliche Vorwissen gewichtet wird und wie es das Selbstregulationsgeschehen beeinflusst.

- *Vorwissensaktivierung als früher Prozess der Selbstregulationszyklen:* Pintrich (2000) führt das Vorwissen in seiner Vier-x-vier-Matrix als Bestandteil der kognitiven Dimension auf und dort in der Planungs- und Aktivierungsphase, also der ersten Phase seines Modells. Die Vorwissensaktivierung kann automatisch und ungesteuert oder gezielt und damit strategisch erfolgen. Wie sich die Vorwissensaktivierung auf die weiteren Phasen auswirkt, bleibt im Modell implizit. Jedoch wird in der darauffolgenden Phase

(Überwachen) gemäß Modellbeschreibung nahegelegt, dass das aktivierte Vorwissen dazu genutzt wird, das Verstehen bei der Beurteilung des Gelernten (Judgements of Learning) zu überwachen. Auch wenn damit die Rolle des Vorwissens unterspezifiziert ist, wird es jedoch als früher Bestandteil im zeitlichen Geschehen im Sinne einer Ressource verortet.

Im COPES-Modell von Winne und Hadwin (1998) wird das Vorwissen ebenfalls in der ersten Phase benannt (s. Teilkap. [6.2.1.2](#)) und außerdem als kognitive Bedingung im Modell aufgenommen. Das Vorwissen bzw. dessen Einschätzung dient im COPES-Modell dazu, eine möglichst präzise Aufgabendefinition vorzunehmen, welche wiederum – dann als kognitives Produkt der ersten Phase entstanden und zum Evaluationsstandard avanciert – für metakognitives Überwachen den Vergleichsmaßstab bildet (Winne, 2001). Insofern ähneln sich Pintrichs sowie Winnes und Hadwins Modelle darin, dass angemessenes Vorwissen dabei hilft, einen verlässlichen, gleichwohl dynamischen Soll-Zustand zu definieren, welcher für das metakognitive Überwachen als interne Ressource fungiert. Die Wirkung des Vorwissens wird demnach prominent, aber vermittelt über seine Nutzung für metakognitive Zielklärungen angenommen.

- *Vorwissen als Ersparnis von Selbstregulation bei Personen mit hohem Domänenwissen.* Borkowski et al. (2000) gehen in ihren Prämissen, nämlich dem Idealtypus des guten Informationsverarbeiters, davon aus, dass selbstregulierte Lernerinnen und Lerner unter anderem ein hohes Welt- und Domänenwissen haben. In ihrem Modell, welches sie aus einer Quasi-Entwicklungslogik mit mehr und komplexer interagierenden Bestandteilen sukzessive elaborieren, taucht das inhaltliche Vorwissen jedoch erst sehr spät auf, nämlich in der letzten und komplexesten Variante des Modells. In diesem Fall sprechen Borkowski et al. (2000, S. 9) von einer „Bypass“-Funktion, also einer Überbrückung: Ist das Vorwissen hoch, sind den Autorinnen und Autoren zufolge die Personen nicht mehr darauf angewiesen, komplexe Selbstregulationsprozesse zu initiieren, zu überwachen und zu modifizieren, sondern können aufgrund der Wissensbestände selbst komplexere Aufgaben lösen. In diesem Fall hat das Vorwissen eine kompensierende Funktion, welche aufwändige Selbstregulationsprozesse erübrigt. Derlei ist jedoch an die Bedingung hohen Wissens gebunden.
- *Vorwissen als eher distaler Bestandteil der Selbstregulationsprozesse:* Muis (2007) weist dem Vorwissen einen distalen Status für das selbstregulierte Vorgehen zu, indem das Vorwissen für sie als ein Vorläufer epistemischer Überzeugungen gilt. Da in ihrem Modell die epistemischen Überzeugungen Bedingungen im Sinne individueller kognitiver Merkmale darstellen, sind Effekte des Vorwissens nur mittelbar anzunehmen. In der Weiterentwicklung des Modells, in dem epistemische Emotionen ergänzt wurden (Muis et al., 2018), wird das Vorwissen wiederum anders im Modell behandelt, nämlich

als Vergleichsgrundlage für neue Informationen, wobei dieser Vergleich zu den epistemischen Emotionen führt. In beiden Fällen kommt dem Vorwissen eine je eigene Funktion zu, die in beiden Varianten des Modells nicht identisch ist. Jedoch ist in beiden Modellvarianten die Art unbestimmt, wie das eigene Vorwissen als Einflussfaktor auf das selbstregulatorische Geschehen präzise einwirkt.

- *Vernachlässigung inhaltlichen Vorwissens bei der Selbstregulation.* Zu guter Letzt gibt es auch Modelle der Selbstregulation, in denen das Konstrukt Vorwissen keine Rolle spielt. Besonders deutlich wird dies in den Modellen von Efklides (2011) sowie Zimmerman (2000), welche das inhaltliche Vorwissen vollständig außen vor lassen. Das Modell von Boekaerts (2011) widmet sich ebenfalls diesem Konstrukt nicht mehr, selbst wenn in Vorläuferversionen das domänenspezifische Wissen expliziter Gegenstand der grafischen Modelldarstellung, aber nur impliziter Gegenstand der schriftlichen Modellexplikation war (Boekaerts & Niemivirta, 2000). Hadwin et al. (2018) benennen das inhaltliche Vorwissen von einzelnen Personen bzw. den Mitgliedern von Gruppen als individuelle und soziale Bedingung des selbstregulierten Lernens zwar, in der Ausführung der Einflüsse und Wechselwirkungen bleiben sie jedoch eine Vertiefung schuldig, sodass es legitimierbar wirkt, dieses Modell der Kategorie der Vernachlässigung des Vorwissens zuzuordnen.

Die listenartige Darstellung hebt hervor, dass dem Vorwissen eine eher mittelbare Funktion zugewiesen wird, was freilich nur für jene Modelle gilt, welche das Vorwissen als Bedingung im weitesten Sinne beschreiben. Das Vorwissen fungiert in ihnen einerseits vorgängig für Zielformulierungen und andererseits synchron als metakognitive Bemessungsgrundlage.

Wie steht es um die empirischen Effekte des Vorwissens? Vorwissensbestände werden, wo sie in den Blick der lesenahen Selbstregulationsforschung geraten, häufig als Kovariate oder im Rahmen von Längsschnittstudien als Prä- und Posttestmessungen gemessen, um die Veränderungen im Erwerb von Wissen zu erfassen. Die Auswertungen erfolgen dann typischerweise auf eine von zwei Arten: entweder variablenzentriert (Wie ist das Verhältnis zwischen Vorwissen und Prozessmerkmalen des selbstregulierten Lernens mit Texten?; s. Teilkap. [6.3.3.2](#)) oder personenzentriert (Unterscheiden sich Personen mit divergierendem Niveau an Wissen darin, wie sie selbstreguliert vorgehen?; s. Teilkap. [6.3.3.3](#)).

6.3.3.2 Ein prozessualer Blick auf die Zusammenhänge von Vorwissen und Strategien

Die studienübergreifenden Muster zu den variablenzentrierten Analysen lassen sich unter zwei Gesichtspunkten bündeln:

- *Höhere Vorwissensbestände korrelieren mit mehr metakognitiven Aktivitäten.* Wenn Personen zu dem Sachverhalt ein höheres Wissen aufweisen, dann zeigen sich positive Zusammenhänge zwischen Vorwissen und planerischen und Überwachungsaktivitäten (Bernacki et al., 2012; Moos & Azevedo, 2008b). Vor allem die Überwachung des eigenen Verstehens als Facette der metakognitiven Überwachung ist positiv mit dem Ausmaß an Vorwissen verbunden (Moos & Azevedo, 2009). Ebenfalls ließ sich ein günstiges Zusammenwirken von Vorwissen und qualitativ hochwertigen Lernstrategien in einem Summenmaß eruieren: Wer ein hohes Vorwissen hat und zusätzlich dazu diverse Lernstrategien sinnvoll einzusetzen wusste, erwarb mehr Wissen (Greene, Costa et al., 2010).
- *Die Zusammenhänge von Vorwissen und kognitiven Strategien wirken weniger klar.* Während Vorwissen als individuelle Ressource für metakognitive Kontroll- und Überwachungsprozesse von Nutzen zu sein scheint, ist dies für aggregierte Daten bei den kognitiven Lernstrategien mitunter weniger deutlich der Fall, weil sich keine signifikanten Korrelationen haben ermitteln lassen (Bernacki et al., 2012). Teils gibt es sogar negative Zusammenhänge zwischen Vorwissen und dem Einsatz kognitiver Strategien: Wer mehr wusste, nutzte weniger kognitive Strategien während des Lernens (Moos & Azevedo, 2008b). Zudem ist ein positiver bzw. generell ein direkter Effekt des Vorwissens auf die Frequenz der Strategienutzung nicht durchgängig ermittelbar gewesen (Bernacki et al., 2012; Deekens et al., 2018; Dinsmore & Alexander, 2016; Greene, Bolick & Robertson, 2010; Greene, Costa et al., 2010; Greene et al., 2014, 2020).

Damit wirken die empirischen Befunde so, als stützten sie die Bedeutung des Vorwissens auf die Art, wie es das Modell von Winne (2001; Winne & Hadwin, 1998) bzw. Pintrich (2000) proklamieren: Metakognitive Maße scheinen eher positive Zusammenhänge mit der metakognitiv gesteuerten Ausführung von Lernstrategien aufzuweisen. Für die Häufigkeit der Anwendung kognitiver Strategien gelten solche (linearen) Zusammenhänge anscheinend nicht.

6.3.3.3 Ein personenzentrierter Blick auf differenzielle selbstregulatorische Vorgehensweisen von Personen mit hohem oder geringem Vorwissen

Studien, in denen verschiedene Gruppen dezidiert nach ihrem Vorwissen in Gruppen untergliedert und das selbstregulatorische Vorgehen in diesen Gruppen genauer betrachtet wurde, sind noch selten. Doch die bisherige Empirie lässt sich folgendermaßen verdichten: *Personengruppen mit unterschiedlichen Wissensbeständen neigen zu unterschiedlichen Vorgehensweisen.* Studien, die

teils deskriptiv vorgingen und teils auf statistische Signifikanz prüfend Personenunterschiede ermittelten, kommen zum Ergebnis, dass Personen ein partiell unterschiedliches Verhalten demonstrieren haben:

- Differenzen offenbaren sich darin, dass *Personen mit geringem Vorwissen* dazu neigten, sich den Gegenstand eher oberflächlich anzunähern (durch Notieren, Zusammenfassen, Memorieren) (Moos & Azevedo, 2008b; Trevors et al., 2014). Jedoch ergaben sich wenig Sequenzen von Strategien in ihrem Vorgehen, in denen korrekte metakognitive Einschätzungen oder mehrere kognitive Strategien in ihrem Verbund vorkamen (Taub & Azevedo, 2019).
- *Personen mit höherem Vorwissen* verknüpften hingegen die Informationen stärker beim Lesen über Inferenzen, suchten stärker gezielt nach Informationen und beurteilten den Inhalt stärker. Zusätzlich nutzen sie ihr Vorwissen als Ressource, indem sie es vor dem Bearbeiten der Aufgabe gezielt aktivierten (Moos & Azevedo, 2008b). Ferner zeigen sich Auffälligkeiten bei den metakognitiven Strategien, welche die Personen mit höherem Vorwissen teils in höherem Ausmaß demonstrieren (Taub et al., 2014). Konkret setzten sie sich mehr eigene Ziele beim Planen und überwachten ihren Verstehensfortschritt mehr (Winters & Azevedo, 2005). Differenzen ergaben sich zudem in den häufigsten Sequenzen von aufeinanderfolgenden Strategien im Verbund: Die Personen mit hohem Wissen nutzten besonders häufig ein Vorgehen, bei dem Planen, Vorwissensaktivierung, Verstehensüberwachung und dann ein Anfertigen von schriftlichen Zwischenprodukten anzutreffen war (Taub et al., 2014). Zusätzlich waren sie auffällig darin überlegen, Sequenzen mit einer Kombination von kognitiven und metakognitiven Strategien erfolgreich auszuführen (Taub & Azevedo, 2019).

Wie aus der Aufzählung oben hervorgeht, bildet das bessere Vorwissen eine Ressource für das Anwenden von Strategien, die eine tiefere Verarbeitung und eine Überwachung dieser Verarbeitungstiefe erlauben. Demgegenüber wirken Personen mit geringem Vorwissen bereits von der eher oberflächlichen Aneignung des jeweiligen Lerngegenstands stark gefordert. Das ist deshalb bemerkenswert, weil in den hier zusammengetragenen Studien Aufgaben gewählt wurden, in denen komplexe systemische Sachverhalte den Gegenstand des Lesens und Lernens bildeten.

6.3.3.4 Abschlussbemerkung

Das inhaltliche Vorwissen erhält als Faktor für das selbstregulatorische Geschehen beim Lernen in Theorie und Forschung mehr und mehr Zuwendung. Aus der Warte der Theorie besteht zunächst die Auffälligkeit, dass die Rolle des Vorwissens je nach Modell divergierend konzeptualisiert wird. Dazu gehört, dass das

Vorwissen in einigen Modellen keinerlei gesonderte Berücksichtigung erfahren hat, während andere Modelle wie jenes von Pintrich (2000) und dasjenige von Winne und Hadwin (1998) es systematisch als kognitive Bedingung einbinden. Auf den allgemeinsten Nenner gebracht wird in den beiden genannten Modellen das inhaltliche Vorwissen in einer Domäne als Ressource angesehen, welches zur Zielklärung genutzt wird und im späteren Verlauf als interner Vergleichsmaßstab für die metakognitive Überwachung und etwaige erneute Planungsprozesse fungiert. Deshalb ist in dieser Art der Vorwissensberücksichtigung in Modellen die Wirkung des Vorwissens als mittelbar anzunehmen.

Empirisch konvergieren zwei Forschungszugänge in ihren Befundmustern, welche freilich noch von sehr wenigen Studien unterfüttert sind. Variablenzentrierte Studien weisen auf engere Zusammenhänge zwischen metakognitiven Strategien und Vorwissen hin. Personenzentrierte Studien mit Vergleichen von Personen mit unterschiedlich gut ausgebautem Vorwissen ergeben ein komplementäres Bild: Personen mit höherem Vorwissen zeigen ein stärker metakognitiv gesteuertes Vorgehen, welches markantere Züge einer Tiefenaneignung des Inhaltes aufweist. Deshalb wirkt es im Gesamt so, als wäre mangelndes Vorwissen nicht nur ein Problem für tiefes Verstehen, sondern auch eine mangelnde Basis für ausreichend kalibrierte metakognitive Überwachungs- und Planungsprozesse selbstregulierten Lernens.

6.4 Probleme

Selbstregulation als dynamischer Prozessverbund verlangt den Personen viel ab. Daher ist es erwartbar, dass nicht jede Person bei jedem Anlass die optimale Vorgehensweise wählt, aufrechterhält und modifiziert. Im Bereich des (digitalen) selbstregulierten Lesens sind zuvorderst Fehlanwendungen von Strategien auf der Objektebene anzuführen, und derlei ist bereits ausgiebiger in jeweiligen Teilkapiteln zu den anderen vier Komponenten des digitalen Lesens Gegenstand gewesen ([2.3](#), [3.4](#), [4.4](#), [5.4](#)). Daneben sind aus metakognitiver Warte eigene Probleme der Regulation eigener Kognitionen und Verhaltensweisen ein ganz eigener Bereich, der einer Vertiefung bedarf. Diese Vertiefung erfolgt nun in Bezug auf drei Facetten der Problemlagen: Erstens bestehen Probleme darin, die eigene Aufmerksamkeit auf die verstehensnötigen Kognitionen zu lenken ([6.4.1](#)), zweitens in der Fähigkeit, das eigene Verstehen präzise kalibriert metakognitiv einzuschätzen ([6.4.2](#)), und drittens manifestieren sich in Studien regelmäßig Defizite, geeignete, auf Tiefendurchdringung abzielende und eine solche Vorgehensweise überwachende Strategien spontan im Verbund zu bemühen ([6.4.3](#)).

6.4.1 Problemfeld 1: missglückende Aufmerksamkeitslenkung (Probleme der Kontrolle)

Das erste hier betrachtete Problemfeld behandelt zwei Facetten, wie Personen es schwerfällt, ihre Aufmerksamkeit so zu lenken, dass die für die Leseaufgabe und das Leseziel wichtigen Kognitionen konsequent aufrechterhalten werden. Beim Multitasking bilden hierfür eher externe Reize die Grundlage einer ablenkenden Störung (6.4.1.1), bei den abschweifenden Gedanken eher interne Gedanken (6.4.1.2).

6.4.1.1 Negative Effekte des Multitaskings auf kognitive Maße

Multitasking ist ein Begriff, der dafür verwendet wird, wenn eine Person mehr als eine Aufgabe ausführt, die ihre Aufmerksamkeit benötigt und dabei die zu bewältigenden Aufgaben um die kognitiven Ressourcen konkurrieren (Lang & Chrzan, 2015). Einigen Autoren zufolge lässt sich das Konstrukt Multitasking aber auch noch deutlich engführen, nämlich als die Ablenkung von einer eigentlich zu priorisierenden Aufgabe wie dem Lesen eines Textes (Aagaard, 2019). Es entstehen dadurch zwei Kosten, nämlich indem man entweder Zeit verliert (Verdrängungshypothese) oder/und die lernrelevanten Informationen nicht ausreichend verarbeitet, weil die limitierten kognitiven Ressourcen des Arbeitsgedächtnisses nicht vollumfänglich für das Lernen/Verstehen genutzt werden (Hypothese der begrenzten Informationsverarbeitung; van der Schuur et al., 2015). Für die zweite Hypothese sprechen metaanalytische Befunde, nach denen Personen mit Hang zu viel multimedialem Multitasking schlechtere Ergebnisse in zwei Bereichen der exekutiven Funktionen haben als Personen mit geringem Multitasking (Kong et al., 2023): Erstens zeichnen sich Tendenzen ab, dass die Arbeitsgedächtniskapazität geringer ist, und zweitens gelingt es den Personen weniger gut, ihre inhibitorische Kontrolle zu nutzen. Anders gesagt: Die wenigen Inhalte, die Personen kurzzeitig verarbeiten werden, fallen eher anderen Reizen bei der Aufmerksamkeitssteuerung zum Opfer.

Multitasking als potenzielle Gefährdung von Leistungen, die Aufmerksamkeit benötigen, wurde in so vielen einzelnen Studien untersucht (May & Elder, 2018), dass inzwischen eine Vielzahl von Metaanalysen vorgelegt wurden, die teilweise spezifisch für das Lesen durchgeführt wurden, meist jedoch einen weiter gefassten Bereich abhängiger Variablen fokussierten. Die bislang vorgelegten Metaanalysen unterscheiden sich darin, ob sie entweder diverse kognitive Leistungsmaße kombiniert (Jeong & Hwang, 2016; Wiradhany & Nieuwenstein, 2017) oder einzelne kognitive Leistungen genauer betrachten (Clinton-Lisell, 2021; Jeong & Hwang, 2016; Kämpfe et al., 2011; Parry & Le Roux, 2021; Vasilev et al., 2018). Dasselbe gilt für die moderatoranalytische Frage danach, ob spezifische Aufgaben bzw. Merkmale des Multitaskings eine stärker störende Wirkung haben (spezifische

Effekte untersuchen z. B. Clinton-Lisell, 2021; Jeong & Hwang, 2016; Kämpfe et al., 2011; Parry & Le Roux, 2021; Vasilev et al., 2018; Wiradhany & Koerts, 2021; Wiradhany & Nieuwenstein, 2017). Die metaanalytische Essenz lautet in Bezug auf kognitive Maße, zunächst ausgehend von den allgemeinen Befunden hin zu spezifischeren Ergebnissen und Resultaten aus Moderatoranalysen:

- *Negative Effekte betreffen diverse kognitive Maße:* Zu diesem Ergebnis gelangen die Metaanalysen unisono in Bezug auf die jeweils in ihnen fokussierten kognitiven Indikatoren, die Effekte selbst sind in ihrer Ausprägung als schwach bis mittel einzustufen. Negative Effekte bestehen speziell bei Maßen zum Aufrechterhalten von Aufmerksamkeit (Parry & Le Roux, 2021; Wiradhany & Koerts, 2021), dem Verfügbarhalten relevanter Informationen im Arbeitsgedächtnis (Parry & Le Roux, 2021; Wiradhany & Koerts, 2021), dem Unterdrücken von störenden Impulsen (Parry & Le Roux, 2021; Wiradhany & Koerts, 2021), metakognitiv gesteuerten Facetten der Selbstregulation (Wiradhany & Koerts, 2021) und beim Leseverstehen (Clinton-Lisell, 2021; Vasilev et al., 2018). Hinzu kommen die generischen Effekte in aggregiert ausgewerteten kognitiven Leistungstests verschiedenster Couleur (Jeong & Hwang, 2016; Wiradhany & Nieuwenstein, 2017). Hinsichtlich der Effekte einzelner Arten von Störungen lassen sich in Moderatoranalysen zum Teil nicht immer eindeutige Schlüsse ziehen, welche Merkmale der Ablenkung sich stärker oder schwächer auswirken (Jeong & Hwang, 2016; Parry & Le Roux, 2021; Wiradhany & Nieuwenstein, 2017).
- *Wenn man liest und im Hintergrund akustische (sprachliche) Ablenkung besteht, wirkt sich das negativ auf Leseleistungen aus:* Nicht direkt dem Multitasking im strengen Sinne zugehörend, aber doch weiter gefasst als Ablenkung ist auch der negative Effekt von Hintergrundgeräuschen, unverständlicher Sprache und Musik auf das Leseverstehen und teils und in geringerem Ausmaß auf die Lesegeschwindigkeit belegt (Kämpfe et al., 2011; Vasilev et al., 2018). Dieser störende Effekt eines akustischen Hintergrunds bestand altersunabhängig und unterschied sich nach Art der akustischen Stimuli: Insbesondere wenn verständliche Sprache oder Musik mit Text im Hintergrund zu vernehmen waren, bestand eine erhöhte Wahrscheinlichkeit eines störenden Effekts im Sinne eines geringeren Leseverstehens (Vasilev et al., 2018). Allerdings gibt es auch unklare Muster, ob sich verschiedene Arten von Hintergrundmusik nicht doch positiv bemerkbar machen (Kämpfe et al., 2011). In diesem letztgenannten Fall lassen sich die Primärstudien jedoch kaum auf sinnvolle Weise miteinander vergleichen, sodass diese Frage notwendigerweise derzeit offenbleibt.
- *Negative Effekte des Multitaskings existieren vor allem dann, wenn die Personen das Tempo bei der Aufgabenbearbeitung nicht selbst kontrollieren können.* In den Metaanalysen fallen negative Effekte dann größer aus, wenn die

Zeit entweder vorgegeben wurde (Clinton-Lisell, 2021) oder die Testpersonen keine oder nur wenig Möglichkeiten hatten, Zwischenstopps einzulegen (Jeong & Hwang, 2016). Externale zeitliche Begrenzungen verschärfen demnach die problematischen Effekte des Multitaskings.

- *Wenn die Aufgaben des Multitaskings Distanz implizieren, ist der negative Effekt des Multitaskings höher.* Hatten die beiden Aufgaben ein gleiches Ziel oder ähnliche Ziele, wiesen also eine gewisse Komplementarität auf (z. B. das Zuhören bei einer Vorlesung und eine Internetrecherche), so waren die negativen Effekte nicht so stark wie in dem Fall, in dem die beiden Aufgaben *kein gemeinsames Ziel* hatten (also etwa Musik hören und Hausaufgaben erledigen). Dennoch waren diese Effekte in beiden Fällen ungünstig (Jeong & Hwang, 2016), es geht also nur um einen relativ geringeren Nachteil bei komplementären Aufgaben. Ebenfalls erwies es sich als ungünstiger, wenn die beiden Aufgaben *in unterschiedlichen Medien* stattfanden, also etwa nicht beide am Computer, sondern eine computerbasierte und eine mit analogen Medien (Jeong & Hwang, 2016).

Diese Ergebnisse zu empirisch beobachtbaren Minderleistungen im kognitiven Bereich, spezifisch jene aus den letzten drei Spiegelstrichen der Aufzählung, reißen sich ein in Prinzipien der Gestaltung günstiger Multimedia-Lernumgebungen (Fiorella & Mayer, 2022; Noetel et al., 2022). Diese Prinzipien greifen nämlich gezielt auf, dass es Merkmale von externen Umgebungen gibt, die hinsichtlich vermeidbarer Überlastungen des Arbeitsgedächtnisses das Lernen erschweren bzw. begünstigen. Hier ist Multitasking aber in aller Regel nicht als Folge von zu vielen zeitgleich zu verarbeitenden, samt und sonders verstehensnötigen Informationen anzusehen (sog. „essenzielle Überlastung“; Mayer & Fiorella, 2022). Stattdessen lässt sich Multitasking als unnötige Erschwernis einer kognitiven Aufgabe werten, bei der die sekundären Aufgaben mit der primären Aufgabe um Aufmerksamkeit im Arbeitsgedächtnis konkurrieren (sog. „zusätzliche, sachfremde Belastung“).

Die in diesem Teilkapitel berichtete Liste von negativen Effekten des Multitaskings vermengt leseunspezifische und -spezifische Effekte. Um abschließend die Lesespezifik nochmals herauszuarbeiten, enthält der Kasten „Illustrierendes Beispiel 30“ eine Vertiefung. Die darin berichteten Ergebnisse aus einer aktuellen Metaanalyse betreffen Effekte sowohl auf die Prozesse als auch die Produkte.

Illustrierendes Beispiel 30

Schlechteres Leseverstehen bei höherem Zeitaufwand – der Preis des Multitaskings beim Lesen

A) Was war das Erkenntnisinteresse?

Die Metaanalyse (Clinton-Lisell, 2021) widmete sich den einzelstudienübergreifenden Befunden zu Effekten experimenteller Studien, in denen Personen Texte lasen (in der jeweiligen Kontrollgruppe) oder Texte lasen und eine weitere Sekundäraufgabe im Sinne des Multitaskings bearbeiteten (in der jeweiligen Experimentalgruppe). Bei den Sekundäraufgaben handelte es sich um das Anschauen von Videos, Anhören von Tonaufnahmen, Beantworten von Kurznachrichten etc. Geprüft wurden Effekte dieses Multitaskings auf zwei Variablen: Leseverstehen und Lesetempo.

B) Wer wurde untersucht?

In die Studie flossen 22 Studien ein (20 lieferten Daten für das Leseverstehen, 9 für das Lesetempo). Diese Veröffentlichungen stammten aus den Jahren 2000–2020 und umfassten 1.686 Personen (davon 1.397 Erwachsene, mehrheitlich Studierende).

C) Welche Instrumente kamen zum Einsatz?

Gemessen wurde in den Primärstudien das Leseverstehen mittels diverser Tests (Erinnerungen, offene Fragen, Multiple-Choice-Aufgaben – mehrheitlich bei Sachtexten getestet, nämlich in 90 Prozent der Studien). Die Lesezeit wurde über die Dauer erfasst, ausgewertet wurde diese aber nur für jene Studien, in denen die Testpersonen kein Zeitlimit bei der Aufgabenbearbeitung hatten. Verglichen wurden jeweils die Differenzen zwischen Experimental- und Kontrollgruppe, indem eine gewichtete Effektstärke (g) berechnet wurde. Daneben wurden für die Metaanalyse noch mehrere potenzielle Moderatorvariablen kodiert, von denen sich annehmen ließ, dass sie einen Effekt auf die Ausprägung der Leistungsdifferenzen der Gruppen haben könnten.

D) Was waren die Hauptbefunde?

Die Personen aus der Experimentalgruppe mit Multitasking-Auftrag hatten ein schlechteres Leseverstehen ($g = -.28$; entspricht umgerechnet einer Differenz von 28 Punkten bei PISA). Dieser Nachteil wurde von zwei Moderatorvariablen beeinflusst. Gab es ein Zeitlimit, war der Nachteil ungefähr viermal so stark ($g = -.54$) wie beim Bearbeiten ohne zeitliche Vorgabe ($g = -.14$). Auch das Lesemedium spielte eine Rolle: Personen, die analog lasen, schnitten etwas schlechter ab ($g = -.36$) als Personen, die am Bildschirm Leseaufgaben lösten ($g = -.23$). Bei der Lesezeit ergaben sich im Set der Experimente ohne Zeitvorgaben ebenfalls statistisch signifikante Effekte: Wer Multitasking-Aufgaben absolvierte, brauchte mehr Zeit für die Leseaufgaben ($g = .52$).

E) Diskussion

Multitasking hat seinen Preis beim Lesen: Personen brauchen mehr Zeit und lösen Leseverstehensaufgaben schlechter. Die schlechteren Leistungstests im Leseverstehen fallen weniger durch das Lesemedium auf (hier allerdings zuungunsten des Lesens gedruckter Texte) als durch den Zeitdruck: Personen mit einer vorgegebenen Bearbeitungsdauer lösen Aufgaben erheblich schlechter als jene ohne zeitliche Limitierung. Es ist mithin auffällig, dass diese Ergebnisse zu den negativen Effekten des Multitasking im Gesamt den Faktor Zeit betreffen: Wo es keine zeitliche Begrenzung gab, fielen die Differenzen kaum ins Gewicht (gleichwohl waren die Leistungen immer noch etwas schlechter), führten aber zu einer substantiellen Verlängerung der Bearbeitungsdauer. War diese von vornherein begrenzt, schlug sich Multitasking in markant schlechterem Leseverstehen nieder. Den Preis des Multitaskings bildet also je nach Perspektive entweder schlechtere Leseleistungen oder eine Ausweitung der Bearbeitungsdauer.

6.4.1.2 Negative Effekte abschweifender Gedanken (Mind Wandering)

Abschweifende resp. nicht-aufgabenbezogene Gedanken (Englisch: „mind wandering“ bzw. „task-unrelated thoughts“) bezeichnen ein Phänomen, bei dem Personen während einer kognitiven Aktivität die inneren Kognitionen von dieser Aufgabe entkoppeln und es nicht glückt, diese Gedanken ausreichend wieder auf das eigentliche Primärziel auszurichten (Smallwood, 2011, 2013; Smallwood & Schooler, 2015). Dafür gibt es verschiedene Erklärungsansätze, z. B. dass die Inhalte der abschweifenden Gedanken eine höheren Anreiz bzw. eine höhere Salienz aufweisen (Hypothese der gegenwärtigen Überlegungen), dass eine mangelnde Kontrolle der exekutiven Aufmerksamkeitslenkung vorliegt (Hypothese der versagenden Aufmerksamkeit) oder dass es Personen nicht glückt, in ihrem Metabewusstsein zu erkennen, dass ihre Gedanken abgeschweift sind (Metabewusstseinshypothese; Smallwood, 2013).

Was macht abschweifende Gedanken in ihrem Kern aus? Da das Konstrukt „abschweifender Gedanken“ gegenwärtig eine gewisse Vorläufigkeit und durchaus unterschiedliche Konzeptualisierungen aufweist (s. z. B. McVay & Kane, 2010; Smallwood & Schooler, 2006, 2015), ist ein eher eingrenzendes Vorgehen als Antwort auf die Frage sinnvoll. In ihrem Forschungsüberblick zum Gedankenabschweifen schlagen D’Mello und Mills (2021) in diesem Sinne eine lesespezifische Arbeitsdefinition des Konstrukts vor, welche die aktuellen Unschärfen und das sich dynamisch entwickelnde Verständnis des Konstrukts aufnimmt. Abschweifende Gedanken beim Lesen müssen demnach *drei Kriterien* erfüllen: Die Gedanken *stimmen* erstens *nicht mit dem Ziel des Leseverstehens überein*. Das Abschweifen beim Lesen durch verstehensinkongruente Gedanken erfolgt zweitens *ungewollt*. Und drittens ist das Abschweifen ein insgesamt *dynamischer Prozess*. Im Falle, dass es der lesenden Person missglückt, metakognitiv ihre

Aufmerksamkeit wieder auf das Lesen zu richten, sind unvollständige kognitive Repräsentationen des Gelesenen zu erwarten.

Ob abschweifende Gedanken alleinig Ausdruck mangelnder exekutiver Kontrolle sind, wird wissenschaftlich diskutiert (McVay & Kane, 2010; Schooler et al., 2011), und auch die ausschließlich negativen Konsequenzen abschweifender Gedanken sind uneinheitlich und betreffen anscheinend nur einen Teil von Leistungen und Merkmalen (Mooneyham & Schooler, 2013). Dennoch scheinen abschweifende Gedanken vor allem ein empirisch ermittelbares Problem für anspruchsvolle kognitive Leistungen darzustellen. Mehrere Metaanalysen haben sich mit den Auswirkungen abschweifender Gedanken auf kognitive Leistungen befasst, aber auch damit, welche kognitiven bzw. soziodemografischen Faktoren in einem Zusammenhang zum Ausmaß der abschweifenden Gedanken stehen. Drei Ergebnislinien lassen sich erkennen:

- *Abschweifende Gedanken tauchen häufig auf, vor allem bei jüngeren (erwachsenen) Personen.* Abschweifende Gedanken werden auf unterschiedlichste Weise gemessen (Randall et al., 2014; Smallwood & Schooler, 2006; Weinstein, 2018). Eine besonders oft zum Einsatz gelangende Variante besteht darin, parallel zur Aufgabe auf Signale hin zu erfassen, ob die Person abschweifende Gedanken hat und ggf. worauf sie sich beziehen (Weinstein, 2018). D'Mello (2018) ermittelte in seiner Metaanalyse über 25 Studien hinweg, dass in diesen Experimenten in 28 Prozent der zeitlich wiederkehrenden Aufforderungen die Probandinnen und Probanden per Selbstauskunft abschweifende Gedanken äußerten. Wong et al. (2022) sowie Kawashima et al. (2023) stellten metaanalytisch mit 31 bzw. 35 Prozent einen sehr ähnlichen Wert fest; jüngst ermittelten Zanesco et al. (in Druck) in ihrer Metaanalyse sogar 40 Prozent, wobei sie ebenfalls feststellten, dass Personen zu Beginn von Aufgaben eine deutlich geringere Auftretenshäufigkeit abschweifender Gedanken haben – im Lauf von Aufgabenbearbeitungen mehrten sich also die abschweifenden Gedanken. Anders und die Metaanalysenbefunde zusammenfassend gesagt: In Minimum mehr als einem Viertel der erfragten Zeitpunkte waren die erwachsenen Testpersonen gedanklich mit etwas anderem befasst als mit dem Lösen der gestellten (Lese-)Aufgabe. Dabei – das deutet sich in einer weiteren Synopse von Forschungsergebnissen an – spielt es eine Rolle, wie alt die Testpersonen sind. Dem Effekt des Alters auf die Häufigkeit abschweifender Gedanken ging eine aktuelle Metaanalyse mit Erwachsenen nach (Jordão et al., 2019). Laut dieser Metaanalyse wiesen ältere Testpersonen deutlich geringere Auftretenshäufigkeiten abschweifender Gedanken auf. Demnach sind jüngere Personen weniger dazu in der Lage, ihre Aufmerksamkeit zu kontrollieren. Auf ein ähnliches, wenngleich statistisch nicht signifikantes Ergebnis kommen zudem Wong et al. (2022): Ältere Erwachsene (21 Prozent Auftretenshäufigkeit) haben im Vergleich zu jüngeren

Erwachsenen (33 Prozent) seltener abschweifende Gedanken. Zusammenge-
nommen weisen metaanalytische Ergebnisse darauf hin, dass abschweifende
Gedanken eine häufigere Erscheinung unter jüngeren Erwachsenen sind.

- *Abschweifende Gedanken korrespondieren mit geringeren Leistungen, insbesondere wenn die Aufgaben komplex sind.* Metaanalysen berichten über weit gefasste kognitive Leistungen negative Zusammenhänge zwischen Häufigkeit abschweifender Gedanken und Leistungen (D’Mello, 2018; Randall et al., 2014; Wong et al., 2022), auch und gerade beim Lesen (Bonifacci et al., 2023; D’Mello & Mills, 2021). Ein wichtiges Ergebnis einer Moderatoranalyse in der Metaanalyse von Randall et al. (2014) betraf die Aufgabenkomplexität. Man interessierte sich dafür, ob Leistungen in komplexen Aufgaben stärker darunter litten, wenn sie von abschweifenden Gedanken flankiert wurden. Als komplex galten Aufgaben dann, wenn a) sie eine hohe Anzahl an Elementen aufwiesen, b) diese Elemente miteinander zu koordinieren waren bzw. c) wenn sie in ihrem Ablauf dynamisch waren. Waren zwei der drei Kriterien erfüllt, und darunter fiel auch das Leseverstehen, dann wurden die Primärstudien dieser Gruppe von Aufgaben zugeteilt. Das Ergebnis der Moderatoranalyse: Tatsächlich war der negative Zusammenhang in komplexen Aufgaben erheblich stärker ausgeprägt als in kognitiv weniger anspruchsvollen Aufgaben. Diese Ergebnisse illustrieren mithin recht konsistent, dass abschweifende Gedanken für das Leseverstehen problematisch sind. Zu einem vergleichbaren Resultat gelangten Wong et al. (2022): Das Lesen war besonders deutlich von einem negativen Zusammenhang von der Häufigkeit abschweifender Gedanken und Leistungen betroffen.
- *Je stärker ausgeprägt die individuellen kognitiven Ressourcen sind, desto seltener treten abschweifende Gedanken auf.* Nicht nur ein höheres Alter scheint ein protektiver Faktor zu sein (s. o. die Ergebnisse von Bonifacci et al., 2023, und Jordão et al., 2019), kognitive Ressourcen sind es ebenfalls. Das geht aus den Ergebnissen der metaanalytischen Korrelationen von Randall et al. (2014) hervor. Sie fassten Maße der allgemeinen Intelligenz und der Arbeitsgedächtniskapazität zusammen und ermittelten, dass diese beiden Indikatoren kognitiver Ressourcen negativ mit dem Vorkommen von abschweifenden Gedanken zusammenhängen, dafür aber positiv mit aufgabenbezogenen Gedanken und erbrachten Leistungen. Die gefundenen Korrelationen sind als schwach bis mittel einzustufen, doch ergeben sie in sich ein kohärentes Bild: Personen mit mehr kognitiven Ressourcen sind leistungsbezogener in ihren Prozessen (weniger abschweifende, mehr aufgabenfokussierte Kognitionen) und leistungsstärker in ihren Produkten. Natürlich ist damit nicht entwirrt, wie die kausalen Zusammenhänge sich gestalten, doch auch dafür mehren sich Befunde, z. B. die Experimentalstudie, die im Kasten „Illustrierendes Beispiel 31“ beschrieben wird.

Illustrierendes Beispiel 31

Ist es das Arbeitsgedächtnis, das für abschweifende Gedanken zuständig ist, oder ist es die Fähigkeit, die Aufmerksamkeit zu kontrollieren?

A) Was war das Erkenntnisinteresse?

Die Studie von McVay und Kane (2012) wollte eruieren, ob Arbeitsgedächtnis oder Aufmerksamkeitskontrolle für negative Effekte auf das Leseverstehen verantwortlich sind und ob dies über die Häufigkeit der abschweifenden Gedanken erklärbar ist.

B) Wer wurde untersucht?

Die Testpersonen waren 253 Studierende im Alter zwischen 18 und 35 Jahren aus den USA. Die Personen sprachen Englisch als Muttersprache, zu weiteren soziodemografischen Merkmalen gab es keine Informationen.

C) Welche Instrumente kamen zum Einsatz?

Die Studie testete umfassend kognitive Fähigkeiten aus vier Bereichen. Die a) Arbeitsgedächtnisspanne wurde mit drei separaten Tests gemessen, die b) Aufmerksamkeitskontrolle wurde ebenfalls mit drei verschiedenen Tests geprüft. Das c) Leseverstehen erfasste man über das Lesen von vier Sach- und drei literarischen Texten und dazu beantworteten Fragen. Als weiteres Maß für das Leseverstehen wurde das Ergebnis eines Studierfähigkeitstests herbeigezogen. Die d) abschweifenden Gedanken wurden bei zwei längeren Leseaufgaben und zwei Aufgaben bei der Aufmerksamkeitskontrolle erfasst, stammten also aus vier Testgelegenheiten und gaben die Häufigkeit der abschweifenden Gedanken an. Diese Testergebnisse wertete das Forschungsteam mittels Strukturgleichungsanalysen aus, weil es sich insbesondere für die Mediation, also die vermittelnden Effekte einer dritten Variable, interessierte.

D) Was waren die Hauptbefunde?

Das Hauptergebnis war, dass man Arbeitsgedächtnisspanne und Aufmerksamkeitskontrolle aufgrund der hohen Zusammenhänge beider Faktoren bündelte und dann prüfte, ob dieser gemeinsame neue Faktor (exekutive Aufmerksamkeit) positiv das über sieben Tests erfasste Leseverstehen vorhersagen konnte. Hierbei wurde ein Mediationseffekt ermittelt: Je besser die Testergebnisse in der exekutiven Aufmerksamkeit waren, desto seltener traten die abschweifenden Gedanken auf. Und je weniger diese aufgetreten waren, desto besser waren die Leseleistungen.

Ein anderes Ergebnis bestand darin, dass man geprüft hatte, ob die Arbeitsgedächtnisspanne und die Aufmerksamkeitskontrolle als separate Konstrukte direkte oder indirekte Effekte auf das Leseverstehen hatten. Die Ergebnisse sind in Abbildung 24 dargestellt. Aus ihr geht hervor, dass Aufmerksamkeitskontrolle und Arbeitsgedächtnisspanne hoch korrelieren ($r = .73$), dass aber nur die Aufmerksamkeitskontrolle negativ die Auftretenshäufigkeit vorhersagt ($\beta = -.55$), die ihrerseits negativ das Leseverstehen prädiziert ($\beta = -.44$).

Es ist hier also ein vollständiger Mediationseffekt nachgewiesen worden. Demnach schützt eine erhöhte Aufmerksamkeitskontrolle vor Gedanken, die abschweifen und ein schlechteres Leseverstehen nach sich ziehen.

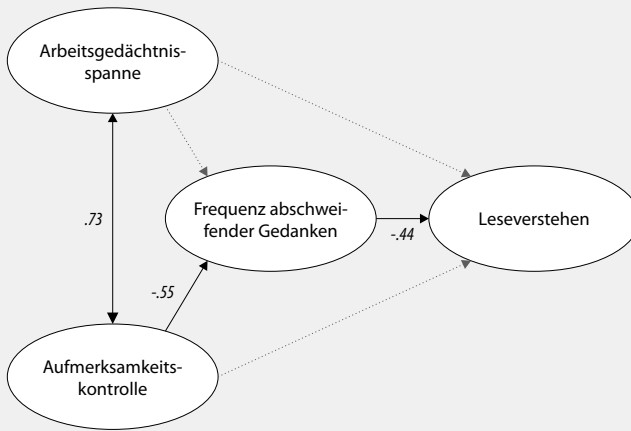


Abbildung 24: Zusammenhänge von Arbeitsgedächtnisspanne, Aufmerksamkeitskontrolle, abschweifenden Gedanken und Leseverstehen im Rahmen einer strukturanalytischen Mediationsanalyse (Quelle der vereinfachten Darstellung: McVay & Kane, 2012, S. 313, gestrichelte graue Linien geben nicht-signifikante Pfade an)

E) Diskussion

Die Studie ist eine der wenig bislang existierenden, die ihre Konstrukte über viele einzelne Messungen abstützten und sich für die Frage nach komplexeren Wirkgefügen interessierten. Besonders aufschlussreich ist die Mediationsanalyse, nach der eine mangelnde Aufmerksamkeitskontrolle bedeutend für abschweifende Gedanken ist, welche dann das Leseverstehen allgemein unterminieren. Im Licht dieser Befunde sind häufige abschweifende Gedanken Folge einer ungenügenden Kontrolle der Aufmerksamkeit.

6.4.2 Problemfeld 2: ungenaues Überwachen der eigenen Verstehensleistungen (Probleme des Monitorings)

Selbstregulation hat zum Ziel, die Diskrepanz zwischen einem Ist-Zustand und einem erwünschten Soll-Zustand zu minimieren, wozu internal ablaufende, mehrfach zu absolvierende Feedback-Schleifen dienen. Deshalb ist die Frage bedeutsam, woran das Erreichen bzw. Nicht-Erreichen eines Zielzustandes festgemacht wird und wie präzise dieses Urteil ist. Die metakognitiv orientierte Forschung hat sich darum mit einer Facette des Monitorings befasst, die in der Literatur als „Metaverstehen“ (Dunlosky & Lipko, 2007), lernbezogene „Urteile“ bzw. „Einschätzungen“ (Nelson & Narens, 1990; Tarricone, 2011), „Akkuratheit des Verstehens“ (Prinz et al., 2020a) oder „Kalibrierung“ (Bol & Hacker, 2012;

Hacker & Bol, 2019) bezeichnet werden. Trotz divergierender Nomenklatur eint diese Konstrukte inhaltlich, dass die metakognitive Überwachung der eigenen Leistung im Zentrum steht, z. B. beim Lesen und dort zum Beispiel zum angestrebten Kohärenzstandard (s. Teilkap. 2.1.2). Es handelt sich mithin um eine Form der Urteilsgenauigkeit (Griffin et al., 2019). Es müssen hierfür zwei Informationen miteinander metakognitiv verglichen werden: eine zur eigenen Leistung (Ist) und eine zu einem definierten Ziel (Soll). Diese Information benötigen Lernende dazu, bei der Selbstregulation im Falle einer Diskrepanz zwischen Soll und Ist ihr Vorgehen gesamthaft im Sinne metakognitiver Kontrollprozesse zu modifizieren, z. B. indem sie Texte erneut lesen, die sie nicht ausreichend verstanden haben (Butler & Winne, 1995; de Bruin & van Gog, 2012; Griffin et al., 2013; Metcalfe, 2009; Panadero et al., 2017; Thiede & de Bruin, 2018). Um dies zu tun, müssen die Urteile angemessen realistisch sein, und hierin liegt die Ursache, warum dieses Thema in diesem Buch als ein Problemfeld dargestellt wird. Denn: Die eingeschätzten Leistungen und die tatsächlich erbrachten Leistungen sind oftmals wenig realistisch vonseiten der Lernenden beurteilt worden.

Für diese Diskrepanz werden in Studien von Personen zwei Arten von Daten verwendet. Die eine betrifft Leistungsdaten, welche über Tests wie Leseverstehensleistungen erhoben werden, z. B. über die Anzahl korrekt gelöster Aufgaben, in aller Regel mithilfe mehrerer Texte, um zu validen Ergebnissen zu gelangen. Die andere Art von Daten sind vorgängige Urteile über die noch zu erbringende Performanz oder nachträgliche Einschätzungen zur erbrachten Leistung (etwa: die vermutete Anzahl richtig beantworteter Fragen). Aus diesen Daten lassen sich absolute oder relative Maße ermitteln, welche für diesen Forschungszweig wichtig sind und unterschiedliche Facetten des Konstrukts der Urteilsgenauigkeit fokussieren (Dunlosky & Lipko, 2007; Schraw, 2009):

- *Absolute Maße* erfassen, ob sich die Personen in ihrer Einschätzung über- oder unterschätzen oder eine realistisch kalibrierte Einschätzung äußern; dafür wird als Indikator ein Differenzmaß bemüht. Es geht hier also darum, ob jemand seine eigene Leistung gemessen an der faktisch erbrachten Leistung korrekt einschätzt.
- *Relative Maße* nutzen Intrapersonen-Korrelationen, um den Zusammenhang zwischen Urteilsgenauigkeit und Leistungen abzubilden.

Diverse Metaanalysen (Dunlosky & Lipko, 2007; Maki, 1998; Prinz et al., 2020a; Thiede et al., 2009; Yang et al., 2023) weisen in ihren Resultaten zur relativen Urteilsgenauigkeit seit mehr als zwei Dekaden darauf hin, dass Selbsteinschätzungen zum Leseverstehen bei Aufgaben und tatsächlich erbrachter Leistungen nur in geringem Maß miteinander statistisch zusammenhängen, was als Symptom für eine geringe Urteilsgenauigkeit gilt. Diese Metaanalysen erbringen in den gefundenen Korrelationskoeffizienten konsistent recht ähnliche Ergebnisse

($r = .18-.27$). Solche Korrelationen belegen Zusammenhänge, doch sind sie nicht sonderlich stark. Allerdings ändert sich das Bild dann, wenn nicht vorgängig Urteile zu bilden sind, sondern *nach* dem Lösen (Rhodes & Tauber, 2011). Maki (1998) berichtet hier Korrelationskoeffizienten in nahezu doppelt hoher Ausprägung ($r = .48$). Dies deutet auf erheblich genauere relative Urteile hin, wobei diese Befunde nicht durchgängig berichtet worden sind (Prinz et al., 2020a).

Unter den zuvor genannten Metaanalysen hat jene von Prinz et al. (2020a) vertiefende Moderatoranalysen vorgenommen, die sich auf verschiedene Merkmale beziehen, von denen für die Zwecke dieses Buchs jene zu Personen und zu Merkmalen der Texte wichtig sind.

- *Es gibt Altersunterschiede bzw. zwischen den Bildungsetappen – die am besten kalibrierten Urteile stammen von Studierenden.* Demgegenüber sind schulpflichtige Personen wie Primarschulkinder und Sekundarschuljugendliche wenig dazu in der Lage, ihre relative Leistung sinnvoll einschätzen. Die gefundenen Korrelationen sind in diesen Altersgruppen nicht nur gering, sondern teils nicht einmal zufallskritisch abzusichern.
- *Je mehr Texte die Personen gelesen haben, desto unsicherer werden sie in ihrer Beurteilung.* Es wirkt so, als würden Personen immer diffuser in ihrer Selbsteinschätzung, wenn sie bei verschiedenen Texten und damit in verschiedenen Anforderungssituationen ein kalibriertes Urteil abgeben sollen.
- *Längere Texte erbrachten höhere Urteilsgenauigkeit.* Das bedeutet: Die Urteilssicherheit wurde umso präziser, je mehr Wörter die Texte umfassten.
- *Lesen am Bildschirm geht mit einer etwas besseren Urteilsgenauigkeit einher.* Dieser Befund war in der Metaanalyse unerwartet, doch zeigten sich leicht höhere Korrelationen zwischen Leistung und metakognitivem Urteil, wenn das Lesen am Bildschirm erfolgte. Unerwartet ist dieser Befund auch, wenn man ihn mit jenem aus Clintons (2019) metaanalytischem Ergebnis kontrastiert, das auf absoluten Urteilsmaßen basierte: Hier neigten die Personen, die digital präsentierte Texte lasen, zu einer Überschätzung ihrer Leistung (s. dazu den Kasten „Illustrierendes Beispiel 32“). Es gab also durchaus einen negativen Effekt, dessen Gegenläufigkeit sich auch darüber erklären lässt, dass verschiedene Maße der Urteilsgenauigkeit zu sehr unterschiedlichen Indikatoren führen, die nicht einmal notwendigerweise empirisch zusammenhängen (Schraw, 2009).

Illustrierendes Beispiel 32

Mehr Schein als Sein? Die absolute Urteilsgenauigkeit von Studierenden und ihr Zusammenhang mit faktischen Leistungen beim Lesen digitaler und analoger Texte

A) Was war das Erkenntnisinteresse?

Im Rahmen eines Experiments interessierte sich ein Forschungsteams dafür, ob sich Monitoringprozesse beim Lernen mit Texten dann unterscheiden, wenn diese Texte entweder analog oder digital gelesen werden (Ackerman & Goldsmith, 2011). Die folgende Darstellung bezieht sich allein auf das zweite Experiment in der Publikation.

B) Wer wurde untersucht?

An dem Experiment nahmen 74 israelische Studierende teil, davon zwei Drittel Frauen. Die Studierenden waren im Durchschnitt 25 Jahre alt.

C) Welche Instrumente kamen zum Einsatz?

Die Studierenden wurden unterschiedlichen Experimentalgruppen zugeteilt, nämlich grundsätzlich einer Gruppe von Personen, die dieselben Texte digital als Word-Dokument lasen, und einer zweiten Gruppe, die die Texte analog zur Verfügung gestellt bekommen hatten. Innerhalb beider Gruppen gab es noch eine ca. hälftige Teilung in zwei Gruppen, nämlich dahingehend, ob die Personen entweder alle drei Minuten beim Lesen eine Einschätzung zu ihrer erwarteten Leistung und nach dem Ende der Lesephase abgaben (Gruppe: Parallelurteile) oder nur ein einziges Urteil am Ende der Lesephase (Gruppe: Schlussurteile). Hier stellten sich aber keine Differenzen heraus, sodass die Ergebnisse nicht weiter ausdifferenziert werden, sondern nur für einen Teil der Gruppen, nämlich den beiden Gruppen Parallelurteil in den Konditionen digitales und analoges Lesen.

Für jede Gruppe wurden verschiedene Maße erfasst: a) die Bearbeitungsdauer pro Text, b) Urteile zum Erfolg bei bevorstehenden Leseverstehensaufgaben und c) die Ergebnisse aus den Multiple-Choice-Leseaufgaben in Bezug auf zwei Arten von unterschiedlich anspruchsvollen Tests. Auf das Vorgehen bei sechs Texten stützte sich diese Studie, der Ablauf war immer der gleiche, Selbiges galt für die Testinstrumente.

Die Studierenden lasen sechs, 1.000 bis 1.200 Wörter lange Sachtexte mit multimedialen Inhalten ohne Zeitvorgabe und sollten je nach Gruppe entweder regelmäßig und final oder nur final am Ende des Lesens pro Text zwei Fragen beantworten. Diese Antworten bezogen sich auf den Anteil korrekter Antworten, die sie glaubten, bei fünf Multiple-Choice-Fragen zu memorierten Details und fünf Fragen zu höherem Leseverstehen zu erzielen, ehe sie diese Tests bearbeiteten. Der Anteil richtiger Antworten wurde als Mittelwert als Leistungsmaß für zwei Arten des Leseverstehens genutzt. Die Urteile zum eingeschätzten Erfolg bei der Testung bezogen sich auf die Frage, wie viel Prozent der beiden Fragetypen die Person glaubte, in der Testphase direkt nach dem Lesen korrekt zu beantworten.

D) Was waren die Hauptbefunde?

Im Zentrum stand der Gruppenvergleich von digital bzw. analog lesenden Personen. Die Ergebnisse lassen sich folgendermaßen bündeln:

- *Keine Differenz bei den Lesezeiten.* Hierbei wurden nur die beiden Gruppen mit den Schlussurteilen verglichen, da das Experiment mit den Gruppen, die Parallelurteile abgeben sollte, verzerrte Ergebnisse nach sich gezogen hätte. Wenngleich es nicht statistisch auffällig war, so lasen die Personen mit digitalen Texten diese etwas kürzer (rund 9 Minuten) als jene mit analogen Texten (10 min).
- *Keine Differenz bei den Urteilen über den Anteil richtiger Antworten:* Beide Gruppen attestierten sich ähnliche Werte. In der Gruppe der Parallelurteile ergaben sich allerdings Differenzen je nach Zeitpunkt der Urteile: Digital lesende Personen schätzten ihren Verstehenserfolg anfänglich deutlich höher ein als analog lesende Probandinnen und Probanden. Im Laufe der Aufgabenbearbeitung nivellierte sich diese Differenz und beide Gruppen steigerten ihre positive Einschätzung sukzessive.
- *Signifikant schlechtere Verstehensleistungen in der Gruppe mit digitalen Texten:* Wer analoge Texte las, beantwortete 72 Prozent aller Fragen korrekt, wer digitale Texte las, nur 63 Prozent. Damit deuten sich schon erwartbare Differenzen für die Ergebnisse der nachstehenden Analysen an.
- *Signifikante Differenzen in der Urteilsverzerrung (absolute Maße) je nach Verstehensleistungen und Lesemedium:* Die Personen überschätzten den Anteil mutmaßlich korrekter Antworten deutlich, wenn sie digital lasen, und sie taten es auch – unabhängig vom Lesemedium – bei den anspruchsvolleren Testfragen. Es gab hingegen keine Unterschiede in relativen Maßen.
- *Die Studierenden aus der Gruppe mit Parallelurteilen unterschieden sich in ihrem Verhalten je nach Lesemedium.* Dazu wurde berechnet, bei wie vielen Texten die Personen nach dem höchsten Urteil über den erwarteten Erfolg bei den Tests das Lesen beendet hatten. Anders gesagt: Es wurde geprüft, ob der Zeitpunkt für das Beenden des Lesens damit korrespondierte, ob sich die Personen sicher waren, die nachfolgenden Fragen mit dem höchsten Zutrauen in sich selbst zu beantworten. Hier gab es markante Differenzen: Die analog Lesenden waren sich sehr ähnlich darin, dass sie praktisch ausnahmslos erst dann die Fragen bearbeiteten, als sich ihre Urteile nicht mehr steigerten. Bei der Gruppe mit digitalen Dokumenten lag der Wert deutlich darunter. Diese Gruppe gab sich also früher mit dem Lesen zufrieden.

E) Diskussion

Diese Studie ist bemerkenswert, weil das Experiment differenzierte Befunde erbrachte. Die Leistungen von digital und analog lesenden Personen in zwei Facetten des Leseverstehens bei mehreren Texten unterschieden sich zugunsten der analog lesenden Teilnehmerinnen und Teilnehmer. Diese Leistungsdifferenz setzte sich in den absoluten Akkuratheitsurteilen zu einem prospektiven Erfolg insofern fort, dass die digital lesende Gruppe ihre Fähigkeiten positiver einschätzte als die Vergleichsgruppe. Sie tat dies teils bereits früher im Aufgabenprozess, und sie beendete auch bei mehr Texten den Leseprozess, obwohl

ihr Akkuratheitsurteil noch gar nicht gefestigt war. Diese erst durch sehr komplexe Auswertungen aufgetretenen Muster vermögen zu erklären, dass in der mangelnden Urteils-
kalibration ein potenzieller Mechanismus in den Minderleistungen beim Leseverstehen digitaler Texte besteht.

Die Befunde von Prinz und Kollegen (2020a) sind vor allem mit Blick auf die drei textbezogenen Faktoren insofern überraschend, weil eine höhere Urteils-
genauigkeit beim Leseverstehen digitaler Texte im Widerspruch dazu steht, dass
anders gelagerte Metaanalysen von schlechteren Leistungen berichten, wenn
Personen am Bildschirm lesen (Clinton, 2019; Delgado et al., 2018; Kong et al.,
2018). Hier scheint sich also eine Gegenläufigkeit abzuzeichnen, wenn man die
Mediumseffekte beim Verhältnis zweier Maße zueinander sucht oder absolute
Leseverstehensleistungen vergleicht. Ein weiterer Befund, der eine Kommentie-
rung wert ist, ist, dass das Lesen von mehr und von kürzeren Texten sich negativ
in der Urteilsgenauigkeit niederschlägt. Dies sind aber wiederum typische Eigen-
schaften von Texten, die im Internet üblicherweise gelesen werden. Das bedeutet,
dass sich hier Fragen ergeben, die aus lesedidaktischer Sicht von hoher Bedeutung
sind, wenn es darum geht, die Urteilsgenauigkeit zu erhöhen.

Die mangelnde Urteilskalibration lässt sich in Interventionen korrigieren (Lin
& Zabrocky, 1998). Zwei wichtige Ergebnisse der metaanalytischen Forschung
weisen den Weg: Eine realistischere Selbsteinschätzung lässt sich mittels Vermitt-
lung von *Lernstrategien* erreichen, die auf ein *tiefes Verstehen* abzielen (Gutierrez
de Blume, 2022), oder sich durch eine Breite verschiedener instruktionaler Ansät-
ze mit verschiedenen *Strategien des generativen Lernens* bewirken (Prinz et al.,
2020b; Yang et al., 2023). Diese beiden Ansätze eint, dass sie dabei helfen, durch
eine tiefenorientierte Verarbeitung des Lerngegenstands und verarbeitungsinten-
sive Prozesse eine umfassendere Basis für die Einschätzung des Lernerfolgs zu
gewinnen (sog. situationsmodellbasierte Fördermaßnahmen; Griffin et al., 2019).

6.4.3 Problemfeld 3: ungenügende spontane Strategie- anwendung (Verbund von Strategien auf Meta- und Objektebene)

Ein in der Lernstrategieforschung immer wieder thematisierter Topos besteht
darin, dass Personen Strategien falsch oder oberflächlich anwenden bzw. diese
Strategien nicht hilfreich sind, um das angestrebte Ziel des Lesens bzw. Lernens
zu erreichen (Dunlosky et al., 2013; Miyatsu et al., 2018). Das lässt sich zum ei-
nen über einen Mangel an metakognitivem Lernstrategiewissen erklären, aber
auch über Phasen-Modelle der Aneignung, welche eine Sequenz systematischer
lernpsychologischer Hürden postulieren, die Fehlanwendungen erwarten lassen
(Spörer & Brunstein, 2006).

Der Mangel an einer spontanen und korrekten Anwendung von Strategien lässt sich gut an jenen Studien illustrieren, in denen der Lerngegenstand komplex ist und das Vorwissen der Lernenden mutmaßlich eher als mittel einzustufen ist; denn in solchen Situationen lässt sich annehmen, dass selbstreguliertes Lernen erforderlich ist (Azevedo et al., 2010). In der Arbeitsgruppe rund um Azevedo sind mehrere Studien aus dem Feld der angewandten Forschung (Interventionsstudien) durchgeführt worden, um die Wirksamkeit von technisch oder personell unterstützten Umgangsweisen mit digitalen computerbasierten Lernumgebungen zu testen. Durch den Vergleich der jeweils geförderten Interventions- und der nicht geförderten Kontrollgruppe ließen sich Differenzen in Produkten und Prozessen feststellen.

Die *produktbezogenen* Maße wurden primär als Vorher-Nachher-Vergleiche in einem Verständnistest mit einem offenen Test zum Verstehen von systemischen Zusammenhängen (in diesem Fall: dem menschlichen Blutkreislauf) erfasst. Die Studien kommen in der Regel zum Resultat, dass die jeweils genutzten Varianten der Förderung dabei helfen, höhere Lernzuwächse zu erzielen (Azevedo & Cromley, 2004; Azevedo, Cromley & Seibert, 2004; Azevedo et al., 2005, 2007, 2008, 2011; Greene & Azevedo, 2007a; Moos & Azevedo, 2008a).

Die Personengruppen wurden je nach Gruppenzugehörigkeit zur Kontroll- bzw. Interventionsgruppe/zu den Interventionsgruppen in ihrem selbstregulatorischen Vorgehen verglichen, sodass prozessbezogene Vergleiche möglich waren. Stellvertretend für die Vergleiche der Personen lässt sich eine Gegenüberstellung anführen: Azevedo und Cromley (2004) haben ihre Ergebnisse prägnant zusammengefasst und die Aufgabebearbeitung der nicht-geförderten Personen als *ineffizientes extern fokussiertes Vorgehen* charakterisiert, während die geförderten Personen sich durch ein *internal fokussiertes und selbstreguliertes Vorgehen* auszeichneten. Das ineffiziente Vorgehen ähnelt auffallend jenem der leistungsschwächeren Personen bei Vergleichen aus Studien der Grundlagenforschung (s. Teilkap. 6.3.1). Um die inhaltliche Redundanz gering zu halten, enthält die nachfolgende Liste nur die wichtigsten Befunde stark kondensiert.

Personen haben Defizite in der spontanen Anwendung von selbstregulatorischen metakognitiven und kognitiven Strategien, was nachfolgend an drei Punkten feststellbar ist. Die ersten beiden Spiegelstriche betreffen metakognitive Kontroll- bzw. Überwachungsprozesse, der letzte Spiegelstrich in der Aufzählung vor allem die kognitiven Strategien auf der Objektebene:

- *Es misslingt den Personen eher, ihren Strategieeinsatz zu planen.* Auffällige Differenzen ergeben sich darin, das eigene Vorwissen zu aktivieren und einen Handlungsplan aufzustellen, was den in Interventionsstudien geförderten Personen besser glückte (Azevedo & Cromley, 2004; Azevedo, Cromley & Seibert, 2004; Azevedo et al., 2005, 2007, 2008, 2011; Greene & Azevedo, 2007a;

Moos & Azevedo, 2008a). Solche Strategien kamen gemäß dem analysierten Verhalten in den jeweiligen Kontrollgruppen seltener vor.

- *Ebenfalls weisen Personen ohne gezielte Förderung die Tendenz auf, sich weniger auf ihr eigenes Verstehen und die Verstehensüberwachung zu konzentrieren.* Die Personen zeichnen sich dadurch aus, dass ihre metakognitive Beurteilung und Überwachung sich weniger auf ihre eigenen Verstehensleistungen erstrecken, sondern eher auf die Verwertbarkeit von Informationen abzielen (Azevedo & Cromley, 2004; Azevedo, Cromley & Seibert, 2004; Azevedo et al., 2005, 2007, 2011). Darin deutet sich eine eher auf die äußere Umwelt orientierte Überwachung an anstatt einer Überwachung der eigenen Prozesse und Ziele.
- *Personen orientieren sich in ihren Lernstrategien eher an Oberflächen- und kontextuellen Merkmalen als an einer tiefenorientierten Durchdringung des Lerngegenstands.* So gibt es in den Kontrollgruppen der Studien die Tendenz, dass die Personen sich mehr auf umfassende Suchen zu begeben, statt sich länger auf das Lesen, Zusammenfassen, Annotieren und Strukturieren von Informationen zu konzentrieren (Azevedo & Cromley, 2004; Azevedo, Cromley & Seibert, 2004; Azevedo et al., 2007, 2008, 2011). Das Vorgehen wirkt damit erratisch.

Mehr als ein halbes Dutzend Studien mit unterschiedlichen Populationen und divergierenden Fördermaßnahmen kommt mit der gleichen Aufgabe und einem ähnlichen Auswertungsverfahren also zu einem deutlich musterhaften Gesamtergebnis: Die Kontrollgruppenmitglieder, die sowohl aus der Sekundarstufe als auch aus der Tertiärbildung stammten, neigten zu einem weniger stark metakognitiv gesteuerten Vorgehen bei gleichzeitig eher oberflächenorientierter Bearbeitung des Gegenstands. Das ist deshalb so bemerkenswert, weil die jeweilige Fördermaßnahmen teils darauf abstellten, vorab selbstregulatorische Vorgehensweisen zu erläutern oder adaptive Rückmeldungen zu geben. Beides sind vom Aufwand eher niederschwellige und kurzfristige Maßnahmen gewesen, welche sich jedoch in Produkten und Prozessen niederschlugen. Ohne solche eine Förderung jedoch mangelte es dem spontanen Vorgehen an einer erkennbaren selbstregulatorischen Herangehensweise.

6.5 Fazit

Die Selbstregulation beim Lesen digitaler Dokumente bezeichnet die Fähigkeit, zielgerichtet internale und externale Ressourcen zu nutzen und Prozesse zu planen, zu steuern und zu überwachen. Diese wird als dynamischer Prozess mit Rekursionen modelliert. Selbstregulation ist ein Sammelbegriff für verschiedene Perspektiven darauf, wie Personen ihre eigenen Aktivitäten koordinieren, um

einen Fortschritt auf ein Ziel hin zu erreichen – bzw. auch aus dynamischen Zielen Priorisierungen vorzunehmen (Neal et al., 2017). Dafür bedarf es der Nutzung des eigenen kognitiven Systems, um das Verhalten zu steuern. Entsprechend sind immer wieder Berührungspunkte zwischen Selbstregulation und den sogenannten *exekutiven Funktionen* aus Sicht der Theorie postuliert worden (Diamond, 2013). Drei Bestandteile der exekutiven Funktionen gelten als Kernbestand: das Arbeitsgedächtnis mit seinen begrenzten Fähigkeiten, Informationen für eine aktuelle Verarbeitung zur Verfügung zu halten, die inhibitorische Kontrolle als Fähigkeit, die Aufmerksamkeit zu fokussieren und Störungen zu minimieren, und schließlich die kognitive Flexibilität als Vermögen, zwischen Aufgaben und Zielen zielführend zu wechseln. Diese drei Bestandteile der exekutiven Funktionen erlauben es mit entsprechender gezielter Nutzung, selbstreguliert vorzugehen. Für das jeweilige aktuelle Vorgehen beim digitalen Lesen sind dann jedoch weitere prozedurale und teils domänenspezifische Metakognitionen nötig, also Lesestrategien im Umgang mit digitalen Texten im Besonderen und dem Lesen im Allgemeinen.

Selbstregulation bzw. ein Ausschnitt daraus, nämlich das selbstregulierte Lernen, welches wiederum das selbstregulierte Lesen unter sich vereint, stellt ein seit Langem bearbeitetes Forschungsfeld dar. In Theorie und Empirie haben diverse vorgelegte Arbeiten zu einer erheblichen Unübersichtlichkeit geführt, sodass in diesem Kapitel eine konzeptuelle Klärung anhand einer modellübergreifenden Betrachtung des Konstrukts und seiner theoretischen Modellierung den Ausgangspunkt markierte und dabei mehr Modelle berücksichtigte als bisherige Modellvergleiche (Panadero, 2017; Puustinen & Pulkkinen, 2001).

Verschiedene Modelle, von denen acht in diesem Kapitel eingangs inspiziert und anhand verschiedener Vergleichsdimensionen analysiert wurden, haben im Einzelnen bedeutende Unterschiede, wie sie das Zusammenspiel von unterschiedlich gewichteten Bestandteilen modellieren. Deutliche Unterschiede ergeben sich darin, wie Kontext, Verhalten und Motivation bzw. Emotionen berücksichtigt wurden. Diese divergierenden Fokussierungen erklären sich über die selbstregulatorischen Kernbereiche der jeweiligen Modelle und ihre Granularität. Zugleich zählen aller Divergenz zum Trotz zwei Bestandteile zum geteilten Kernbestand der Modelle in toto. Zum einen modellieren sämtliche Modelle Selbstregulation als *mehrphasiges Geschehen*, in welchem in mehr oder minder distinkten Phasen spezifische Komponenten in einem Funktionsverbund zusammenwirken. Zum anderen sind *metakognitives Wissen und dessen Anwendung* in einem Wechselspiel von vorgängiger Top-down-Planung, synchroner bottom-up-geleiteter Überwachung und Anpassung sowie abschließender metakognitiver Reflexion anzuführen. Metakognition bzw. metakognitiv gesteuerte Selbstregulation scheint damit der kleinste gemeinsame Nenner der gegenwärtigen Modellierungen zu sein, selbst wenn die Modelle in dieser Hinsicht variieren.

Metakognitive Selbstregulation als Nutzung lernstrategischer Kognitionen wird als Mehrebenen-Geschehen modelliert, bei dem auf einer Metaebene die Ziele für das aktuelle Vorgehen und die deklarativen, prozeduralen und konditionalen Wissensbestände ebenso zu lokalisieren sind wie auch eine Repräsentation der Aktivitäten auf einer zweiten, darunter befindlichen Ebene. Diese Objektebene bildet den Ort der unmittelbaren kognitiven Auseinandersetzung mit dem Gegenstand, also der Interaktion von Leseprozessen und Textinhalten. Was auf der Objektebene geschieht, wird mittels sogenannter Kontrollprozesse initiiert und modifiziert, doch es gibt auch einen gegenläufigen Informationsfluss, nämlich die metakognitive Überwachung. Kontroll- und Überwachungsprozesse verbinden somit die beiden Ebenen und sorgen für eine kontinuierliche metakognitive Selbstregulation, welche die Dynamik des Fortschritts fokussiert und als treibende Kraft fungiert.

Eines der Modelle, welches insbesondere im Kontext des Lernens mit digitalen Texten einschlägig wirkt und immer wieder explizit als Rückgrat empirischer Studien fungiert, ist aus metakognitiver Sicht das COPES-Modell (Greene & Azevedo, 2007b; Winne, 2001; Winne & Hadwin, 1998). Es nutzt die etablierte Unterscheidung von metakognitiven planerisch-regulierenden Kontroll- und den Fortschritt fokussierenden Überwachungsprozessen und verbindet diese mit zentralen Komponenten. Insbesondere der interne (und im Modell auch über externe Leistungsrückmeldungen stattfindende) Vergleich mit einem Soll (Standards) ist treibende Kraft für die Prozesse und deren Ergebnisse, die Produkte. Das Modell impliziert – wie so viele andere Modelle auch – eine zeitliche Trennung in Phasen, welche sich aber als lediglich idealtypische Sequenz entpuppt, da in der Realität eher ein rekursives und dynamisches Zusammenspiel vorzufinden ist, also ein kontingentes Geschehen. Der Hintergrund ist, dass sämtliche beteiligten Faktoren und Komponenten im Laufe der Aufgabenbearbeitung zu sich verändernden Konstellationen führen, welche über Rückkopplungen jederzeit neue Prozesse initiieren und vorantreiben. Darum ist die Adaptivität im Vorgehen entscheidend, also eine Anpassung des eigenen Vorgehens an Aufgabenerfordernisse und den Verlauf der Aufgabenbearbeitung.

Diesem Gedanken folgend lassen sich diverse Strategien, wie sie aus der Leseforschung unter dem Mantelbegriff des konstruktiv-responsiven Lesens seit längerer Zeit zusammengetragen wurden, als spezifische Varianten in Funktionsgruppen des selbstregulierten Vorgehens werten. Verschiedene Strategien erfüllen unterschiedliche Zwecke, und es kommt hierbei auf den Verbund an. In der Regel sind komplexe Aufgaben nicht mit einer Strategie allein zu bewältigen, was sich bereits aus dem teils sehr offenen Aufgabencharakter ergibt, zu dem es gleichfalls zählt, metakognitiv unterstützt Ziele festzulegen und die geeigneten Lese Strategien so auszuwählen, auszuführen und ggf. zu modifizieren, dass sie den Fortschritt zu erreichen helfen, den die lesende Person anstrebt. Die Erkenntnisse aus der Leseforschung sind hilfreich dafür, spezifische und

allgemeine Strategien mit variierender Nutzungsbreite zu kartieren. So werden deutliche Konstanten ebenso sichtbar wie Spezifika, die sich beim digitalen Lesen abzuzeichnen beginnen, darunter der Umgang mit Kohärenzproblemen oder das Finden von Informationen und Navigationspfaden. Dabei zeichnet sich bei aller Vorläufigkeit des Forschungsstandes ab, dass digitales Lesen sowohl Strategien des analogen Lesens einzelner Texte erfordert als auch für das digitale Lesen spezifisch erforderliche Vorgehensweisen verlangt. Dies unterstreicht die Notwendigkeit eines breiten Strategierepertoires.

Umso wichtiger ist es, dass Einflussfaktoren auf die Strategienutzung bekannt sind und in theoretische Modelle einfließen, welche wiederum empirisch unterfüttert und getestet werden. Die theoretischen Modelle mit ihren teils explizierten individuellen Bedingungen geben erste Hinweise. Epistemische Überzeugungen und inhaltliches Vorwissen sind zwei solcher Bedingungen, die beide als Ressourcen fungieren. Fortgeschrittenes Wissen und fortgeschrittene Überzeugungen wirken auf spezifische Weisen als günstige Ressource. Ein höheres Wissen scheint eine stärkere metakognitive Verstehensüberwachung ebenso zu ermöglichen wie eine höhere Informationsintegration, die über kognitive Schlussfolgerungen (Inferenzen) realisiert wird. Epistemische Überzeugungen als subjektive Theorien darüber, was Wissen und dessen Herstellung ausmacht, ist demgegenüber ein entfernterer Einflussfaktor. Diese subjektiven Überzeugungen wirken eher indirekt, indem sie als filternde Wahrnehmungen die Elaboriertheit von Standards beeinflussen. Naive Überzeugungen wirken eher kongruent zu oberflächenorientiertem Vorgehen, während elaborierte Überzeugungen aufwändigere, komplexere und auf tieferes Verstehen abzielende Strategienutzungen nahelegen. Und tatsächlich gibt es hierfür Hinweise, wenn man Leistungsgruppen vergleicht, die teils eigens danach gebildet wurden, dass Lernzuwächse besonders stark waren. Leistungsstärkere Personen nutzen speziell in schwierigen Aufgaben mehr metakognitive Überwachungsstrategien, sie wenden evaluierende, elaborierende und damit das eigene Vorwissen stärker beanspruchende kognitive Strategien eher an, und sie orchestrieren zudem ihre Strategien rekursiver und flexibler. Ihr Vorgehen wirkt damit angemessener und anspruchsvoller aus Sicht der Selbstregulation.

Dies leitet über zu der Frage, welche Schwierigkeiten hierbei bestehen. Die Fähigkeit, das eigene Vorgehen beim Lernen und beim Lesen mithilfe der eigenen kognitiven, metakognitiven, motivationalen, behavioralen und kontextuellen Ressourcen positiv und zielbezogen zu beeinflussen, kann allein aufgrund der Vielzahl beteiligter Komponenten gemäß den Theorien als anspruchsvoll deklariert werden. Die Vielzahl zu koordinierender und zu synchronisierender Prozesse, die zudem adaptiv auf Erfordernisse der Aufgaben und die Kontingenz im Gesamtprozess reagieren, birgt die Gefahr der Fehlanwendungen und Probleme. Drei ausgewählte dieser Problemlagen hat dieses Kapitel fokussiert, und alle drei Problemlagen eint, dass sich Defizite in der Anwendung von metakognitiven Strategien manifestiert haben.

Diese betreffen erstens die mangelnde Planung im Sinne von Top-down-Prozessen von der Meta- auf die Objektebene. Hierbei wurden Probleme der Aufmerksamkeitslenkung genauer fokussiert, die sich aus zwei Quellen speisen können, nämlich zum einen dem Multitasking als dem verstehensmindernden Impetus, neben dem Lesen auch noch eine weitere Aufgabe zu bewältigen, und zum anderen aus abschweifenden Gedanken, die den Fokus der Aufmerksamkeit verlagern und damit eine angemessene Verstehensleistung unterwandern. Im strengen Sinne zeigen sich also nicht mangelnde Planungen des Vorgehens, sondern vielmehr Probleme darin, auf eine adäquate Weise die Limitationen des eigenen kognitiven Systems, darunter die begrenzte Verarbeitungskapazität des Arbeitsgedächtnisses und vor allem die inhibitorische Kontrolle, so zu adressieren, dass die Primäraufgabe des Lesens erfolgreich bewältigt wird.

Zweitens ergeben sich Probleme darin, im Sinne eines Bottom-up-Prozesses die Verstehens(miss)erfolge kalibriert metakognitiv zu überwachen. Dutzende Studien haben das stabile Muster erbracht, dass Leserinnen und Leser nur eine vage bzw. eine ungenaue Vorstellung davon haben, was sie faktisch zu leisten imstande sind. Teils lässt sich das darauf zurückführen, ob metakognitive Urteile zum Lernen aus Texten vor oder nach dem Bearbeiten zu treffen waren. Individuelle Merkmale wie ein höheres Alter und stabilere Urteilsgrundlagen wie weniger zu beurteilende Aufgaben oder längere Texte wirken hierbei mildernd. Da sich das Lesen im Internet in aller Regel auf kürzere und diverse (Hyper-) Texte bezieht (Bråten et al., 2020), zeichnet sich dieses Lesen durch potenzielle Erschwernisse aus.

Ein drittes Problemfeld, das sich aus einem Set von vergleichbaren Experimentalstudien mit unterschiedlichen Populationen insbesondere in den jeweiligen Kontrollgruppen abzeichnet, betrifft sowohl meta- als auch kognitive Strategien beim Aufbau mentaler Repräsentationen komplexer Sachverhalte. In diesen Studien waren durch Prä-Post-Vergleiche die Veränderungen in Abhängigkeit experimenteller Variationen Gegenstand des Interesses. Studienübergreifend kristallisierte sich heraus, dass die Kontrollgruppenmitglieder gemäß den Produktdaten wenig dazugelernt hatten, was mutmaßlich dem selbstregulatorisch defizienten Vorgehen geschuldet war. Die Prozessdaten weisen darauf hin, dass eine starke Außenorientierung mit dem Finden von Informationen in einer digitalen Lernumgebung statt einer tiefenorientierten Verarbeitung auf der Objektebene den dominanten Modus ausmachte. Hinzu kommen auf der Metaebene zu lokalisierende Probleme der mangelnden Planung (vor allem der Vorwissensaktivierung) und der kontinuierlichen Verstehensüberwachung. Damit wirkt das Vorgehen in seiner Gänze eher unangemessen und ungeeignet, wenn es darum geht, sich gemessen an den Aufgabenerfordernissen mit den Gegenständen kognitiv ausreichend tief auseinanderzusetzen.

Dass die Problemlagen bei der Selbstregulation beim Lernen aus Texten an dieser Stelle so ausführlich thematisiert wurden, ist nicht zufällig erfolgt. Diese Problemlagen illustrieren, dass es selbst erwachsenen Personen nicht ausreichend glückt, beim (digitalen) Lesen ihre Ressourcen zielführend einzusetzen. Spiegelbildlich zu den hohen Anforderungen, welche das digitale Lesen an die Leserinnen und Leser stellt – und hier sind nur die kognitiven und metakognitiven Facetten zur Sprache gekommen –, wirkt es nötig, diese Fähigkeiten systematisch zu schulen und das metakognitive Wissen und dessen Anwendung in der Regulation aufzubauen.

Literaturverzeichnis

- Aagaard, J. (2019). Multitasking as Distraction: A Conceptual Analysis of Media Multitasking Research. *Theory & Psychology*, 29(1), 87–99. <https://doi.org/10.1177/0959354318815766>
- Abel, F. & Barzilai, S. (2023). Can Students Evaluate Scientific YouTube Videos? Examining Students' Strategies and Criteria for Evaluating Videos versus Webpages on Climate Change. *Journal of Computer Assisted Learning*, 39(2), 558–577. <https://doi.org/10.1111/jcal.12762>
- Abel, R., Roelle, J. & Stadler, M. (in Druck). Whom to Believe? Fostering Source Evaluation Skills with Interleaved Presentation of Untrustworthy and Trustworthy Social Media Sources. *Discourse Processes*. <https://doi.org/10.1080/0163853X.2024.2339733>
- Abendroth, J. & Richter, T. (2020). Text-Belief Consistency Effects in Adolescents' Comprehension of Multiple Documents from the Web. *Journal for the Study of Education and Development*, 43(1), 60–100.
- Abendroth, J. & Richter, T. (2021). How to Understand What You Don't Believe: Metacognitive Training Prevents Belief-Biases in Multiple Text Comprehension. *Learning and Instruction*, 71, Artikel 101394, 1–16. <https://doi.org/10.1016/j.learninstruc.2020.101394>
- Abendroth, J. & Richter, T. (2023). Reading Perspectives Moderate Text-Belief Consistency Effects in Eye Movements and Comprehension. *Discourse Processes*, 60(2), 119–140. <https://doi.org/10.1080/0163853X.2023.2172300>
- Ackerman, J. M. (1991). Reading, Writing, and Knowing: The Role of Disciplinary Knowledge in Comprehension and Composing. *Research in the Teaching of English*, 25(2), 133–178.
- Ackerman, R. & Goldsmith, M. (2011). Metacognitive Regulation of Text Learning: On Screen versus on Paper. *Journal of Experimental Psychology: Applied*, 17(1), 18–32. <https://doi.org/10.1037/a0022086>
- Adams, E. J., Nguyen, A. T. & Cowan, N. (2018). Theories of Working Memory: Differences in Definition, Degree of Modularity, Role of Attention, and Purpose. *Language, Speech, and Hearing Services in Schools*, 49(3), 340–355. https://doi.org/10.1044/2018_LSHSS-17-0114
- Afflerbach, P. & Cho, B.-Y. (2009). Identifying and Describing Constructively Responsive Comprehension Strategies in New and Traditional Forms of Reading. In S. E. Israel & G. G. Duffy (Hrsg.), *Handbook of Research on Reading Comprehension* (S. 69–90). Routledge.
- Afflerbach, P. & Cho, B.-Y. (2010). Determining and Describing Reading Strategies: Internet and Traditional Forms of Reading. In H. S. Waters & W. Schneider (Hrsg.), *Metacognition, Strategy Use, and Instruction* (S. 201–225). Guilford Press.
- Afflerbach, P., Cho, B.-Y. & Kim, J.-Y. (2014). Inaccuracy and Reading in Multiple Text and Internet/Hypertext Environment. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 403–424). MIT Press.
- Afflerbach, P., Cho, B.-Y., Kim, J.-Y., Crassas, M. E. & Doyle, B. (2013). Reading: What Else Matters besides Strategies and Skills? *The Reading Teacher*, 66(6), 440–448. <https://doi.org/10.1002/TRTR.1146>
- Ahmed, Y. & Wagner, R. (2020). A „Simple“ Illustration of a Joint Model of Reading and Writing Using Metaanalytic Structural Equation Modeling (MASEM). In R. A. Alves, T. Limpo & R. M. Joshi (Hrsg.), *Reading-Writing Connections: Towards Integrative Literacy Science* (S. 55–75). Springer. <https://doi.org/10.1007/978-3-030-38811-94>
- Ainsworth, S. (1999). The Functions of Multiple Representations. *Computers & Education*, 33(2), 131–152. [https://doi.org/10.1016/S0360-1315\(99\)00029-9](https://doi.org/10.1016/S0360-1315(99)00029-9)
- Ainsworth, S. (2006). DeFT: A Conceptual Framework for Considering Learning with Multiple Representations. *Learning and Instruction*, 16(3), 183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Ainsworth, S. (2022). The Multiple Representations Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 158–170). Cambridge University Press. <https://doi.org/10.1017/9781108894333.016>
- Alamargot, D. & Chanquoy, L. (2001). *Through the Models of Writing*. Kluwer Academic Publishers.

- Aleknavičiūtė, V., Lehtinen, E. & Södervik, I. (2023). Thirty Years of Conceptual Change Research in Biology – A Review and Meta-Analysis of Intervention Studies. *Educational Research Review*, 41, Artikel 100556, 1–25. <https://doi.org/10.1016/j.edurev.2023.100556>
- Alemdag, E. & Cagiltay, K. (2018). A Systematic Review of Eye Tracking Research on Multimedia Learning. *Computers & Education*, 125, 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>
- Alexander, P. A. (1986). College Students' Use of Task-Related Information during Studying. *Reading Research and Instruction*, 25(2), 91–101. <https://doi.org/10.1080/19388078609557865>
- Alexander, P. A. (2019). The Art (and Science) of Seduction: Why, When, and for Whom Seductive Details Matter. *Applied Cognitive Psychology*, 33(1), 142–148. <https://doi.org/10.1002/acp.3510>
- Alexander, P. A. (2023). „Here Be Dragons!“ Mapping the Realm of Higher-Order, Critical, and Critical-Analytic Thinking. *Educational Psychology Review*, 35(2), Artikel 42, 1–18. <https://doi.org/10.1007/s10648-023-09762-0>
- Alexander, P. A., Dinsmore, D. L., Parkinson, M. M. & Winters, F. I. (2011). Self-Regulated Learning in Academic Domains. In B. J. Zimmerman & D. H. Schunk (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (S. 393–407). Routledge.
- Alexander, P. A. & Disciplined Reading and Learning Research Laboratory. (2020). Relational Reasoning: The Bedrock of Integration within and across Multiple Representations, Documents, and Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 401–424). Routledge.
- Alexander, P. A. & Fox, E. (2018). Reading Research and Practice Over the Decades: A Historical Analysis. In D. E. Alvermann, N. J. Unrau, M. Sailors & R. B. Ruddell (Hrsg.), *Theoretical Models and Processes of Literacy* (7. Aufl., S. 35–64). Routledge.
- Alexander, P. A., Fox, E., Maggioni, L., Loughlin, S., Baggetta, P., Dinsmore, D. L., Grossnickle, E. M., List, A., Parkinson, M. M., Winters, F. I. & Dumas, D. (2012). Reading into the Future: Competence for the 21st Century. *Educational Psychologist*, 47(4), 259–280. <https://doi.org/10.1080/00461520.2012.722511>
- Alexander, P. A., Graham, St. & Harris, K. R. (1998). A Perspective on Strategy Research: Progress and Prospects. *Educational Psychology Review*, 10(2), 129–154. <https://doi.org/10.1023/A:1022185502996>
- Alexander, P. A., Grossnickle, E. M., Dumas, D. & Hattan, C. (2018). A Retrospective and Prospective Examination of Cognitive Strategies and Academic Development: Where Have We Come in Twenty-Five Years? In A. O'Donnell (Hrsg.), *The Oxford Handbook of Educational Psychology*. Oxford University Press.
- Alexander, P. A. & Schoute, E. C. (2022). *Knowledge, Knowing, and Information: Their Meaning and Meaningfulness to Learning and Development*. Routledge. <https://doi.org/10.4324/9781138609877-REF203-1>
- Allen, L. K. & McNamara, D. S. (2020). Defining Deep Reading Comprehension for Diverse Readers. In E. B. Moje, P. Afflerbach, P. Enciso & N. K. Lesaux (Hrsg.), *Handbook of Reading Research: Volume V* (S. 261–276). Routledge.
- Alpizar, D., Adesope, O. O. & Wong, R. M. (2020). A Meta-Analysis of Signaling Principle in Multimedia Learning Environments. *Educational Technology Research and Development*, 68(4), 2095–2119. <https://doi.org/10.1007/s11423-020-09748-7>
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman.
- Anderson, R. C. & Pichert, J. W. (1978). Recall of Previously Unrecallable Information Following a Shift in Perspective. *Journal of Verbal Learning and Verbal Behavior*, 17(1), 1–12. [https://doi.org/10.1016/S0022-5371\(78\)90485-1](https://doi.org/10.1016/S0022-5371(78)90485-1)
- Anderson, R. C., Pichert, J. W. & Shirey, L. L. (1983). Effects of the Reader's Schema at Different Points in Time. *Journal of Educational Psychology*, 75(2), 271–279. <https://doi.org/10.1037/0022-0663.75.2.271>
- Andre, T. (1979). Does Answering Higher-Level Questions while Reading Facilitate Productive Learning? *Review of Educational Research*, 49(2), 280–318. <https://doi.org/10.2307/1169962>

- Andresen, A., Anmarkrud, Ø. & Bråten, I. (2019). Investigating Multiple Source Use among Students with and without Dyslexia. *Reading and Writing*, 32(5), 1149–1174. <https://doi.org/10.1007/s11145-018-9904-z>
- Anmarkrud, Ø., Bråten, I., Florit, E. & Mason, L. (2022). The Role of Individual Differences in Sourcing: A Systematic Review. *Educational Psychology Review*, 34(2), 749–792. <https://doi.org/10.1007/s10648-021-09640-7>
- Anmarkrud, Ø., Bråten, I. & Strømso, H. I. (2014). Multiple-Documents Literacy: Strategic Processing, Source Awareness, and Argumentation when Reading Multiple Conflicting Documents. *Learning and Individual Differences*, 30, 64–76. <https://doi.org/10.1016/j.lindif.2013.01.007>
- Anmarkrud, Ø., McCrudden, M. T., Bråten, I. & Strømso, H. I. (2013). Task-Oriented Reading of Multiple Documents: Online Comprehension Processes and Offline Products. *Instructional Science*, 41(5), 873–894. <https://doi.org/10.1007/s11251-013-9263-8>
- Anttonen, R., Räikkönen, E., Kiili, K. & Kiili, C. (in Druck). Sixth Graders Evaluating Online Texts: Self-Efficacy Beliefs Predict Confirming But Not Questioning the Credibility. *Scandinavian Journal of Educational Research*. <https://doi.org/10.1080/00313831.2023.2228834>
- Appel, M., Marker, C. & Gnamb, T. (2020). Are Social Media Ruining Our Lives? A Review of Meta-Analytic Evidence. *Review of General Psychology*, 24(1), 60–74. <https://doi.org/10.1177/1089268019880891>
- Appel, M. & Schreiner, C. (2014). Digitale Demenz? Mythen und wissenschaftliche Befundlage zur Auswirkung von Internetnutzung. *Psychologische Rundschau*, 65(1), 1–10. <https://doi.org/10.1026/0033-3042/a000186>
- Arias-Gundin, O. & Robledo, P. (2023). An Analysis of the Psychometric Properties of the Writing-Specific Cognitive Strategies Questionnaire for Undergraduate Students. *Frontiers in Psychology*, 14, Artikel 1274478, 1–9. <https://doi.org/10.3389/fpsyg.2023.1274478>
- Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, 63(1), 18–25. <https://doi.org/10.1026/0033-3042/a000106>
- Artelt, C., Schiefele, U., Schneider, W. & Stanat, P. (2002). Leseleistungen deutscher Schülerinnen und Schüler im internationalen Vergleich (PISA): Ergebnisse und Erklärungsansätze. *Zeitschrift für Erziehungswissenschaft*, 5(1), 6–27. <https://doi.org/10.1007/s11618-002-0002-1>
- Artelt, C. & Schneider, W. (2015). Cross-Country Generalizability of the Role of Metacognitive Knowledge in Students' Strategy Use and Reading Competence. *Teachers College Record*, 117(1), 1–32. <https://doi.org/10.1177/016146811511700109>
- Artmann, B., Scheibenzuber, C., Fendt, M. & Nistor, N. (2023). Lateral Reading Against Misinformation: The Cognitive Processing Level Mediates the Intervention Effect on the News Credibility Assessment Performance. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou & T. Papathoma (Hrsg.), *Responsive and Sustainable Educational Futures: 18th European Conference on Technology Enhanced Learning, EC-TEL 2023, Aveiro, Portugal, September 4–8, 2023, Proceedings* (S. 531–535). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42682-7_37
- Ayres, P. & Sweller, J. (2022). The Split-Attention Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 199–211). Cambridge University Press. <https://doi.org/10.1017/9781108894333.020>
- Azevedo, R. (2009). Theoretical, Conceptual, Methodological, and Instructional Issues in Research on Metacognition and Self-Regulated Learning: A Discussion. *Metacognition and Learning*, 4(1), 87–95. <https://doi.org/10.1007/s11409-009-9035-7>
- Azevedo, R. (2020). Reflections on the Field of Metacognition: Issues, Challenges, and Opportunities. *Metacognition and Learning*, 15(2), 91–98. <https://doi.org/10.1007/s11409-020-09231-x>
- Azevedo, R. & Cromley, J. G. (2004). Does Training on Self-Regulated Learning Facilitate Students' Learning with Hypermedia? *Journal of Educational Psychology*, 96(3), 523–535. <https://doi.org/10.1037/0022-0663.96.3.523>
- Azevedo, R., Cromley, J. G., Moos, D. C., Greene, J. A. & Winters, F. I. (2011). Adaptive Content and Process Scaffolding: A Key to Facilitating Students' Self-Regulated Learning with Hypermedia. *Psychological Test and Assessment Modeling*, 53(1), 106–140.

- Azevedo, R., Cromley, J. G. & Seibert, D. (2004). Does Adaptive Scaffolding Facilitate Students' Ability to Regulate Their Learning with Hypermedia? *Contemporary Educational Psychology*, 29(3), 344–370. <https://doi.org/10.1016/j.cedpsych.2003.09.002>
- Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C. & Greene, J. A. (2005). Adaptive Human Scaffolding Facilitates Adolescents' Self-Regulated Learning with Hypermedia. *Instructional Science*, 33(5), 381–412. <https://doi.org/10.1007/s11251-005-1273-8>
- Azevedo, R., Greene, J. A. & Moos, D. C. (2007). The Effect of a Human Agent's External Regulation upon College Students' Hypermedia Learning. *Metacognition and Learning*, 2(2), 67–87. <https://doi.org/10.1007/s11409-007-9014-9>
- Azevedo, R., Guthrie, J. T. & Seibert, D. (2004). The Role of Self-Regulated Learning in Fostering Students' Conceptual Understanding of Complex Systems with Hypermedia. *Journal of Educational Computing Research*, 30(1-2), 87–111. <https://doi.org/10.2190/DVWX-GMIT-6THQ-5WC7>
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I. & Cromley, J. G. (2008). Why Is Externally-Facilitated Regulated Learning More Effective than Self-Regulated Learning with Hypermedia? *Educational Technology Research and Development*, 56(1), 45–72. <https://doi.org/10.1007/s11423-007-9067-0>
- Azevedo, R., Moos, D. C., Johnson, A. M. & Chauncey, A. D. (2010). Measuring Cognitive and Metacognitive Regulatory Processes During Hypermedia Learning: Issues and Challenges. *Educational Psychologist*, 45(4), 210–223. <https://doi.org/10.1080/00461520.2010.515934>
- Azevedo, R. & Witherspoon, A. (2008). Detecting, Tracking, and Modeling Self-Regulatory Processes during Complex Learning with Hypermedia. In A. V. Samsonovic (Hrsg.), *Biologically Inspired Cognitive Architectures: Papers from the AAAI Fall Symposium* (S. 16–26). AAAI Press.
- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Baggetta, P. & Alexander, P. A. (2016). Conceptualization and Operationalization of Executive Function. *Mind, Brain, and Education*, 10(1), 10–33. <https://doi.org/10.1111/mbe.12100>
- Baillet, S. D. & Keenan, J. M. (1986). The Role of Encoding and Retrieval Processes in the Recall of Text. *Discourse Processes*, 9(3), 247–268. <https://doi.org/10.1080/01638538609544643>
- Bannert, M., Reimann, P. & Sonnenberg, C. (2014). Process Mining Techniques for Analysing Patterns and Strategies in Students' Self-Regulated Learning. *Metacognition and Learning*, 9(2), 161–185. <https://doi.org/10.1007/s11409-013-9107-6>
- Barzilai, S. & Eshet-Alkalai, Y. (2015). The Role of Epistemic Perspectives in Comprehension of Multiple Author Viewpoints. *Learning and Instruction*, 36, 86–103. <https://doi.org/10.1016/j.learninstruc.2014.12.003>
- Barzilai, S. & Ka'adan, I. (2017). Learning to Integrate Divergent Information Sources: The Interplay of Epistemic Cognition and Epistemic Metacognition. *Metacognition and Learning*, 12(2), 193–232. <https://doi.org/10.1007/s11409-016-9165-7>
- Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T. & Ben-Yishai, R. (2020). Making Sources Visible: Promoting Multiple Document Literacy with Digital Epistemic Scaffolds. *Computers & Education*, 157, Artikel 103980, 1–17. <https://doi.org/10.1016/j.compedu.2020.103980>
- Barzilai, S., Tal-Savir, D., Abed, F., Mor-Hagani, S. & Zohar, A. R. (2023). Mapping Multiple Documents: From Constructing Multiple Document Models to Argumentative Writing. *Reading and Writing*, 36(4), 809–847. <https://doi.org/10.1007/s11145-021-10208-8>
- Barzilai, S., Thomm, E. & Shlomi-Elooz, T. (2020). Dealing with Disagreement: The Roles of Topic Familiarity and Disagreement Explanation in Evaluation of Conflicting Expert Claims and Sources. *Learning and Instruction*, 69, Artikel 101367, 1–15. <https://doi.org/10.1016/j.learninstruc.2020.101367>
- Barzilai, S., Tzadok, E. & Eshet-Alkalai, Y. (2015). Sourcing while Reading Divergent Expert Accounts: Pathways from Views of Knowing to Written Argumentation. *Instructional Science*, 43(6), 737–766. <https://doi.org/10.1007/s11251-015-9359-4>
- Barzilai, S. & Weinstock, M. (2020). Beyond Trustworthiness: Comprehending Multiple Source Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 123–140). Routledge.
- Barzilai, S. & Zohar, A. (2012). Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cognition and Instruction*, 30(1), 39–85. <https://doi.org/10.1080/07370008.2011.636495>

- Barzilai, S. & Zohar, A. (2014). Reconsidering Personal Epistemology as Metacognition: A Multi-faceted Approach to the Analysis of Epistemic Thinking. *Educational Psychologist*, 49(1), 13–35. <https://doi.org/10.1080/00461520.2013.863265>
- Barzilai, S., Zohar, A. R. & Mor-Hagani, S. (2018). Promoting Integration of Multiple Texts: A Review of Instructional Approaches and Practices. *Educational Psychology Review*, 30(3), 973–999. <https://doi.org/10.1007/s10648-018-9436-8>
- Beker, K., Jolles, D. & van den Broek, P. W. (2017). Meaningful Learning from Texts: The Construction of Knowledge. In J. A. León & I. Escudero (Hrsg.), *Reading Comprehension in Educational Settings* (S. 29–62). John Benjamins.
- Ben-Eliyahu, A. & Bernacki, M. L. (2015). Addressing Complexities in Self-Regulated Learning: A Focus on Contextual Factors, Contingencies, and Dynamic Relations. *Metacognition and Learning*, 10(1), 1–13. <https://doi.org/10.1007/s11409-015-9134-6>
- Bereiter, C. & Scardamalia, M. (1987). *The Psychology of Written Composition*. Lawrence Erlbaum.
- Bernacki, M. L., Byrnes, J. P. & Cromley, J. G. (2012). The Effects of Achievement Goals and Self-Regulated Learning Behaviors on Reading Comprehension in Technology-Enhanced Learning Environments. *Contemporary Educational Psychology*, 37(2), 148–161. <https://doi.org/10.1016/j.cedpsych.2011.12.001>
- Binbasaran Tüysüzöğlu, B. & Greene, J. A. (2015). An Investigation of the Role of Contingent Metacognitive Behavior in Self-Regulated Learning. *Metacognition and Learning*, 10(1), 77–98. <https://doi.org/10.1007/s11409-014-9126-y>
- Binks, A., Toniolo, A. & Nacenta, M. A. (2022). Representational Transformations: Using Maps to Write Essays. *International Journal of Human-Computer Studies*, 165, Artikel 102851, 1–20. <https://doi.org/10.1016/j.ijhcs.2022.102851>
- Biswas, G., Baker, R. S. & Paquette, L. (2018). Data Mining Methods for Assessing Self-Regulated Learning. In D. H. Schunk & Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 388–403). Routledge. <https://doi.org/10.4324/9781315697048-25>
- Blaum, D., Griffin, T. D., Wiley, J. & Britt, M. A. (2017). Thinking about Global Warming: Effect of Policy-Related Documents and Prompts on Learning about Causes of Climate Change. *Discourse Processes*, 54(4), 303–316. <https://doi.org/10.1080/0163853X.2015.1136169>
- Blöbaum, B. (2016). Key Factors in the Process of Trust. On the Analysis of Trust under Digital Conditions. In B. Blöbaum (Hrsg.), *Trust and Communication in a Digitized World: Models and Concepts of Trust Research* (S. 3–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-28059-2_1
- Boekaerts, M. (2006). Self-Regulation and Effort Investment. In I. E. Sigel & K. A. Renninger (Hrsg.), *Handbook of Child Psychology, Vol. 4: Child Psychology in Practice* (6. Aufl., S. 345–377). Wiley. <https://doi.org/10.1002/9780470147658.chpsy0409>
- Boekaerts, M. (2011). Emotions, Emotion Regulation, and Self-Regulation of Learning. In B. J. Zimmerman & D. H. Schunk (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (S. 408–425). Routledge.
- Boekaerts, M. & Niemivirta, M. (2000). Self-Regulated Learning: Finding a Balance between Learning Goals and Ego-Protective Goals. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Handbook of Self-Regulation* (S. 417–450). Academic Press.
- Boekaerts, M., Pintrich, P. R. & Zeidner, M. (Hrsg.). (2000). *Handbook of Self-Regulation*. Academic Press.
- Boer, H., Donker, A. S., Kostons, D. D. & van der Werf, G. P. (2018). Long-Term Effects of Metacognitive Strategy Instruction on Student Academic Performance: A Meta-Analysis. *Educational Research Review*, 24, 98–115. <https://doi.org/10.1016/j.edurev.2018.03.002>
- Bohn-Gettler, C. M. & Kendeou, P. (2014). The Interplay of Reader Goals, Working Memory, and Text Structure during Reading. *Contemporary Educational Psychology*, 39(3), 206–219. <https://doi.org/10.1016/j.cedpsych.2014.05.003>
- Bohn-Gettler, C. M. & McCrudden, M. T. (2018). Effects of Task Relevance Instructions and Topic Beliefs on Reading Processes and Memory. *Discourse Processes*, 55(4), 410–433. <https://doi.org/10.1080/0163853X.2017.1292824>
- Bol, L. & Hacker, D. J. (2012). Calibration Research: Where Do We Go from Here? *Frontiers in Psychology*, 3, Artikel 229, 1–6. <https://doi.org/10.3389/fpsyg.2012.00229>

- Bonifacci, P., Viroli, C., Vassura, C., Colombini, E. & Desideri, L. (2023). The Relationship between Mind Wandering and Reading Comprehension: A Meta-Analysis. *Psychonomic Bulletin & Review*, 30(1), 40–59. <https://doi.org/10.3758/s13423-022-02141-w>
- Borkowski, J. G., Chan, L. K. S. & Muthukrishna, N. (2000). A Process-Oriented Model of Metacognition: Links between Motivation and Executive Functioning. In G. Schraw & J. C. Impara (Hrsg.), *Issues in the Measurement of Metacognition* (S. 1–41). Buros Institute of Mental Measurements and Department of Educational Psychology.
- Boscolo, P., Arfè, B. & Quarisa, M. (2007). Improving the Quality of Students' Academic Writing: An Intervention Study. *Studies in Higher Education*, 32(4), 419–438. <https://doi.org/10.1080/03075070701476092>
- Boudreau, R. L., Wood, E., Willoughby, T. & Specht, J. (1999). Evaluating the Efficacy of Elaborative Strategies for Remembering Expository Text. *Alberta Journal of Educational Research*, 45(2), 170–183.
- Braasch, J. L. G., Anmarkrud, Ø., Andresen, A., Ferguson, L. E. & Kardash, C. M. (2023). Beliefs about the Malleability of Working Memory Guide College Students' Evaluations of Belief-Inconsistent and Belief-Consistent Journal Articles. *Reading Psychology*, 44(6), 649–671. <https://doi.org/10.1080/02702711.2023.2179145>
- Braasch, J. L. G. & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) Model: Basic Assumptions and Preliminary Evidence. *Educational Psychologist*, 52(3), 167–181. <https://doi.org/10.1080/00461520.2017.1323219>
- Braasch, J. L. G., Bråten, I., Britt, M. A., Steffens, B. & Strømsø, H. I. (2014). Sensitivity to Inaccurate Argumentation in Health News Articles: Potential Contributions of Readers' Topic and Epistemic Beliefs. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 117–137). MIT Press.
- Braasch, J. L. G., Bråten, I. & McCrudden, M. T. (2018). Introduction to Research on Multiple Source Use. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 1–13). Routledge.
- Braasch, J. L. G., Bråten, I., Strømsø, H. I. & Anmarkrud, Ø. (2014). Incremental Theories of Intelligence Predict Multiple Document Comprehension. *Learning and Individual Differences*, 31, 11–20. <https://doi.org/10.1016/j.lindif.2013.12.012>
- Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø. & Ferguson, L. E. (2013). Promoting Secondary School Students' Evaluation of Source Features of Multiple Documents. *Contemporary Educational Psychology*, 38(3), 180–195. <https://doi.org/10.1016/j.cedpsych.2013.03.003>
- Braasch, J. L. G., Haverkamp, Y. E., Latini, N., Shaw, S., Arshad, M. S. & Bråten, I. (2022). Belief Bias when Adolescents Read to Comprehend Multiple Conflicting Texts. *Reading and Writing*, 35(8), 1759–1785. <https://doi.org/10.1007/s11145-022-10262-w>
- Braasch, J. L. G. & Kessler, E. D. (2021). Working toward a Theoretical Model for Source Comprehension in Everyday Discourse. *Discourse Processes*, 58(5–6), 449–467. <https://doi.org/10.1080/0163853X.2021.1905393>
- Braasch, J. L. G., Killion, S. C. & Bråten, I. (2021). Contextual Factors That Affect Adolescents' Detection of and Memory for Conflicts across Multiple Texts. *Journal of Research in Reading*, 44(2), 418–433. <https://doi.org/10.1111/1467-9817.12348>
- Braasch, J. L. G., Lawless, K. A., Goldman, S. R., Manning, F. H., Gomez, K. W. & Macleod, S. M. (2009). Evaluating Search Results: An Empirical Analysis of Middle School Students' Use of Source Attributes to Select Useful Sources. *Journal of Educational Computing Research*, 41(1), 63–82. <https://doi.org/10.2190/EC.41.1.c>
- Braasch, J. L. G., McCabe, R. M. & Daniel, F. (2016). Content Integration across Multiple Documents Reduces Memory for Sources. *Reading and Writing*, 29(8), 1571–1598. <https://doi.org/10.1007/s11145-015-9609-5>
- Braasch, J. L. G., Rouet, J.-F., Vibert, N. & Britt, M. A. (2012). Readers' Use of Source Information in Text Comprehension. *Memory & Cognition*, 40(3), 450–465. <https://doi.org/10.3758/s13421-011-0160-6>

- Braasch, J. L. G. & Scharrer, L. (2020). The Role of Cognitive Conflict in Understanding and Learning from Multiple Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 205–222). Routledge.
- Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L. & van Gog, T. (2017). Source Evaluation of Domain Experts and Novices during Web Search. *Journal of Computer Assisted Learning*, 33(3), 234–251. <https://doi.org/10.1111/jcal.12162>
- Brand-Gruwel, S., Wopereis, I. & Walraven, A. (2009). A Descriptive Model of Information Problem Solving while Using Internet. *Computers & Education*, 53(4), 1207–1217. <https://doi.org/10.1016/j.compedu.2009.06.004>
- Brante, E. W. & Strømsø, H. I. (2018). Sourcing in Text Comprehension: A Review of Interventions Targeting Sourcing Skills. *Educational Psychology Review*, 30(3), 773–799. <https://doi.org/10.1007/s10648-017-9421-7>
- Bråten, I., Anmarkrud, Ø., Brandmo, C. & Strømsø, H. I. (2014). Developing and Testing a Model of Direct and Indirect Relationships between Individual Differences, Processing, and Multiple-Text Comprehension. *Learning and Instruction*, 30, 9–24. <https://doi.org/10.1016/j.learninstruc.2013.11.002>
- Bråten, I. & Braasch, J. L. G. (2018). The Role of Conflict in Multiple Source Use. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 184–201). Routledge.
- Bråten, I., Braasch, J. L. G. & Salmerón, L. (2020). Reading Multiple and Non-Traditional Texts: New Opportunities and New Challenges. In E. B. Moje, P. Afflerbach, P. Enciso & N. K. Lesaux (Hrsg.), *Handbook of Reading Research: Volume V* (S. 79–98). Routledge.
- Bråten, I., Braasch, J. L. G., Strømsø, H. I. & Ferguson, L. E. (2015). Establishing Trustworthiness when Students Read Multiple Documents Containing Conflicting Scientific Evidence. *Reading Psychology*, 36(4), 315–349. <https://doi.org/10.1080/02702711.2013.864362>
- Bråten, I., Brandmo, C. & Kammerer, Y. (2019). A Validation Study of the Internet-Specific Epistemic Justification Inventory with Norwegian Pre-Service Teachers. *Journal of Educational Computing Research*, 57(4), 877–900.
- Bråten, I., Brante, E. W. & Strømsø, H. I. (2018). What Really Matters: The Role of Behavioural Engagement in Multiple Document Literacy Tasks. *Journal of Research in Reading*, 41(4), 680–699. <https://doi.org/10.1111/1467-9817.12247>
- Bråten, I., Brante, E. W. & Strømsø, H. I. (2019). Teaching Sourcing in Upper Secondary School: A Comprehensive Sourcing Intervention with Follow-Up Data. *Reading Research Quarterly*, 54(4), 481–505. <https://doi.org/10.1002/rrq.253>
- Bråten, I., Britt, M. A., Strømsø, H. I. & Rouet, J.-F. (2011). The Role of Epistemic Beliefs in the Comprehension of Multiple Expository Texts: Toward an Integrated Model. *Educational Psychologist*, 46(1), 48–70. <https://doi.org/10.1080/00461520.2011.538647>
- Bråten, I., Ferguson, L. E., Anmarkrud, Ø. & Strømsø, H. I. (2013). Prediction of Learning and Comprehension when Adolescents Read Multiple Texts: The Roles of Word-Level Processing, Strategic Approach, and Reading Motivation. *Reading and Writing*, 26(3), 321–348. <https://doi.org/10.1007/s11145-012-9371-x>
- Bråten, I., Ferguson, L. E., Strømsø, H. I. & Anmarkrud, Ø. (2013). Justification Beliefs and Multiple-Documents Comprehension. *European Journal of Psychology of Education*, 28(3), 879–902. <https://doi.org/10.1007/s10212-012-0145-2>
- Bråten, I., Ferguson, L. E., Strømsø, H. I. & Anmarkrud, Ø. (2014). Students Working with Multiple Conflicting Documents on a Scientific Issue: Relations between Epistemic Cognition while Reading and Sourcing and Argumentation in Essays. *British Journal of Educational Psychology*, 84(4), 58–85. <https://doi.org/10.1111/bjep.12005>
- Bråten, I., Gil, L. & Strømsø, H. I. (2011). The Role of Different Task Instructions and Reader Characteristics when Learning from Multiple Expository Texts. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 95–122). Information Age.
- Bråten, I., Latini, N. & Salmerón, L. (in Druck). Source Evaluation. In C. A. MacArthur, St. Graham & J. Fitzgerald (Hrsg.), *Handbook of Writing Research* (3. Aufl.). Guilford.

- Bråten, I., McCrudden, M. T., Stang Lund, E., Brante, E. W. & Strømsø, H. I. (2018). Task-Oriented Learning with Multiple Documents: Effects of Topic Familiarity, Author Expertise, and Content Relevance on Document Selection, Processing, and Use. *Reading Research Quarterly*, 53(3), 345–365. <https://doi.org/10.1002/rrq.197>
- Bråten, I., Salmerón, L. & Strømsø, H. I. (2016). Who Said That? Investigating the Plausibility-Induced Source Focusing Assumption with Norwegian Undergraduate Readers. *Contemporary Educational Psychology*, 46, 253–262. <https://doi.org/10.1016/j.cedpsych.2016.07.004>
- Bråten, I. & Samuelstuen, M. S. (2007). Measuring Strategic Processing: Comparing Task-Specific Self-Reports to Traces. *Metacognition and Learning*, 2(1), 1–20. <https://doi.org/10.1007/s11409-007-9004-y>
- Bråten, I., Stadler, M. & Salmerón, L. (2017). The Role of Sourcing in Discourse Comprehension. In M. F. Schober, D. N. Rapp & M. A. Britt (Hrsg.), *The Routledge Handbook of Discourse Processes* (2. Aufl., S. 141–168). Routledge.
- Bråten, I. & Strømsø, H. I. (2003). A Longitudinal Think-Aloud Study of Spontaneous Strategic Processing during the Reading of Multiple Expository Texts. *Reading and Writing*, 16(3), 195–218. <https://doi.org/10.1023/A:1022895207490>
- Bråten, I. & Strømsø, H. I. (2006a). Constructing Meaning from Multiple Information Sources as a Function of Personal Epistemology. *Information Design Journal*, 14(1), 56–67.
- Bråten, I. & Strømsø, H. I. (2006b). Effects of Personal Epistemology on the Understanding of Multiple Texts. *Reading Psychology*, 27(5), 457–484. <https://doi.org/10.1080/02702710600848031>
- Bråten, I. & Strømsø, H. I. (2010a). Effects of Task Instruction and Personal Epistemology on the Understanding of Multiple Texts About Climate Change. *Discourse Processes*, 47(1), 1–31. <https://doi.org/10.1080/01638530902959646>
- Bråten, I. & Strømsø, H. I. (2010b). When Law Students Read Multiple Documents about Global Warming: Examining the Role of Topic-Specific Beliefs about the Nature of Knowledge and Knowing. *Instructional Science*, 38(6), 635–657. <https://doi.org/10.1007/s11251-008-9091-4>
- Bråten, I. & Strømsø, H. I. (2011). Measuring Strategic Processing when Students Read Multiple Texts. *Metacognition and Learning*, 6(2), 111–130. <https://doi.org/10.1007/s11409-011-9075-7>
- Bråten, I. & Strømsø, H. I. (2020). On the Roles of Dispositions and Beliefs in Learning from Multiple Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 141–163). Routledge.
- Bråten, I., Strømsø, H. I. & Andreassen, R. (2016). Sourcing in Professional Education: Do Text Factors Make Any Difference? *Reading and Writing*, 29(8), 1599–1628. <https://doi.org/10.1007/s11145-015-9611-y>
- Bråten, I., Strømsø, H. I. & Britt, M. A. (2009). Trust Matters: Examining the Role of Source Evaluation in Students' Construction of Meaning within and across Multiple Texts. *Reading Research Quarterly*, 44(1), 6–28. <https://doi.org/10.1598/RRQ.44.1.1>
- Bråten, I., Strømsø, H. I. & Ferguson, L. E. (2016). The Role of Epistemic Beliefs in the Comprehension of Single and Multiple Texts. In P. Afflerbach (Hrsg.), *Handbook of Individual Differences in Reading: Reader, Text, and Context* (S. 67–79). Routledge.
- Bråten, I., Strømsø, H. I. & Salmerón, L. (2011). Trust and Mistrust when Students Read Multiple Information Sources about Climate Change. *Learning and Instruction*, 21(2), 180–192. <https://doi.org/10.1016/j.learninstruc.2010.02.002>
- Bråten, I., Strømsø, H. I. & Samuelstuen, M. S. (2008). Are Sophisticated Students Always Better? The Role of Topic-Specific Personal Epistemology in the Understanding of Multiple Expository Texts. *Contemporary Educational Psychology*, 33(4), 814–840. <https://doi.org/10.1016/j.cedpsych.2008.02.001>
- Breakstone, J., McGrew, S. & Smith, M. D. (2024). Measuring What Matters: Investigating What New Types of Assessments Reveal about Students' Online Source Evaluations. *Harvard Kennedy School Misinformation Review*, 5(1), 1–15. <https://doi.org/10.37016/mr-2020-133>
- Breakstone, J., McGrew, S., Smith, M. D., Ortega, T. & Wineburg, S. S. (2018). Why We Need a New Approach to Teaching Digital Literacy. *Phi Delta Kappan*, 99(6), 27–32. <https://doi.org/10.1177/0031721718762419>

- Breakstone, J., Smith, M. D., Connors, P., Ortega, T., Kerr, D. & Wineburg, S. S. (2021). Lateral Reading: College Students Learn to Critically Evaluate Internet Sources in an Online Course. *Harvard Kennedy School Misinformation Review*, 2(1), 1–17. <https://doi.org/10.37016/mr-2020-56>
- Breakstone, J., Smith, M. D., Wineburg, S. S., Rapaport, A., Carle, J., Garland, M. & Saavedra, A. (2021). Students' Civic Online Reasoning: A National Portrait. *Educational Researcher*, 50(8), 505–515. <https://doi.org/10.3102/0013189X211017495>
- Breakstone, J., Smith, M. D., Ziv, N. & Wineburg, S. S. (2022). Civic Preparation for the Digital Age: How College Students Evaluate Online Sources about Social and Political Issues. *The Journal of Higher Education*, 93(7), 963–988. <https://doi.org/10.1080/00221546.2022.2082783>
- Britt, M. A. & Aglinskias, C. (2002). Improving Students' Ability to Identify and Use Source Information. *Cognition and Instruction*, 20(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2
- Britt, M. A., Durik, A. & Rouet, J.-F. (2022). Reading Contexts, Goals, and Decisions: Text Comprehension as a Situated Activity. *Discourse Processes*, 59(5–6), 361–378. <https://doi.org/10.1080/0163853X.2022.2068345>
- Britt, M. A., Perfetti, C. A., Sandak, R. & Rouet, J.-F. (1999). Content Integration and Source Separation in Learning from Multiple Texts. In S. R. Goldman, A. C. Graesser & P. W. van den Broek (Hrsg.), *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (S. 209–233). L. Erlbaum Associates.
- Britt, M. A., Richter, T. & Rouet, J.-F. (2014). Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist*, 49(2), 104–122. <https://doi.org/10.1080/00461520.2014.916217>
- Britt, M. A. & Rouet, J.-F. (2012). Learning with Multiple Documents: Component Skills and Their Acquisition. In J. R. Kirby & M. J. Lawson (Hrsg.), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes* (S. 276–314). Cambridge University Press. <https://doi.org/10.1017/CBO9781139048224.017>
- Britt, M. A. & Rouet, J.-F. (2020). *Multiple Document Comprehension* [Oxford Research Encyclopedia of Education]. <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-867>
- Britt, M. A., Rouet, J.-F. & Braasch, J. L. G. (2013). Documents as Entities: Extending the Situation Model Theory of Comprehension. In M. A. Britt, S. R. Goldman & J.-F. Rouet (Hrsg.), *Reading: From Words to Multiple Texts* (S. 160–179). Routledge.
- Britt, M. A., Rouet, J.-F. & Durik, A. M. (2018). *Literacy beyond Text Comprehension: A Theory of Purposeful Reading*. Routledge. <https://doi.org/10.4324/9781315682860>
- Britt, M. A., Rouet, J.-F., Georgi, M. C. & Perfetti, C. A. (1994). Learning from History Texts: From Causal Analysis to Argument Models. In G. Leinhardt, I. L. Beck & C. Stainton (Hrsg.), *Teaching and Learning in History* (S. 47–84). Erlbaum.
- Britt, M. A. & Sommer, J. (2004). Facilitating Textual Integration with Macro-Structure Focusing Tasks. *Reading Psychology*, 25(4), 313–339. <https://doi.org/10.1080/02702710490522658>
- Britton, B. K., Glynn, S. M., Muth, K. D. & Penland, M. J. (1985). Instructional Objectives in Text: Managing the Reader's Attention. *Journal of Reading Behavior*, 17(2), 101–113. <https://doi.org/10.1080/10862968509547533>
- Broadbent, J. & Poon, W. L. (2015). Self-Regulated Learning Strategies & Academic Achievement in Online Higher Education Learning Environments: A Systematic Review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Brodsky, J. E., Brooks, P. J., Pavlounis, D. & Johnston, J. L. (2023). Instruction Increases Canadian Students' Preference for and Use of Lateral Reading Strategies to Fact-Check Online Information. *AERA Open*, 9(1), 1–10. <https://doi.org/10.1177/23328584231192106>
- Brodsky, J. E., Brooks, P. J., Scimeca, D., Galati, P., Todorova, R. & Caulfield, M. (2021). Associations between Online Instruction in Lateral Reading Strategies and Fact-Checking COVID-19 News Among College Students. *AERA Open*, 7(1), 1–17. <https://doi.org/10.1177/23328584211038937>
- Brodsky, J. E., Brooks, P. J., Scimeca, D., Todorova, R., Galati, P., Batson, M., Grosso, R., Matthews, M., Miller, V. & Caulfield, M. (2021). Improving College Students' Fact-Checking Strategies through Lateral Reading Instruction in a General Education Civics Course. *Cognitive Research: Principles and Implications*, 6, Artikel 23, 1–18. <https://doi.org/10.1186/s41235-021-00291-4>

- Broekkamp, H., van den Bergh, H., van Hout-Wolters, B. H. A. M. & Rijlaarsdam, G. (2002). Will That Be on the Test? Perceived Task Demands and Test Performance in a Classroom Context. *European Journal of Psychology of Education*, 17(1), 75–92. <https://doi.org/10.1007/BF03173206>
- Broekkamp, H. & van Hout-Wolters, B. H. A. M. (2007). Students' Adaptation of Study Strategies when Preparing for Classroom Tests. *Educational Psychology Review*, 19(4), 401–428. <https://doi.org/10.1007/s10648-006-9025-0>
- Broekkamp, H., van Hout-Wolters, B. H. A. M., Rijlaarsdam, G. & van den Bergh, H. (2002). Importance in Instructional Text: Teachers' and Students' Perceptions of Task Demands. *Journal of Educational Psychology*, 94(2), 260–271. <https://doi.org/10.1037/0022-0663.94.2.260>
- Bromme, R. & Goldman, S. R. (2014). The Public's Bounded Understanding of Science. *Educational Psychologist*, 49(2), 59–69. <https://doi.org/10.1080/00461520.2014.921572>
- Bromme, R., Kienhues, D. & Porsch, T. (2010). Who Knows What and Who Can We Believe? Epistemological Beliefs Are Beliefs about Knowledge (Mostly) to Be Attained from Others. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice* (S. 163–194). Cambridge University Press. <https://doi.org/10.1017/CBO9780511691904.006>
- Bromme, R., Kienhues, D. & Stahl, E. (2008). Knowledge and Epistemological Beliefs: An Intimate but Complicate Relationship. In M. S. Khine (Hrsg.), *Knowing, Knowledge and Beliefs: Epistemological Studies across Diverse Cultures* (S. 423–441). Springer. https://doi.org/10.1007/978-1-4020-6596-5_20
- Bromme, R., Pieschl, S. & Stahl, E. (2010). Epistemological Beliefs Are Standards for Adaptive Learning: A Functional Theory about Epistemological Beliefs and Metacognition. *Metacognition and Learning*, 5(1), 7–26. <https://doi.org/10.1007/s11409-009-9053-5>
- Bromme, R., Pieschl, S. & Stahl, E. (2014). Epistemological Beliefs and Students' Adaptive Perception of Task Complexity. In S. Krolak-Schwerdt, S. Glock & M. Böhmer (Hrsg.), *Teachers' Professional Development: Assessment, Training, and Learning* (S. 123–151). SensePublishers. https://doi.org/10.1007/978-94-6209-536-6_8
- Bromme, R., Stadler, M. & Scharer, L. (2018). The Provenance of Certainty: Multiple Source Use and the Public Engagement with Science. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 269–284). Routledge.
- Buehl, M. M. (2008). Assessing the Multidimensionality of Students' Epistemic Beliefs across Diverse Cultures. In M. S. Khine (Hrsg.), *Knowing, Knowledge and Beliefs: Epistemological Studies across Diverse Cultures* (S. 65–112). Springer. https://doi.org/10.1007/978-1-4020-6596-5_4
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M. & Finkel, E. J. (2013). Mind-Sets Matter: A Meta-Analytic Review of Implicit Theories and Self-Regulation. *Psychological Bulletin*, 139(3), 655–701. <https://doi.org/10.1037/a0029531>
- Burton, C. & Daneman, M. (2007). Compensating for a Limited Working Memory Capacity during Reading: Evidence from Eye Movements. *Reading Psychology*, 28(2), 163–186. <https://doi.org/10.1080/02702710601186407>
- Butcher, K. R. & Davies, S. (2015). Inference Generation during Online Study and Multimedia Learning. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Hrsg.), *Inferences during Reading* (S. 321–347). Cambridge University Press.
- Butler, D. L. (2021). Enabling Educators to Become More Effective Supporters of SRL: Commentary on a Special Issue. *Metacognition and Learning*, 16(3), 667–684. <https://doi.org/10.1007/s11409-021-09282-8>
- Butler, D. L. & Cartier, S. C. (2004). Promoting Effective Task Interpretation as an Important Work Habit: A Key to Successful Teaching and Learning. *Teachers College Record*, 106(9), 1729–1758. <https://doi.org/10.1111/j.1467-9620.2004.00403.x>
- Butler, D. L. & Cartier, S. C. (2018). Advancing Research and Practice about Self-Regulated Learning: The Promise of In-Depth Case Study Methodologies. In D. H. Schunk & Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 352–369). Routledge.
- Butler, D. L., Cartier, S. C., Schnellert, L., Gagnon, F. & Giammarino, M. (2011). Secondary Students' Self-Regulated Engagement in Reading: Researching Self-Regulation as Situated in Context. *Psychological Test and Assessment Modeling*, 53(1), 73–105.

- Butler, D. L. & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Butterfuss, R. & Kendeou, P. (2018). The Role of Executive Functions in Reading Comprehension. *Educational Psychology Review*, 30(3), 801–826. <https://doi.org/10.1007/s10648-017-9422-6>
- Butterfuss, R. & Kendeou, P. (2021). KReC-MD: Knowledge Revision with Multiple Documents. *Educational Psychology Review*, 33(4), 1475–1497. <https://doi.org/10.1007/s10648-021-09603-y>
- Butterfuss, R., Kim, J. & Kendeou, P. (2020). *Reading Comprehension*. Oxford Research Encyclopedias, Education.
- Buyuktas Kara, M., van Steendam, E., Rijlaarsdam, G. & Kuru, H. (2018). The Effect of Two Modes of Strategy Instruction: Modeling vs. Presentational. *International Online Journal of Education and Teaching*, 5(2), 460–495.
- Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H. & Oades, L. G. (2021). Activity Achievement Emotions and Academic Performance: A Meta-Analysis. *Educational Psychology Review*, 33(3), 1051–1095. <https://doi.org/10.1007/s10648-020-09585-3>
- Camp, G., Surma, T. & Kirschner, P. A. (2022). Foundations of Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 17–24). Cambridge University Press. <https://doi.org/10.1017/9781108894333.004>
- Carney, R. N. & Levin, J. R. (2002). Pictorial Illustrations Still Improve Students' Learning from Text. *Educational Psychology Review*, 14(1), 5–26. <https://doi.org/10.1023/A:1013176309260>
- Carretti, B., Borella, E., Cornoldi, C. & Beni, R. (2009). Role of Working Memory in Explaining the Performance of Individuals with Specific Reading Comprehension Difficulties: A Meta-Analysis. *Learning and Individual Differences*, 19(2), 246–251. <https://doi.org/10.1016/j.lindif.2008.10.002>
- Cartiff, B. M. & Greene, J. A. (2020). Using Critical Thinking Frameworks to Understand Integration of Multiple Inputs. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 425–440). Routledge.
- Cartwright, K. B. (2009). The Role of Cognitive Flexibility in Reading Comprehension: Past, Present, and Future. In S. E. Israel & G. G. Duffy (Hrsg.), *Handbook of Research on Reading Comprehension* (S. 115–139). Routledge.
- Casado-Ledesma, L., Cuevas, I., van den Bergh, H., Rijlaarsdam, G., Mateos, M., Granado-Peinado, M. & Martín, E. (2021). Teaching Argumentative Synthesis Writing through Deliberative Dialogues: Instructional Practices in Secondary Education. *Instructional Science*, 49(4), 515–559. <https://doi.org/10.1007/s11251-021-09548-3>
- Casado-Ledesma, L. & Tarchi, C. (2024). The Influence of Textual Genre in Multiple-Text Comprehension. *Learning and Instruction*, 93, Artikel 101947, 1–13. <https://doi.org/10.1016/j.learninstruc.2024.101947>
- Castells, N., Minguela, M. & Nadal, E. (2023). Writing a Synthesis versus Reading: Strategies Involved and Impact on Comprehension. *Reading and Writing*, 36(4), 849–880. <https://doi.org/10.1007/s11145-022-10341-y>
- Castells, N., Minguela, M., Solé, I., Miras, M., Nadal, E. & Rijlaarsdam, G. (2022). Improving Questioning – Answering Strategies in Learning from Multiple Complementary Texts: An Intervention Study. *Reading Research Quarterly*, 57(3), 879–912. <https://doi.org/10.1002/rrq.451>
- Castro-Alonso, J. C., Koning, B. B. de, Fiorella, L. & Paas, F. G. W. C. (2021). Five Strategies for Optimizing Instructional Materials: Instructor- and Learner-Managed Cognitive Load. *Educational Psychology Review*, 33(4), 1379–1407. <https://doi.org/10.1007/s10648-021-09606-9>
- Catrysse, L., Gijbels, D., Donche, V., Maeyer, S. de, Lesterhuis, M. & van den Bossche, P. (2018). How Are Learning Strategies Reflected in the Eyes? Combining Results from Self-Reports and Eye-Tracking. *British Journal of Educational Psychology*, 88(1), 118–137. <https://doi.org/10.1111/bjep.12181>
- Cerdán, R., Candel, C. & Leppink, J. (2018). Cognitive Load and Learning in the Study of Multiple Documents. *Frontiers in Education*, 3, Artikel 59, 1–7. <https://doi.org/10.3389/feduc.2018.00059>
- Cerdán, R., Gil, L. & Vidal-Abarca, E. (2011). Question-Driven Processing in Single and Multiple Texts. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 295–319). Information Age.

- Cerdán, R., Gilabert, R. & Vidal-Abarca, E. (2011). Selecting Information to Answer Questions: Strategic Individual Differences when Searching Texts. *Learning and Individual Differences*, 21(2), 201–205. <https://doi.org/10.1016/j.lindif.2010.11.007>
- Cerdán, R., Gilabert, R. & Vidal-Abarca, E. (2013). Self-Generated Explanations on the Question Demands Are Not Always Helpful. *The Spanish Journal of Psychology*, 16, 1–11. <https://doi.org/10.1017/sjp.2013.45>
- Cerdán, R., Máñez, I. & Serrano-Mendizábal, M. (2021). Reading from Multiple Documents: The Role of Text Availability and Question Type. *Reading Research Quarterly*, 56(S1), S209–S220. <https://doi.org/10.1002/rrq.380>
- Cerdán, R., Marín, M. C. & Candel, C. (2013). The Role of Perspective on Students' Use of Multiple Documents to Solve an Open-Ended Task. *Psicología Educativa*, 19(2), 89–94. [https://doi.org/10.1016/S1135-755X\(13\)70015-0](https://doi.org/10.1016/S1135-755X(13)70015-0)
- Cerdán, R., Martínez, T., Vidal-Abarca, E., Gilabert, R., Gil, L. & Rouet, J.-F. (2008). Search and Comprehension Processes in Learning from Text. In J.-F. Rouet, R. Lowe & W. Schnotz (Hrsg.), *Understanding Multimedia Documents* (S. 121–147). Springer US. https://doi.org/10.1007/978-0-387-73337-1_7
- Cerdán, R., Pérez, A., Vidal-Abarca, E. & Rouet, J.-F. (2019). To Answer Questions from Text, One Has to Understand What the Question Is Asking: Differential Effects of Question Aids as a Function of Comprehension Skill. *Reading and Writing*, 32(8), 2111–2124. <https://doi.org/10.1007/s11145-019-09943-w>
- Cerdán, R. & Vidal-Abarca, E. (2008). The Effects of Tasks on Integrating Information from Multiple Documents. *Journal of Educational Psychology*, 100(1), 209–222. <https://doi.org/10.1037/0022-0663.100.1.209>
- Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R. & Gil, L. (2009). Impact of Question-Answering Tasks on Search Processes and Reading Comprehension. *Learning and Instruction*, 19(1), 13–27. <https://doi.org/10.1016/j.learninstruc.2007.12.003>
- Chang, S. (2023). Structural Relations between Processing Variables and Text Quality while Reading and Writing from Multiple Documents. *Korean Language Education Research*, 58(5), 5–43. <https://doi.org/10.20880/kler.2023.58.5.5>
- Chau, L. T., Leijten, M., Bernolet, S. & Vangehuchten, L. (2022). Envisioning Multilingualism in Source-Based Writing in L1, L2, and L3: The Relation between Source Use and Text Quality. *Frontiers in Psychology*, 13, Artikel 914125, 1–20. <https://doi.org/10.3389/fpsyg.2022.914125>
- Chen, Y.-C. (2019). Writing as an Epistemological Tool: Perspectives from Personal, Disciplinary, and Sociocultural Landscapes. In V. Prain & B. Hand (Hrsg.), *Theorizing the Future of Science Education Research* (S. 115–132). Springer International Publishing. https://doi.org/10.1007/978-3-030-24013-4_8
- Chen, J. (2022). The Effectiveness of Self-Regulated Learning (SRL) Interventions on L2 Learning Achievement, Strategy Employment and Self-Efficacy: A Meta-Analytic Study. *Frontiers in Psychology*, 13, Artikel 1021101, 1–17. <https://doi.org/10.3389/fpsyg.2022.1021101>
- Cheong, C. M., Liu, Y. & Mu, R. (2024). Characteristics of Students' Task Representation and Its Association with Argumentative Integrated Writing performance. *Assessing Writing*, 60, Artikel 100845, 1–16. <https://doi.org/10.1016/j.asw.2024.100845>
- Cheong, C. M., Zhu, X., Li, G. Y. & Wen, H. (2019). Effects of Intertextual Processing on L2 Integrated Writing. *Journal of Second Language Writing*, 44, 63–75. <https://doi.org/10.1016/j.jslw.2019.03.004>
- Cheong, C. M., Zhu, X. & Liao, X. (2018). Differences between the Relationship of L1 Learners' Performance in Integrated Writing with Both Independent Listening and Independent Reading Cognitive Skills. *Reading and Writing*, 31(4), 779–811. <https://doi.org/10.1007/s11145-017-9811-8>
- Cheong, C. M., Zhu, X. & Liu, Y. (2022). The Role of Need for Cognition (NfC) in the Effect of Language Modalities on Integrated Writing Performance. *British Journal of Educational Psychology*, 92(4), 1256–1277. <https://doi.org/10.1111/bjep.12498>
- Chevrier, M., Muis, K. R. & Di Leo, I. (2020). Calibration to Task Complexity: The Role of Epistemic Cognition. *Journal of Experimental Education*, 88(1), 1–26. <https://doi.org/10.1080/00220973.2019.1584740>

- Chinn, C. A., Barzilai, S. & Duncan, R. G. (2020). Disagreeing about How to Know: The Instructional Value of Explorations into Knowing. *Educational Psychologist*, 55(3), 167–180. <https://doi.org/10.1080/00461520.2020.1786387>
- Chinn, C. A., Barzilai, S. & Duncan, R. G. (2021). Education for a „Post-Truth“ World: New Directions for Research and Practice. *Educational Researcher*, 50(1), 51–60. <https://doi.org/10.3102/0013189X20940683>
- Chinn, C. A., Buckland, L. A. & Samarapungavan, A. (2011). Expanding the Dimensions of Epistemic Cognition: Arguments from Philosophy and Psychology. *Educational Psychologist*, 46(3), 141–167. <https://doi.org/10.1080/00461520.2011.587722>
- Chinn, C. A. & Rinehart, R. W. (2016a). Commentary: Advances in Research on Sourcing – Source Credibility and Reliable Processes for Producing Knowledge Claims. *Reading and Writing*, 29(8), 1701–1717. <https://doi.org/10.1007/s11145-016-9675-3>
- Chinn, C. A. & Rinehart, R. W. (2016b). Epistemic Cognition and Philosophy: Developing a New Framework for Epistemic Cognition. In Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of Epistemic Cognition* (S. 460–478). Routledge.
- Chinn, C. A., Rinehart, R. W. & Buckland, L. A. (2014). Epistemic Cognition and Evaluating Information: Applying the AIR Model of Epistemic Cognition. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 425–453). MIT Press.
- Cho, B.-Y. (2014). Competent Adolescent Readers' Use of Internet Reading Strategies: A Think-Aloud Study. *Cognition and Instruction*, 32(3), 253–289. <https://doi.org/10.1080/07370008.2014.918133>
- Cho, B.-Y. & Afflerbach, P. (2017). An Evolving Perspective of Constructively Responsive Reading Comprehension Strategies in Multilayered Digital Text Environments. In S. E. Israel (Hrsg.), *Handbook of Research on Reading Comprehension* (2. Aufl., S. 109–134). Routledge.
- Cho, B.-Y., Afflerbach, P. & Han, H. (2018). Strategic Processing in Accessing, Comprehending, and Using Multiple Sources Online. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 133–150). Routledge.
- Cho, B.-Y., Han, H. & Kucan, L. L. (2018). An Exploratory Study of Middle-School Learners' Historical Reading in an Internet Environment. *Reading and Writing*, 31(7), 1525–1549. <https://doi.org/10.1007/s11145-018-9847-4>
- Cho, B.-Y., Woodward, L. & Li, D. (2018). Epistemic Processing when Adolescents Read Online: A Verbal Protocol Analysis of More and Less Successful Online Readers. *Reading Research Quarterly*, 53(2), 197–221. <https://doi.org/10.1002/rrq.190>
- Cho, B.-Y., Woodward, L., Li, D. & Barlow, W. (2017). Examining Adolescents' Strategic Processing during Online Reading with a Question-Generating Task. *American Educational Research Journal*, 54(4), 691–724. <https://doi.org/10.3102/0002831217701694>
- Choi, W. & Stvilia, B. (2015). Web Credibility Assessment: Conceptualization, Operationalization, Variability, and Models. *Journal of the Association for Information Science and Technology*, 66(12), 2399–2414. <https://doi.org/10.1002/asi.23543>
- Chun, M. M., Golomb, J. D. & Turk-Browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, 62, 73–101. <https://doi.org/10.1146/annurev.psych.093008.100427>
- Cleary, T. J. & Callan, G. L. (2018). Assessing Self-Regulated Learning Using Microanalytic Methods. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 338–351). Routledge.
- Cleary, T. J., Callan, G. L. & Zimmerman, B. J. (2012). Assessing Self-Regulation as a Cyclical, Context-Specific Phenomenon: Overview and Analysis of SRL Microanalytic Protocols. *Education Research International*, Artikel 428639, 1–19. <https://doi.org/10.1155/2012/428639>
- Cleary, T. J., Slemp, J., Reddy, L. A., Alperin, A., Lui, A., Austin, A. & Cedar, T. (2023). Characteristics and Uses of SRL Microanalysis across Diverse Contexts, Tasks, and Populations: A Systematic Review. *School Psychology Review*, 52(2), 159–179. <https://doi.org/10.1080/2372966X.2020.1862627>
- Clinton, V. (2019). Reading from Paper Compared to Screens: A Systematic Review and Meta-Analysis. *Journal of Research in Reading*, 42(2), 288–325. <https://doi.org/10.1111/1467-9817.12269>

- Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E. & Seipel, B. (2020). Inferential Comprehension Differences between Narrative and Expository Texts: A Systematic Review and Meta-Analysis. *Reading and Writing*, 31(5–6), 2223–2248. <https://doi.org/10.1007/s11145-020-10044-2>
- Clinton-Lisell, V. (2021). Stop Multitasking and Just Read: Meta-Analyses of Multitasking's Effects on Reading Performance and Reading Time. *Journal of Research in Reading*, 44(4), 787–816. <https://doi.org/10.1111/1467-9817.12372>
- Coiro, J. (2015). Purposeful, Critical, and Flexible: Vital Dimensions of Online Reading and Learning. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink & P. Thompson (Hrsg.), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices* (S. 53–64). Routledge.
- Coiro, J. (2021). Toward a Multifaceted Heuristic of Digital Reading to Inform Assessment, Research, Practice, and Policy. *Reading Research Quarterly*, 56(1), 9–31. <https://doi.org/10.1002/rrq.302>
- Coiro, J., Coscarelli, C., Maykel, C. & Forzani, E. (2015). Investigating Criteria That Seventh Graders Use to Evaluate the Quality of Online Information. *Journal of Adolescent & Adult Literacy*, 59(3), 287–297. <https://doi.org/10.1002/jaal.448>
- Collins, P., Tate, T. P., won Lee, J., Krishnan, J. A. & Warschauer, M. (2021). A Multi-Dimensional Examination of Adolescent Writing: Considering the Writer, Genre and Task Demands. *Reading and Writing*, 34(8), 2151–2173. <https://doi.org/10.1007/s11145-021-10140-x>
- Connell, L. & Keane, M. T. (2006). A Model of Plausibility. *Cognitive Science*, 30(1), 95–120. https://doi.org/10.1207/s15516709cog0000_53
- Conradi, K., Jang, B. G. & McKenna, M. C. (2014). Motivation Terminology in Reading Research: A Conceptual Review. *Educational Psychology Review*, 26(1), 127–164. <https://doi.org/10.1007/s10648-013-9245-z>
- Cook, A. E. & O'Brien, E. J. (2017). Fundamentals of Inferencing during Reading. *Language and Linguistics Compass*, 11(7), 1–16. <https://doi.org/10.1111/lnc3.12246>
- Cook, D. A. & Artino, A. R. (2016). Motivation to Learn: An Overview of Contemporary Theories. *Medical Education*, 50(10), 997–1014. <https://doi.org/10.1111/medu.13074>
- Cortés Pascual, A., Moyano Muñoz, N. & Quílez Robres, A. (2019). The Relationship between Executive Functions and Academic Performance in Primary Education: Review and Meta-Analysis. *Frontiers in Psychology*, 10, 1582. <https://doi.org/10.3389/fpsyg.2019.01582>
- Costa, A. & Faria, L. (2018). Implicit Theories of Intelligence and Academic Achievement: A Meta-Analytic Review. *Frontiers in Psychology*, 9, Artikel 829, 1–16. <https://doi.org/10.3389/fpsyg.2018.00829>
- Coté, N., Goldman, S. R. & Saul, E. U. (1998). Students Making Sense of Informational Text: Relations between Processing and Representation. *Discourse Processes*, 25(1), 1–53. <https://doi.org/10.1080/01638539809545019>
- Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, 19(1), 51–57. <https://doi.org/10.1177/0963721409359277>
- Cowan, N. (2017). The Many Faces of Working Memory and Short-Term Storage. *Psychonomic Bulletin & Review*, 24(4), 1158–1170. <https://doi.org/10.3758/s13423-016-1191-6>
- Craig, K., Hale, D., Grainger, C. & Stewart, M. E. (2020). Evaluating Metacognitive Self-Reports: Systematic Reviews of the Value of Self-Report in Metacognitive Research. *Metacognition and Learning*, 15(3), 155–213. <https://doi.org/10.1007/s11409-020-09222-y>
- Craik, F. I. M. (2002). Levels of Processing: Past, Present ... and Future? *Memory*, 10(5-6), 305–318. <https://doi.org/10.1080/09658210244000135>
- Credé, M. & Phillips, L. A. (2011). A Meta-Analytic Review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, 21(4), 337–346. <https://doi.org/10.1016/j.lindif.2011.03.002>
- Cromley, J. G. & Azevedo, R. (2006). Self-Report of Reading Comprehension Strategies: What Are We Measuring? *Metacognition and Learning*, 1(3), 229–247. <https://doi.org/10.1007/s11409-006-9002-5>

- Cromley, J. G. & Wills, T. W. (2016). Flexible Strategy Use by Students Who Learn Much versus Little from Text: Transitions within Think-Aloud Protocols. *Journal of Research in Reading*, 39(1), 50–71. <https://doi.org/10.1111/1467-9817.12026>
- Cuevas, I., Mateos, M., Martín, E., Luna, M., Martín, A., Solari, M., González-Lamas, J. & Martínez, I. (2016). Collaborative Writing of an Argumentative Synthesis from Multiple Sources: The Role of Writing Beliefs and Strategies to Deal with Controversy. *Journal of Writing Research*, 8(2), 205–226. <https://doi.org/10.17239/jowr-2016.08.02.02>
- Cumming, A., Lai, C. & Cho, H. (2016). Students' Writing from Sources for Academic Purposes: A Synthesis of Recent Research. *Journal of English for Academic Purposes*, 23, 47–58. <https://doi.org/10.1016/j.jeap.2016.06.002>
- D'Mello, S. K. (2018). What Do We Think about When We Learn? In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Hrsg.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (S. 52–67). Routledge.
- D'Mello, S. K. & Mills, C. S. (2021). Mind Wandering during Reading: An Interdisciplinary and Integrative Review of Psychological, Computing, and Intervention Research and Theory. *Language and Linguistics Compass*, 15(4), Artikel e12412, 1–32. <https://doi.org/10.1111/lnc3.12412>
- Daher, T. A. & Kiewra, K. A. (2016). An Investigation of SOAR Study Strategies for Learning from Multiple Online Resources. *Contemporary Educational Psychology*, 46, 10–21. <https://doi.org/10.1016/j.cedpsych.2015.12.004>
- Daneman, M. & Hannon, B. (2007). What Do Working Memory Span Tasks Like Reading Span Really Measure? In N. Osaka, R. H. Logie, R. Logie & M. D'Esposito (Hrsg.), *The Cognitive Neuroscience of Working Memory* (S. 21–42). Oxford University Press.
- Daneman, M. & Merikle, P. M. (1996). Working Memory and Language Comprehension: A Meta-Analysis. *Psychonomic Bulletin & Review*, 3(4), 422–433. <https://doi.org/10.3758/BF03214546>
- Danışman, Ş. (2017). The Effect of Self-Regulation on Student Achievement. In E. Karadag (Hrsg.), *The Factors Effecting Student Achievement: Meta-Analysis of Empirical Studies* (S. 133–159). Springer International Publishing. https://doi.org/10.1007/978-3-319-56083-0_8
- Davis, D. S., Huang, B. & Yi, T. (2017). Making Sense of Science Texts: A Mixed-Methods Examination of Predictors and Processes of Multiple-Text Comprehension. *Reading Research Quarterly*, 52(2), 227–252. <https://doi.org/10.1002/rrq.162>
- de Bruin, A. B. H., Roelle, J., Carpenter, S. K., Baars, M. & EFG-MRE (2020). Synthesizing Cognitive Load and Self-Regulation Theory: A Theoretical Framework and Research Agenda. *Educational Psychology Review*, 32(4), 903–915. <https://doi.org/10.1007/s10648-020-09576-4>
- de Bruin, A. B. H. & van Gog, T. (2012). Improving Self-Monitoring and Self-Regulation: From Cognitive Psychology to the Classroom. *Learning and Instruction*, 22(4), 245–252. <https://doi.org/10.1016/j.learninstruc.2012.01.003>
- de Bruin, A. B. H. & van Merriënboer, J. J. G. (2017). Bridging Cognitive Load and Self-Regulated Learning Research: A Complementary Approach to Contemporary Issues in Educational Research. *Learning and Instruction*, 51, 1–9. <https://doi.org/10.1016/j.learninstruc.2017.06.001>
- De La Paz, S. & Wissinger, D. R. (2015). Effects of Genre and Content Knowledge on Historical Thinking with Academically Diverse High School Students. *The Journal of Experimental Education*, 83(1), 110–129. <https://doi.org/10.1080/00220973.2013.876228>
- de-la-Peña, C., Chaves-Yuste, B. & Luque-Rojas, M. J. (2024). Bridging the Printed or Digital Controversy: A Meta-Analysis of the Impact of Digital and Print Resources on College Students' Reading Comprehension. *Journal of College Reading and Learning*, 54(2), 79–102. <https://doi.org/10.1080/10790195.2024.2350949>
- Deekens, V. M., Greene, J. A. & Lobczowski, N. G. (2018). Monitoring and Depth of Strategy Use in Computer-Based Learning Environments for Science and History. *British Journal of Educational Psychology*, 88(1), 63–79. <https://doi.org/10.1111/bjep.12174>
- DeLeeuw, K. E. & Mayer, R. E. (2008). A Comparison of Three Measures of Cognitive Load: Evidence for Separable Measures of Intrinsic, Extraneous, and Germane Load. *Journal of Educational Psychology*, 100(1), 223–234. <https://doi.org/10.1037/0022-0663.100.1.223>
- Delgado, P., Ávila, V., Fajardo, I. & Salmerón, L. (2019). Training Young Adults with Intellectual Disability to Read Critically on the Internet. *Journal of Applied Research in Intellectual Disabilities*, 32(3), 666–677. <https://doi.org/10.1111/jar.12562>

- Delgado, P., Stang Lund, E., Salmerón, L. & Bråten, I. (2020). To Click or Not to Click: Investigating Conflict Detection and Sourcing in a Multiple Document Hypertext Environment. *Reading and Writing*, 33(8), 2049–2072. <https://doi.org/10.1007/s11145-020-10030-8>
- Delgado, P., Vargas, C., Ackerman, R. & Salmerón, L. (2018). Don't Throw Away Your Printed Books: A Meta-Analysis on the Effects of Reading Media on Reading Comprehension. *Educational Research Review*, 25, 23–38. <https://doi.org/10.1016/j.edurev.2018.09.003>
- Demir, B., Haverkamp, Y. E., Braasch, J. L. & Bråten, I. (2024). Investigating the Role of Prior Knowledge in Comprehending Intratextual and Intertextual Relationships when Reading Multiple Texts. *Learning and Individual Differences*, 111, Artikel 102442, 1–13. <https://doi.org/10.1016/j.lindif.2024.102442>
- Dent, A. L. & Hoyle, R. H. (2015). A Framework for Evaluating and Enhancing Alignment in Self-Regulated Learning Research. *Metacognition and Learning*, 10(1), 165–179. <https://doi.org/10.1007/s11409-015-9136-4>
- Dent, A. L. & Koenka, A. C. (2016). The Relation between Self-Regulated Learning and Academic Achievement across Childhood and Adolescence: A Meta-Analysis. *Educational Psychology Review*, 28(3), 425–474. <https://doi.org/10.1007/s10648-015-9320-8>
- Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. M. & Carter, S. (2015). Text-Processing Differences in Adolescent Adequate and Poor Comprehenders Reading Accessible and Challenging Narrative and Informational Text. *Reading Research Quarterly*, 50(4), 393–416. <https://doi.org/10.1002/rrq.105>
- Denton, C. A., Muis, K. R., Dubé, A. K. & Armstrong, S. (2022). En-Garde: Source Evaluations in the Digital Age. *Advances in Social Sciences Research Journal*, 9(9), 320–360. <https://doi.org/10.14738/assrj.99.13066>
- Deroover, K., Knight, S., Burke, P. F. & Bucher, T. (2022). Why Do Experts Disagree? The Development of a Taxonomy. *Public Understanding of Science*, 32(2), 224–246. <https://doi.org/10.1177/09636625221110029>
- DeStefano, D. & LeFevre, J.-A. (2007). Cognitive Load in Hypertext Reading: A Review. *Computers in Human Behavior*, 23(3), 1616–1641. <https://doi.org/10.1016/j.chb.2005.08.012>
- Di Vesta, F. J. & Di Cintio, M. J. (1997). Interactive Effects of Working Memory Span and Text Context on Reading Comprehension and Retrieval. *Learning and Individual Differences*, 9(3), 215–231. [https://doi.org/10.1016/S1041-6080\(97\)90007-8](https://doi.org/10.1016/S1041-6080(97)90007-8)
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- DiBenedetto, M. K. & Zimmerman, B. J. (2010). Differences in Self-Regulatory Processes among Students Studying Science: A Microanalytic Investigation. *The International Journal of Educational and Psychological Assessment*, 5, 2–24.
- DiBenedetto, M. K. & Zimmerman, B. J. (2013). Construct and Predictive Validity of Microanalytic Measures of Students' Self-Regulation of Science Learning. *Learning and Individual Differences*, 26, 30–41. <https://doi.org/10.1016/j.lindif.2013.04.004>
- Dignath, C. & Büttner, G. (2008). Components of Fostering Self-Regulated Learning among Students. A Meta-Analysis on Intervention Studies at Primary and Secondary School Level. *Metacognition and Learning*, 3(3), 231–264. <https://doi.org/10.1007/s11409-008-9029-x>
- Dignath, C., Büttner, G. & Langfeldt, H.-P. (2008). How Can Primary School Students Learn Self-Regulated Learning Strategies Most Effectively? A Meta-Analysis on Self-Regulation Training Programmes. *Educational Research Review*, 3(2), 101–129. <https://doi.org/10.1016/j.edurev.2008.02.003>
- Dinsmore, D. L. (2017a). Examining the Ontological and Epistemic Assumptions of Research on Metacognition, Self-Regulation and Self-Regulated Learning. *Educational Psychology*, 37(9), 1125–1153. <https://doi.org/10.1080/01443410.2017.1333575>
- Dinsmore, D. L. (2017b). Toward a Dynamic, Multidimensional Research Framework for Strategic Processing. *Educational Psychology Review*, 29(2), 235–268. <https://doi.org/10.1007/s10648-017-9407-5>
- Dinsmore, D. L. & Alexander, P. A. (2012). A Critical Discussion of Deep and Surface Processing: What It Means, How It Is Measured, the Role of Context, and Model Specification. *Educational Psychology Review*, 24(4), 499–567. <https://doi.org/10.1007/s10648-012-9198-7>

- Dinsmore, D. L. & Alexander, P. A. (2016). A Multidimensional Investigation of Deep-Level and Surface-Level Processing. *The Journal of Experimental Education*, 84(2), 213–244. <https://doi.org/10.1080/00220973.2014.979126>
- Dinsmore, D. L., Alexander, P. A. & Loughlin, S. (2008). Focusing the Conceptual Lens on Metacognition, Self-Regulation, and Self-Regulated Learning. *Educational Psychology Review*, 20(4), 391–409. <https://doi.org/10.1007/s10648-008-9083-6>
- Dinsmore, D. L. & Hattan, C. (2020). Levels of Strategies and Strategic Processing. In D. L. Dinsmore, L. K. Fryer & M. M. Parkinson (Hrsg.), *Handbook of Strategies and Strategic Processing* (S. 29–46). Routledge.
- Dinsmore, D. L., Hattan, C. & List, A. (2018). A Meta-Analysis of Strategy Use and Performance in the Model of Domain Learning. In H. Fives & D. L. Dinsmore (Hrsg.), *The Model of Domain Learning: Understanding the Development of Expertise* (S. 37–55). Routledge.
- Dissen, A., Qadiri, Q. & Middleton, C. J. (2022). I Read It Online: Understanding How Undergraduate Students Assess the Accuracy of Online Sources of Health Information. *American journal of lifestyle medicine*, 16(5), 641–654. <https://doi.org/10.1177/1559827621990574>
- Dong, Y., Tang, Y., Chow, B. W.-Y., Wang, W. & Dong, W.-Y. (2020). Contribution of Vocabulary Knowledge to Reading Comprehension Among Chinese Students: A Meta-Analysis. *Frontiers in Psychology*, 11, Artikel 525369, 1–15. <https://doi.org/10.3389/fpsyg.2020.525369>
- Donker, A. S., Boer, H., Kostons, D., Dignath-van Ewijk, C. & van der Werf, M. P. C. (2014). Effectiveness of Learning Strategy Instruction on Academic Performance: A Meta-Analysis. *Educational Research Review*, 11, 1–26. <https://doi.org/10.1016/j.edurev.2013.11.002>
- Dontre, A. J. (2021). The Influence of Technology on Academic Distraction: A Review. *Human Behavior and Emerging Technologies*, 3(3), 379–390. <https://doi.org/10.1002/hbe2.229>
- Dornisch, M. M. & Sperling, R. A. (2006). Facilitating Learning from Technology-Enhanced Text: Effects of Prompted Elaborative Interrogation. *The Journal of Educational Research*, 99(3), 156–165. <https://doi.org/10.3200/JOER.99.3.156-166>
- Dörrenbächer-Ulrich, L., Weisenfels, M., Russer, L. & Perels, F. (2021). Multimethod Assessment of Self-Regulated Learning in College Students: Different Methods for Different Components? *Instructional Science*, 49(1), 137–163. <https://doi.org/10.1007/s11251-020-09533-2>
- Du, H. & List, A. (2020). Researching and Writing Based on Multiple Texts. *Learning and Instruction*, 66, Artikel 101297, 1–25. <https://doi.org/10.1016/j.learninstruc.2019.101297>
- Du, H. & List, A. (2021). Evidence Use in Argument Writing Based on Multiple Texts. *Reading Research Quarterly*, 56(4), 715–735 |. <https://doi.org/10.1002/rtrq.366>
- Du, H. & List, A. (2022). Reasoning about Text-Based Evidence. *Contemporary Educational Psychology*, 68, Artikel 102038, 1–14. <https://doi.org/10.1016/j.cedpsych.2021.102038>
- Duchastel, P. C. & Brown, B. R. (1974). Incidental and Relevant Learning with Instructional Objectives. *Journal of Educational Psychology*, 66(4), 481–485. <https://doi.org/10.1037/h0036743>
- Dumas, D. (2020). Strategic Processing within and across Domains of Learning. In D. L. Dinsmore, L. K. Fryer & M. M. Parkinson (Hrsg.), *Handbook of Strategies and Strategic Processing* (S. 11–28). Routledge.
- Dunlosky, J. & Lipko, A. R. (2007). Metacomprehension: A Brief History and How to Improve Its Accuracy. *Current Directions in Psychological Science*, 16(4), 228–232. <https://doi.org/10.1111/j.1467-8721.2007.00509.x>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. & Willingham, D. T. (2013). Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Dweck, C. S. & Leggett, E. L. (1988). A Social-Cognitive Approach to Motivation and Personality. *Psychological Review*, 95(2), 256–273. <https://doi.org/10.1037/0033-295X.95.2.256>
- Dwyer, C. P., Hogan, M. J. & Stewart, I. (2014). An Integrated Critical Thinking Framework for the 21st Century. *Thinking Skills and Creativity*, 12, 43–52. <https://doi.org/10.1016/j.tsc.2013.12.004>
- Dyoniziak, Y., Potocki, A. & Rouet, J.-F. (2023). Role of Advanced Theory of Mind in Teenagers' Evaluation of Source Information. *Discourse Processes*, 60(4-5), 363–377. <https://doi.org/10.1080/0163853X.2023.2197691>

- Eccles, J. S. & Wigfield, A. (2020). From Expectancy-Value Theory to Situated Expectancy-Value Theory: A Developmental, Social Cognitive, and Sociocultural Perspective on Motivation. *Contemporary Educational Psychology*, 61, Artikel 101859, 1–13. [https://doi.org/10.1016/j.ced-
psych.2020.101859](https://doi.org/10.1016/j.ced-psych.2020.101859)
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K. & Amazeen, M. A. (2022). The Psychological Drivers of Misinformation Belief and Its Resistance to Correction. *Nature Reviews Psychology*, 1(1), 13–29. [https://doi.org/10.1038/
s44159-021-00006-y](https://doi.org/10.1038/s44159-021-00006-y)
- Efklides, A. (2008). Metacognition: Defining Its Facets and Levels of Functioning in Relation to Self-Regulation and Co-Regulation. *European Psychologist*, 13(4), 277–287. [https://doi.
org/10.1027/1016-9040.13.4.277](https://doi.org/10.1027/1016-9040.13.4.277)
- Efklides, A. (2011). Interactions of Metacognition with Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, 46(1), 6–25. [https://doi.org/10.1080/00461
520.2011.538645](https://doi.org/10.1080/00461520.2011.538645)
- Egelhofer, J. L. & Lecheler, S. (2019). Fake News as a Two-Dimensional Phenomenon: A Framework and Research Agenda. *Annals of the International Communication Association*, 43(2), 97–116. <https://doi.org/10.1080/23808985.2019.1602782>
- Eitel, A., Endres, T. & Renkl, A. (2020). Self-Management as a Bridge between Cognitive Load and Self-Regulated Learning: The Illustrative Case of Seductive Details. *Educational Psychology Review*, 32(4), 1073–1087. <https://doi.org/10.1007/s10648-020-09559-5>
- Elhusseini, S. A., Tischner, C. M., Aspiranti, K. B. & Fedewa, A. L. (2022). A Quantitative Review of the Effects of Self-Regulation Interventions on Primary and Secondary Student Academic Achievement. *Metacognition and Learning*, 17(3), 1117–1139. [https://doi.org/10.1007/s11409-
022-09311-0](https://doi.org/10.1007/s11409-022-09311-0)
- Engelmann, K. & Bannert, M. (2021). Analyzing Temporal Data for Understanding the Learning Process Induced by Metacognitive Prompts. *Learning and Instruction*, 72, Artikel 101205, 1–11. <https://doi.org/10.1016/j.learninstruc.2019.05.002>
- Engle, R. W. & Kane, M. J. (2004). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. In B. H. Ross (Hrsg.), *The Psychology of Learning and Motivation: Advances in Research and Theory: Volume 44* (S. 145–199). Elsevier Science.
- Engle, R. W., Kane, M. J. & Tuholski, S. W. (1999). Individual Differences in Working Memory Capacity and What They Tell Us about Controlled Attention, General Fluid Intelligence, and Functions of the Prefrontal Cortex. In A. Miyake & P. Shah (Hrsg.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (S. 102–134). Cambridge University Press. <https://doi.org/10.1017/CBO9781139174909.007>
- Ericsson, K. A. (2018). Superior Working Memory in Experts. In A. M. Williams, A. Kozbelt, K. A. Ericsson & R. R. Hoffman (Hrsg.), *The Cambridge Handbook of Expertise and Expert Performance* (2. Aufl., S. 696–713). Cambridge University Press. <https://doi.org/10.1017/9781316480748.036>
- Ertmer, P. A. & Newby, T. J. (1996). The Expert Learner: Strategic, Self-Regulated, and Reflective. *Instructional Science*, 24(1), 1–24. <https://doi.org/10.1007/BF00156001>
- Escorcía, D., Passerault, J.-M., Ros, C. & Pylouster, J. (2017). Profiling Writers: Analysis of Writing Dynamics among College Students. *Metacognition and Learning*, 12(2), 233–273. [https://doi.
org/10.1007/s11409-016-9166-6](https://doi.org/10.1007/s11409-016-9166-6)
- Espinás, D. R. & Chandler, B. W. (2024). Correlates of K – 12 Students’ Intertextual Integration. *Educational Psychology Review*, 36(2), Artikel 48, 1–48. <https://doi.org/10.1007/s10648-024-09889-8>
- Fendt, M., Nistor, N., Scheibenzuber, C. & Artmann, B. (2023). Sourcing Against Misinformation: Effects of a Scalable Lateral Reading Training Based on Cognitive Apprenticeship. *Computers in Human Behavior*, 146, Artikel 107820, 1–11. <https://doi.org/10.1016/j.chb.2023.107820>
- Ferguson, L. E. & Bråten, I. (2013). Student Profiles of Knowledge and Epistemic Beliefs: Changes and Relations to Multiple-Text Comprehension. *Learning and Instruction*, 25, 49–61. [https://doi.
org/10.1016/j.learninstruc.2012.11.003](https://doi.org/10.1016/j.learninstruc.2012.11.003)
- Ferguson, L. E., Bråten, I. & Strømso, H. I. (2012). Epistemic Cognition when Students Read Multiple Documents Containing Conflicting Scientific Evidence: A Think-Aloud Study. *Learning and Instruction*, 22(2), 103–120. <https://doi.org/10.1016/j.learninstruc.2011.08.002>

- Ferguson, L. E., Bråten, I., Strømsø, H. I. & Anmarkrud, Ø. (2013). Epistemic Beliefs and Comprehension in the Context of Reading Multiple Documents: Examining the Role of Conflict. *International Journal of Educational Research*, 62, 100–114. <https://doi.org/10.1016/j.ijer.2013.07.001>
- Ferguson-Hessler, M. G. M. & Jong, T. (1990). Studying Physics Texts: Differences in Study Processes between Good and Poor Performers. *Cognition and Instruction*, 7(1), 41–54.
- Fiorella, L. & Mayer, R. E. (2022). Principles for Reducing Extraneous Processing in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 185–198). Cambridge University Press. <https://doi.org/10.1017/9781108894333.019>
- Firetto, C. M. & van Meter, P. N. (2018). Inspiring Integration in College Students Reading Multiple Biology Texts. *Learning and Individual Differences*, 65, 123–134. <https://doi.org/10.1016/j.lindif.2018.05.011>
- Fitzgerald, J. (1987). Research on Revision in Writing. *Review of Educational Research*, 57(4), 481–506. <https://doi.org/10.3102/00346543057004481>
- Flanagin, A. J. & Metzger, M. J. (2007). The Role of Site Features, User Attributes, and Information Verification Behaviors on the Perceived Credibility of Web-Based Information. *New Media & Society*, 9(2), 319–342. <https://doi.org/10.1177/1461444807075015>
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Florit, E., Cain, K. & Mason, L. (2020). Going beyond Children's Single-Text Comprehension: The Role of Fundamental and Higher-Level Skills in 4th Graders' Multiple-Document Comprehension. *British Journal of Educational Psychology*, 90(2), 449–472. <https://doi.org/10.1111/bjep.12288>
- Florit, E., Carli, P., Giunti, G. & Mason, L. (2020). Advanced Theory of Mind Uniquely Contributes to Children's Multiple-Text Comprehension. *Journal of Experimental Child Psychology*, 189, Artikel 104708, 1–17. <https://doi.org/10.1016/j.jecp.2019.104708>
- Flower, L. S. (1990). The Role of Task Representation in Reading-to-Write. In L. S. Flower, V. Stein, J. Ackerman, M. J. Kantz, K. McCormick & W. C. Peck (Hrsg.), *Reading-to-Write: Exploring a Cognitive and Social Process* (S. 35–75). Oxford University Press.
- Fogg, B. J. (2003). Prominence-Interpretation Theory: Explaining How People Assess Credibility Online. In G. Cockton (Hrsg.), *ACM Conferences, CHI '03 Extended Abstracts on Human Factors in Computing Systems* (S. 722–723). ACM. <https://doi.org/10.1145/765891.765951>
- Follmer, D. J. (2018). Executive Function and Reading Comprehension: A Meta-Analytic Review. *Educational Psychologist*, 53(1), 42–60. <https://doi.org/10.1080/00461520.2017.1309295>
- Follmer, D. J. & Sperling, R. A. (2019). Examining the Role of Self-Regulated Learning Microanalysis in the Assessment of Learners' Regulation. *The Journal of Experimental Education*, 87(2), 269–287. <https://doi.org/10.1080/00220973.2017.1409184>
- Follmer, D. J. & Sperling, R. A. (2020). The Roles of Executive Functions in Learning from Multiple Representations and Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 297–313). Routledge.
- Follmer, D. J. & Tise, J. (2022). Effects of an Executive Function-Based Text Support on Strategy Use and Comprehension – Integration of Conflicting Informational Texts. *Reading and Writing*, 35(7), 1731–1758. <https://doi.org/10.1007/s11145-022-10257-7>
- Fong, C. J., Krou, M., Hoff, M. & Lin, S. (2021). LASSI's Great Adventure: A Meta-Analytic Review of the Learning and Study Strategies Inventory and Academic Outcomes. *Educational Research Review*, 34, Artikel 100407, 1–24. <https://doi.org/10.1016/j.edurev.2021.100407>
- Forzani, E. (2018). How Well Can Students Evaluate Online Science Information? Contributions of Prior Knowledge, Gender, Socioeconomic Status, and Offline Reading Ability. *Reading Research Quarterly*, 53(4), 385–390. <https://doi.org/10.1002/rrq.218>
- Forzani, E. (2020). A Three-Tiered Framework for Proactive Critical Evaluation during Online Inquiry. *Journal of Adolescent & Adult Literacy*, 63(4), 401–414. <https://doi.org/10.1002/jaal.1004>
- Forzani, E., Corrigan, J. & Kiili, C. (2022). What Does More and Less Effective Internet Evaluation Entail? Investigating Readers' Credibility Judgments across Content, Source, and Context. *Computers in Human Behavior*, 135, Artikel 107359, 1–16. <https://doi.org/10.1016/j.chb.2022.107359>
- Fox, E. (2009). The Role of Reader Characteristics in Processing and Learning from Informational Text. *Review of Educational Research*, 79(1), 197–261. <https://doi.org/10.3102/0034654308324654>

- Frankel, K. K., Becker, B. L. C., Rowe, M. W. & Pearson, P. D. (2017). From „What is Reading?“ to What is Literacy? *Journal of Education*, 196(3), 7–17. <https://doi.org/10.1177/002205741619600303>
- Fryer, L. K. & Shum, A. (2020). Person-Centered Approaches to Explaining Students' Cognitive Processing Strategies. In D. L. Dinsmore, L. K. Fryer & M. M. Parkinson (Hrsg.), *Handbook of Strategies and Strategic Processing* (S. 361–372). Routledge.
- Gagné, E. D. & Rothkopf, E. Z. (1975). Text Organization and Learning Goals. *Journal of Educational Psychology*, 67(3), 445–450. <https://doi.org/10.1037/h0076617>
- García, V., Amadiéu, F. & Salmerón, L. (2021). Integrating Digital Documents by Means of Concept Maps: Testing an Intervention Program with Eye-Movements Modelling Examples. *Heliyon*, 7(12), Artikel e08607, 1–9. <https://doi.org/10.1016/j.heliyon.2021.e08607>
- Garner, R. (1990). When Children and Adults Do Not Use Learning Strategies: Toward a Theory of Settings. *Review of Educational Research*, 60(4), 517–529. <https://doi.org/10.3102/00346543060004517>
- Garner, R., Gillingham, M. G. & White, C. S. (1989). Effects of „Seductive Details“ on Macroprocessing and Microprocessing in Adults and Children. *Cognition and Instruction*, 6(1), 41–57. <http://www.jstor.org/stable/3233462>
- Gerjets, P., Kammerer, Y. & Werner, B. (2011). Measuring Spontaneous and Instructed Evaluation Processes during Web Search: Integrating Concurrent Thinking-Aloud Protocols and Eye-Tracking Data. *Learning and Instruction*, 21(2), 220–231. <https://doi.org/10.1016/j.learninstruc.2010.02.005>
- Gigerenzer, G. & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, 1(1), 107–143. <https://doi.org/10.1111/j.1756-8765.2008.01006.x>
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62, 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gil, L., Ávila, V. & Ferrer, A. (2011). Note-Taking in Multiple-Text Reading: Analysis of Individual Differences. *Infancia y Aprendizaje*, 34(4), 449–464. <https://doi.org/10.1174/021037011797898476>
- Gil, L., Bråten, I., Vidal-Abarca, E. & Strømsø, H. I. (2010a). Summary versus Argument Tasks when Working with Multiple Documents: Which Is Better for Whom? *Contemporary Educational Psychology*, 35(3), 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>
- Gil, L., Bråten, I., Vidal-Abarca, E. & Strømsø, H. I. (2010b). Understanding and Integrating Multiple Science Texts: Summary Tasks Are Sometimes Better than Argument Tasks. *Reading Psychology*, 31(1), 30–68. <https://doi.org/10.1080/02702710902733600>
- Ginn, P. (2006). Integrating Information: A Meta-Analysis of the Spatial Contiguity and Temporal Contiguity Effects. *Learning and Instruction*, 16(6), 511–525. <https://doi.org/10.1016/j.learninstruc.2006.10.001>
- Goetz, E. T., Schallert, D. L., Reynolds, R. E. & Radin, D. I. (1983). Reading in Perspective: What Real Cops and Pretend Burglars Look for in a Story. *Journal of Educational Psychology*, 75(4), 500–510. <https://doi.org/10.1037/0022-0663.75.4.500>
- Goldman, S. R. (2004). Cognitive Aspects of Constructing Meaning through and across Multiple Texts. In D. Bloome & N. Shuart-Faris (Hrsg.), *Uses of Intertextuality in Classroom and Educational Research* (S. 317–351). Information Age.
- Goldman, S. R. (2015). Reading and the Web: Broadening the Need for Complex Comprehension. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink & P. Thompson (Hrsg.), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices* (S. 89–103). Routledge.
- Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C. & Brodowska, K. (2012). Comprehending and Learning from Internet Sources: Processing Patterns of Better and Poorer Learners. *Reading Research Quarterly*, 47(4), 356–381. <https://doi.org/10.1002/RRQ.027>
- Goldman, S. R. & Brand-Gruwel, S. (2018). Learning from Multiple Sources in a Digital Society. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman & P. Reimann (Hrsg.), *International Handbook of the Learning Sciences* (S. 86–95). Routledge.
- Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C. D., Shanahan, C. & READI Project (2016). Disciplinary Literacies and Learning to Read for Understanding: A Conceptual Framework for Disciplinary Literacy. *Educational Psychologist*, 51(2), 219–246. <https://doi.org/10.1080/00461520.2016.1168741>

- Goldman, S. R., Lawless, K. A., Gomez, K. W., Braasch, J. L. G., Macleod, S. M. & Manning, F. H. (2010). Literacy in the Digital World: Comprehending and Learning from Multiple Sources. In M. G. McKeown & L. Kucan (Hrsg.), *Bringing Reading Research to Life: Essays in Honor of Isabel Beck* (S. 257–284). Guilford.
- Goldman, S. R., Lawless, K. A. & Manning, F. H. (2013). Research and Development of Multiple Source Comprehension Assessment. In M. A. Britt, S. R. Goldman & J.-F. Rouet (Hrsg.), *Reading: From Words to Multiple Texts* (S. 180–199). Routledge.
- Goldman, S. R. & Saul, E. U. (1990). Flexibility in Text Processing: A Strategy Competition Model. *Learning and Individual Differences*, 2(2), 181–219. [https://doi.org/10.1016/1041-6080\(90\)90022-9](https://doi.org/10.1016/1041-6080(90)90022-9)
- Goldman, S. R. & Scardamalia, M. (2013). Managing, Understanding, Applying, and Creating Knowledge in the Information Age: Next-Generation Challenges and Opportunities. *Cognition and Instruction*, 31(2), 255–269. <https://doi.org/10.1080/10824669.2013.773217>
- Golparvar, S. E. & Khafi, A. (2021). The Role of L2 Writing Self-Efficacy in Integrated Writing Strategy Use and Performance. *Assessing Writing*, 47, Artikel 100504, 1–15. <https://doi.org/10.1016/j.asw.2020.100504>
- González-Lamas, J., Cuevas, I. & Mateos, M. (2016). Arguing from Sources: Design and Evaluation of a Programme to Improve Written Argumentation and Its Impact According to Students' Writing Beliefs. *Infancia y Aprendizaje*, 39(1), 49–83. <https://doi.org/10.1080/02103702.2015.1111606>
- Gottschling, S. & Kammerer, Y. (2021). Readers' Regulation and Resolution of a Scientific Conflict Based on Differences in Source Information: An Eye-Tracking Study. *Discourse Processes*, 58(5–6), 468–490. <https://doi.org/10.1080/0163853X.2021.1930808>
- Gottschling, S., Kammerer, Y. & Gerjets, P. (2019). Readers' Processing and Use of Source Information as a Function of Its Usefulness to Explain Conflicting Scientific Claims. *Discourse Processes*, 56(5–6), 429–446.
- Gottschling, S., Kammerer, Y., Thomm, E. & Gerjets, P. (2020). How Laypersons Consider Differences in Sources' Trustworthiness and Expertise in Their Regulation and Resolution of Scientific Conflicts. *International Journal of Science Education, Part B*, 10(4), 335–354. <https://doi.org/10.1080/21548455.2020.1849856>
- Graesser, A. C. (2015). Deeper Learning with Advances in Discourse Science and Technology. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 42–50. <https://doi.org/10.1177/2372732215600888>
- Graesser, A. C. & Lehman, B. (2011). Questions Drive Comprehension of Text and Multimedia. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 53–74). Information Age.
- Graesser, A. C., Millis, K. K. & Zwaan, R. A. (1997). Discourse Comprehension. *Annual Review of Psychology*, 48(1), 163–189. <https://doi.org/10.1146/annurev.psych.48.1.163>
- Graesser, A. C., Ozuru, Y. & Sullins, J. (2010). What Is a Good Question? In M. G. McKeown & L. Kucan (Hrsg.), *Bringing Reading Research to Life* (S. 112–141). Guilford Press.
- Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing Inferences during Narrative Text Comprehension. *Psychological Review*, 101(3), 371–395. <https://doi.org/10.1037/0033-295X.101.3.371>
- Graesser, A. C., Swamer, S. S., Baggett, W. B. & Sell, M. A. (1996). New Models of Deep Comprehension. In B. K. Britton & A. C. Graesser (Hrsg.), *Models of Understanding Text* (S. 1–32). Lawrence Erlbaum.
- Graesser, A. C., Wiley, J., Goldman, S. R., O'Reilly, T., Jeon, M. & McDaniel, B. (2007). SEEK Web Tutor: Fostering a Critical Stance while Exploring the Causes of Volcanic Eruption. *Metacognition and Learning*, 2(2–3), 89–105. <https://doi.org/10.1007/s11409-007-9013-x>
- Graham, S. (2020). An Attributional Theory of Motivation. *Contemporary Educational Psychology*, 61, Artikel 101861, 1–11. <https://doi.org/10.1016/j.cedpsych.2020.101861>
- Graham, S. (2020). The Sciences of Reading and Writing Must Become More Fully Integrated. *Reading Research Quarterly*, 55(S1), S35–S44. <https://doi.org/10.1002/rrq.332>
- Graham, S. & Hebert, M. (2011). Writing to Read: A Meta-Analysis of the Impact of Writing and Writing Instruction on Reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>

- Graham, St., Kihara, S. A. & MacKay, M. (2020). The Effects of Writing on Learning in Science, Social Studies, and Mathematics: A Meta-Analysis. *Review of Educational Research*, 90(2), 179–226. <https://doi.org/10.3102/0034654320914744>
- Granado-Peinado, M., Cuevas, I., Olmos, R., Martín, E., Casado-Ledesma, L. & Mateos, M. (2023). Collaborative Writing of Argumentative Syntheses by Low-Performing Undergraduate Writers: Explicit Instruction and Practice. *Reading and Writing*, 36(4), 909–936. <https://doi.org/10.1007/s11145-022-10318-x>
- Granado-Peinado, M., Mateos, M., Martín, E. & Cuevas, I. (2019). Teaching to Write Collaborative Argumentative Syntheses in Higher Education. *Reading and Writing*, 32(8), 2037–2058. <https://doi.org/10.1007/s11145-019-09939-6>
- Greene, C., Symons, S. & Richards, C. (1996). Elaborative Interrogation Effects for Children with Learning Disabilities: Isolated Facts versus Connected Prose. *Contemporary Educational Psychology*, 21(1), 19–42. <https://doi.org/10.1006/ceps.1996.0003>
- Greene, J. A. & Azevedo, R. (2007a). Adolescents' Use of Self-Regulatory Processes and Their Relation to Qualitative Mental Model Shifts while Using Hypermedia. *Journal of Educational Computing Research*, 36(2), 125–148. <https://doi.org/10.2190/G7M1-2734-3JRR-8033>
- Greene, J. A. & Azevedo, R. (2007b). A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Review of Educational Research*, 77(3), 334–372. <https://doi.org/10.3102/003465430303953>
- Greene, J. A. & Azevedo, R. (2009). A Macro-Level Analysis of SRL Processes and Their Relations to the Acquisition of a Sophisticated Mental Model of a Complex System. *Contemporary Educational Psychology*, 34(1), 18–29. <https://doi.org/10.1016/j.cedpsych.2008.05.006>
- Greene, J. A., Azevedo, R. & Torney-Purta, J. (2008). Modeling Epistemic and Ontological Cognition: Philosophical Perspectives and Methodological Directions. *Educational Psychologist*, 43(3), 142–160. <https://doi.org/10.1080/00461520802178458>
- Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C. & McVea, M. (2015). Domain-Specificity of Self-Regulated Learning Processing in Science and History. *Contemporary Educational Psychology*, 42, 111–128. <https://doi.org/10.1016/j.cedpsych.2015.06.001>
- Greene, J. A., Bolick, C. M. & Robertson, J. (2010). Fostering Historical Knowledge and Thinking Skills Using Hypermedia Learning Environments: The Role of Self-Regulated Learning. *Computers & Education*, 54(1), 230–243. <https://doi.org/10.1016/j.compedu.2009.08.006>
- Greene, J. A., Cartiff, B. M. & Duke, R. F. (2018). A Meta-Analytic Review of the Relationship between Epistemic Cognition and Academic Achievement. *Journal of Educational Psychology*, 110(8), 1084–1111. <https://doi.org/10.1037/edu0000263>
- Greene, J. A., Copeland, D. Z., Deekens, V. M. & Yu, S. B. (2018). Beyond Knowledge: Examining Digital Literacy's Role in the Acquisition of Understanding in Science. *Computers & Education*, 117, 141–159. <https://doi.org/10.1016/j.compedu.2017.10.003>
- Greene, J. A., Costa, L.-J. & Dellinger, K. (2011). Analysis of Self-Regulated Learning Processing Using Statistical Models for Count Data. *Metacognition and Learning*, 6(3), 275–301. <https://doi.org/10.1007/s11409-011-9078-4>
- Greene, J. A., Costa, L.-J., Robertson, J., Pan, Y. & Deekens, V. M. (2010). Exploring Relations among College Students' Prior Knowledge, Implicit Theories of Intelligence, and Self-Regulated Learning in a Hypermedia Environment. *Computers & Education*, 55(3), 1027–1043. <https://doi.org/10.1016/j.compedu.2010.04.013>
- Greene, J. A., Deekens, V. M., Copeland, D. Z. & Yu, S. B. (2018). Capturing and Modeling Self-Regulated Learning Using Think-Aloud Protocols. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 323–337). Routledge.
- Greene, J. A., Hutchison, L. A., Costa, L.-J. & Crompton, H. (2012). Investigating How College Students' Task Definitions and Plans Relate to Self-Regulated Learning Processing and Understanding of a Complex Science Topic. *Contemporary Educational Psychology*, 37(4), 307–320. <https://doi.org/10.1016/j.cedpsych.2012.02.002>
- Greene, J. A., Lobczowski, N. G., Freed, R., Cartiff, B. M., Demetriou, C. & Panter, A. T. (2020). Effects of a Science of Learning Course on College Students' Learning With a Computer. *American Educational Research Journal*, 57(3), 947–978. <https://doi.org/10.3102/0002831219865221>

- Greene, J. A., Moos, D. C., Azevedo, R. & Winters, F. I. (2008). Exploring Differences between Gifted and Grade-Level Students' Use of Self-Regulatory Learning Processes with Hypermedia. *Computers & Education*, 50(3), 1069–1083. <https://doi.org/10.1016/j.compedu.2006.10.004>
- Greene, J. A., Muis, K. R. & Pieschl, S. (2010). The Role of Epistemic Beliefs in Students' Self-Regulated Learning with Computer-Based Learning Environments: Conceptual and Methodological Issues. *Educational Psychologist*, 45(4), 245–257. <https://doi.org/10.1080/00461520.2010.515932>
- Greene, J. A., Yu, S. B. & Copeland, D. Z. (2014). Measuring Critical Components of Digital Literacy and Their Relationships with Learning. *Computers & Education*, 76, 55–69. <https://doi.org/10.1016/j.compedu.2014.03.008>
- Greene, S. (1993). The Role of Task in the Development of Academic Thinking through Reading and Writing in a College History Course. *Research in the Teaching of English*, 27(1), 46–75.
- Greene, S. (1994). Students as Authors in the Study of History. In G. Leinhardt, I. L. Beck & C. Stain-ton (Hrsg.), *Teaching and Learning in History* (S. 137–170). Lawrence Erlbaum Associates.
- Griffin, T. D., Jaeger, A. J., Britt, M. A. & Wiley, J. (in Druck). Improving Multiple Document Com-prehension with a Lesson about Multi-Causal Explanations in Science. *Instructional Science*. <https://doi.org/10.1007/s11251-023-09657-1>
- Griffin, T. D., Mielicki, M. K. & Wiley, J. (2019). Improving Students' Metacomprehension Accuracy. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 619–646). Cambridge University Press. <https://doi.org/10.1017/9781108235631.025>
- Griffin, T. D., Wiley, J., Britt, M. A. & Salas, C. R. (2012). The Role of CLEAR Thinking in Learning Science from Multiple-Document Inquiry Tasks. *International Electronic Journal of Elementary Education*, 5(1), 63–78. <https://www.iejee.com/index.php/IEJEE/article/view/10>
- Griffin, T. D., Wiley, J. & Salas, C. R. (2013). Supporting Effective Self-Regulated Learning: The Cri-tical Role of Monitoring. In R. Azevedo & V. Alevin (Hrsg.), *International Handbook of Meta-cognition and Learning Technologies* (S. 19–34). Springer New York. https://doi.org/10.1007/978-1-4419-5546-3_2
- Groeben, N. (2002). Zur konzeptuellen Struktur des Konstrukts „Lesekompetenz“. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (S. 11–21). Juventa.
- Gu, Y. (2012). Learning Strategies: Prototypical Core and Dimensions of Variation. *Studies in Self-Access Learning Journal*, 3(4), 330–356.
- Guo, D., Zhang, S., Wright, K. L. & McTigue, E. M. (2020). Do You Get the Picture? A Meta-Analy-sis of the Effect of Graphics on Reading Comprehension. *AERA Open*, 6(1), 1–20. <https://doi.org/10.1177/2332858420901696>
- Guo, L. (2022a). Seeking Alternatives: How Task Instruction Affects Comprehension of Texts with Conflicting Information. *Reading Psychology*, 43(1), 40–69. <https://doi.org/10.1080/02702711.2021.2008072>
- Guo, L. (2022b). Using Metacognitive Prompts to Enhance Self-Regulated Learning and Learning Outcomes: A Meta-Analysis of Experimental Studies in Computer-Based Learning Environ-ments. *Journal of Computer Assisted Learning*, 38(3), 811–832. <https://doi.org/10.1111/jcal.12650>
- Gutierrez de Blume, A. P. (2022). Calibrating Calibration: A Meta-Analysis of Learning Strategy In-struction Interventions to Improve Metacognitive Monitoring Accuracy. *Journal of Educational Psychology*, 114(4), 681–700. <https://doi.org/10.1037/edu0000674>
- Hacker, D. J. (1998). Self-Regulated Comprehension during Normal Reading. In D. J. Hacker, J. Dun-losky & A. C. Graesser (Hrsg.), *Metacognition in Educational Theory and Practice* (S. 165–191). Erlbaum.
- Hacker, D. J. (2018). A Metacognitive Model of Writing: An Update from a Developmental Perspec-tive. *Educational Psychologist*, 53(4), 220–237. <https://doi.org/10.1080/00461520.2018.1480373>
- Hacker, D. J. & Bol, L. (2019). Calibration and Self-Regulated Learning. In J. Dunlosky & K. A. Raw-son (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 647–677). Cambridge University Press. <https://doi.org/10.1017/9781108235631.026>
- Hadwin, A. F., Boutara, L., Knoetke, T. & Thompson, S. (2004). Cross-Case Study of Self-Regulated Learning as a Series of Events. *Educational Research and Evaluation*, 10(4-6), 365–417. <https://doi.org/10.1080/13803610512331383499>

- Hadwin, A. F., Järvelä, S. & Miller, M. (2011). Self-Regulated, Co-Regulated, and Socially Shared Regulation of Learning. In B. J. Zimmerman & D. H. Schunk (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (S. 65–84). Routledge.
- Hadwin, A. F., Järvelä, S. & Miller, M. (2018). Self-Regulation, Co-Regulation, and Shared Regulation in Collaborative Learning Environments. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 83–106). Routledge.
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J. & Winne, P. H. (2007). Examining Trace Data to Explore Self-Regulated Learning. *Metacognition and Learning*, 2(2–3), 107–124. <https://doi.org/10.1007/s11409-007-9016-7>
- Hadwin, A. F., Winne, P. H., Stockley, D. B., Nesbit, J. C. & Woszczyna, C. (2001). Context Modulates Students' Self-Reports about How They Study. *Journal of Educational Psychology*, 93(3), 477–487. <https://doi.org/10.1037/0022-0663.93.3.477>
- Hagen, Å. M., Braasch, J. L. G. & Bråten, I. (2014). Relationships between Spontaneous Note-Taking, Self-Reported Strategies and Comprehension when Reading Multiple Texts in Different Task Conditions. *Journal of Research in Reading*, 37(1), 141–157. <https://doi.org/10.1111/j.1467-9817.2012.01536.x>
- Hahnel, C., Goldhammer, F., Kroehne, U., Mahlow, N., Artelt, C. & Schoor, C. (2021). Automated and Controlled Processes in Comprehending Multiple Documents. *Studies in Higher Education*, 46(10), 2074–2086. <https://doi.org/10.1080/03075079.2021.1953333>
- Hahnel, C., Goldhammer, F., Kröhne, U. & Naumann, J. (2018). The Role of Reading Skills in the Evaluation of Online Information Gathered from Search Engine Environments. *Computers in Human Behavior*, 78, 223–234. <https://doi.org/10.1016/j.chb.2017.10.004>
- Hahnel, C., Goldhammer, F., Naumann, J. & Kröhne, U. (2016). Effects of Linear Reading, Basic Computer Skills, Evaluating Online Information, and Navigation on Reading Digital Text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>
- Hahnel, C., Kröhne, U., Goldhammer, F., Schoor, C., Mahlow, N. & Artelt, C. (2019). Validating Process Variables of Sourcing in an Assessment of Multiple Document Comprehension. *British Journal of Educational Psychology*, 89(3), 524–537. <https://doi.org/10.1111/bjep.12278>
- Halpain, D. R., Glover, J. A. & Harvey, A. L. (1985). Differential Effects of Higher and Lower Order Questions: Attention Hypotheses. *Journal of Educational Psychology*, 77(6). <https://doi.org/10.1037/0022-0663.77.6.703>
- Hamaker, C. (1986). The Effects of Adjunct Questions on Prose Learning. *Review of Educational Research*, 56(2), 212–242. <https://doi.org/10.3102/00346543056002212>
- Hämäläinen, E. K., Kiili, C., Räikkönen, E., Lakkala, M., Ilomäki, L., Toom, A. & Marttunen, M. (2023). Teaching Sourcing during Online Inquiry – Adolescents with the Weakest Skills Benefited the Most. *Instructional Science*, 51(1), 135–163. <https://doi.org/10.1007/s11251-022-09597-2>
- Hämäläinen, E. K., Kiili, C., Räikkönen, E. & Marttunen, M. (2021). Students' Abilities to Evaluate the Credibility of Online Texts: The Role of Internet-Specific Epistemic Justifications. *Journal of Computer Assisted Learning*, 37(5), 1409–1422. <https://doi.org/10.1111/jcal.12580>
- Hamilton, R. J. (1985). A Framework for the Evaluation of the Effectiveness of Adjunct Questions and Objectives. *Review of Educational Research*, 55(1), 47–85. <https://doi.org/10.3102/00346543055001047>
- Hammer, D. & Elby, A. (2002). On the Form of a Personal Epistemology. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (S. 169–190). Lawrence Erlbaum.
- Hannus, M. & Hyönä, J. (1999). Utilization of Illustrations during Learning of Science Textbook Passages among Low- and High-Ability Children. *Contemporary Educational Psychology*, 24(2), 95–123. <https://doi.org/10.1006/ceps.1998.0987>
- Harris, K. R. (2022). Real Fakes: The Epistemology of Online Misinformation. *Philosophy & Technology*, 35(3), Artikel 83, 1–24. <https://doi.org/10.1007/s13347-022-00581-9>
- Hartmann, U., Kindlinger, M. & Trempler, K. (2021). Integrating Information from Multiple Texts Relates to Pre-Service Teachers' Epistemic Products for Reflective Teaching Practice. *Teaching and Teacher Education*, 97, Artikel 103205, 1–12. <https://doi.org/10.1016/j.tate.2020.103205>

- Hattan, C., Lee, E. & List, A. (2023). Comprehension, Diagram Analysis, Integration, and Interest: A Cross-Sectional Analysis. *Reading Psychology*, 44(7), 731–760. <https://doi.org/10.1080/02702711.2023.2187907>
- Hattie, J. A. C. & Donoghue, G. M. (2016). Learning Strategies: A Synthesis and Conceptual Model. *NPJ Science of Learning*, 1(1), 1–13. <https://doi.org/10.1038/npscilearn.2016.13>
- Hattie, J. A. C., Hodis, F. A. & Kang, S. H. K. (2020). Theories of Motivation: Integration and Ways Forward. *Contemporary Educational Psychology*, 61, 101865. <https://doi.org/10.1016/j.cedpsych.2020.101865>
- Hautala, J., Kiili, C., Kammerer, Y., Loberg, O., Hokkanen, S. & Leppänen, P. H. T. (2018). Sixth Graders' Evaluation Strategies when Reading Internet Search Results: An Eye-Tracking Study. *Behaviour & Information Technology*, 37(8), 761–773. <https://doi.org/10.1080/0144929X.2018.1477992>
- Haverkamp, Y. E., Bråten, I., Latini, N. & Strømso, H. I. (2024). Effects of Media Multitasking on the Processing and Comprehension of Multiple Documents: Does Main Idea Summarization Make a Difference? *Contemporary Educational Psychology*, 77, Artikel 102271, 1–16. <https://doi.org/10.1016/j.cedpsych.2024.102271>
- Heirweg, S., Smul, M. de, Devos, G. & van Keer, H. (2019). Profiling Upper Primary School Students' Self-Regulated Learning through Self-Report Questionnaires and Think-Aloud Protocol Analysis. *Learning and Individual Differences*, 70, 155–168. <https://doi.org/10.1016/j.lindif.2019.02.001>
- Heirweg, S., Smul, M. de, Merchie, E., Devos, G. & van Keer, H. (2020). Mine the Process: Investigating the Cyclical Nature of Upper Primary School Students' Self-Regulated Learning. *Instructional Science*, 48(4), 337–369. <https://doi.org/10.1007/s11251-020-09519-0>
- Hemmerich, J. A. & Wiley, J. (2002). Do Argumentation Tasks Promote Conceptual Change about Volcanoes? *Proceedings of the Annual Meeting of the Cognitive Science Society*, 24, 453–458.
- Hendriks, F., Kienhues, D. & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLoS ONE*, 10(10), Artikel e0139309, 1–20. <https://doi.org/10.1371/journal.pone.0139309>
- Hendriks, F., Kienhues, D. & Bromme, R. (2016). Trust in Science and the Science of Trust. In B. Blöbaum (Hrsg.), *Trust and Communication in a Digitized World: Models and Concepts of Trust Research* (S. 143–159). Springer International Publishing. https://doi.org/10.1007/978-3-319-28059-2_8
- Herrero-Diz, P., Conde-Jiménez, J., Tapia-Frade, A. & Varona-Aramburu, D. (2019). The Credibility of Online News: An Evaluation of the Information by University Students. *Cultura y Educación*, 31(2), 407–435. <https://doi.org/10.1080/11356405.2019.1601937>
- Hidi, S. E. (1995). A Reexamination of the Role of Attention in Learning from Text. *Educational Psychology Review*, 7(4), 323–350. <https://doi.org/10.1007/BF02212306>
- Higgs, K. P., Santuzzi, A. M., Gibson, C., Kopatich, R. D., Feller, D. P. & Magliano, J. P. (2023). Relationships between Task Awareness, Comprehension Strategies, and Literacy Outcomes. *Frontiers in Psychology*, 14, Artikel 1056457, 1–15. <https://doi.org/10.3389/fpsyg.2023.1056457>
- Hilbert, T. S., Nückles, M., Renkl, A., Minarik, C., Reich, A. & Ruhe, K. (2008). Concept Mapping zum Lernen aus Texten. *Zeitschrift für Pädagogische Psychologie*, 22(2), 119–125. <https://doi.org/10.1024/1010-0652.22.2.119>
- Hilbert, T. S. & Renkl, A. (2008). Concept Mapping as a Follow-Up Strategy to Learning from Texts: What Characterizes Good and Poor Mappers? *Instructional Science*, 36(1), 53–73. <https://doi.org/10.1007/s11251-007-9022-9>
- Hofer, B. K. (2001). Personal Epistemology Research: Implications for Learning and Teaching. *Educational Psychology Review*, 13(4), 353–383. <https://doi.org/10.1023/A:1011965830686>
- Hofer, B. K. (2016). Epistemic Cognition as a Psychological Construct: Advancements and Challenges. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of Epistemic Cognition* (S. 19–38). Routledge.
- Hofer, B. K. & Pintrich, P. R. (1997). The Development of Epistemological Theories: Beliefs about Knowledge and Knowing and Their Relation to Learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.2307/1170620>
- Hofmann, W., Schmeichel, B. J. & Baddeley, A. D. (2012). Executive Functions and Self-Regulation. *Trends in Cognitive Sciences*, 16(3), 174–180. <https://doi.org/10.1016/j.tics.2012.01.006>

- Holding, L., Janssen, T. & Rijlaarsdam, G. (2021). The Relationship between Students' Writing Process, Text Quality, and Thought Process Quality in 11th-Grade History and Philosophy Assignments. *Written Communication*, 38(4), 544–586. <https://doi.org/10.1177/07410883211028853>
- Holma, K. & Hyytinen, H. (2015). The Philosophy of Personal Epistemology. *Theory and Research in Education*, 13(3), 334–350. <https://doi.org/10.1177/1477878515606608>
- Howard, J., Bureau, J., Guay, F., Chong, J. X. & Ryan, R. (2021). Student Motivation and Associated Outcomes: A Meta-Analysis from Self-Determination Theory. *Perspectives on Psychological Science*, 16(6), 1300–1323. <https://doi.org/10.1177/1745691620966789>
- Hoyer, J. von, Hoppe, A., Kammerer, Y., Otto, C., Pardi, G., Rokicki, M., Yu, R., Dietze, S., Ewerth, R. & Holtz, P. (2022). The Search as Learning Spaceship: Toward a Comprehensive Model of Psychological and Technological Facets of Search as Learning. *Frontiers in Psychology*, 13, Artikel 827748, 1–19. <https://www.frontiersin.org/article/10.3389/fpsyg.2022.827748>
- Hsu, C.-Y., Tsai, M.-J., Hou, H.-T. & Tsai, C.-C. (2014). Epistemic Beliefs, Online Search Strategies, and Behavioral Patterns While Exploring Socioscientific Issues. *Journal of Science Education and Technology*, 23(3), 471–480. <https://doi.org/10.1007/s10956-013-9477-1>
- Hu, J. & Yu, R. (2021). The Effects of ICT-Based Social Media on Adolescents' Digital Reading Performance: A Longitudinal Study of PISA 2009, PISA 2012, PISA 2015 and PISA 2018. *Computers & Education*, 175, Artikel 104342, 1–20. <https://doi.org/10.1016/j.compedu.2021.104342>
- Huang, C. (2018). Social Network Site Use and Academic Achievement: A Meta-Analysis. *Computers & Education*, 119, 76–83. <https://doi.org/10.1016/j.compedu.2017.12.010>
- Hyönä, J., Lorch, R. F. & Kaakinen, J. K. (2002). Individual Differences in Reading to Summarize Expository Text: Evidence from Eye Fixation Patterns. *Journal of Educational Psychology*, 94(1), 44–55. <https://doi.org/10.1037/0022-0663.94.1.44>
- Hyönä, J. & Nurminen, A.-M. (2006). Do Adult Readers Know How They Read? Evidence from Eye Movement Patterns and Verbal Reports. *British Journal of Psychology*, 97(1), 31–50. <https://doi.org/10.1348/000712605X53678>
- Hyytinen, H., Löfström, E. & Lindblom-Ylänne, S. (2017). Challenges in Argumentation and Paraphrasing among Beginning Students in Educational Sciences. *Scandinavian Journal of Educational Research*, 61(4), 411–429. <https://doi.org/10.1080/00313831.2016.1147072>
- Hyytinen, H., Siven, M., Salminen, O. & Katajavuori, N. (2021). Argumentation and Processing Knowledge in Open-Ended Assignment Tasks: Challenges and Accomplishments among Pharmacy Students. *Journal of University Teaching and Learning Practice*, 18(6), 37–53. <https://doi.org/10.53761/1.18.6.04>
- Hyytinen, H., Toom, A. & Postareff, L. (2018). Unraveling the Complex Relationship in Critical Thinking, Approaches to Learning and Self-Efficacy Beliefs among First-Year Educational Science Students. *Learning and Individual Differences*, 67, 132–142. <https://doi.org/10.1016/j.lindif.2018.08.004>
- Incognito, O. & Tarchi, C. (2024). The Association between Sourcing Skills and Intertextual Integration in Lower Secondary School Students. *European Journal of Psychology of Education*, 39(2), 1485–1500. <https://doi.org/10.1007/s10212-023-00750-0>
- Inzlicht, M., Werner, K. M., Briskin, J. L. & Roberts, B. W. (2021). Integrating Models of Self-Regulation. *Annual Review of Psychology*, 72(1), 319–345. <https://doi.org/10.1146/annurev-psych-061020-105721>
- Ioannou, K. & Iordanou, K. (2020). Elementary School Students' Epistemic Perspective and Learning Strategies in History. *Learning: Research and Practice*, 6(2), 150–166.
- Iordanou, K., Kendeou, P. & Zembylas, M. (2020). Examining My-Side Bias during and after Reading Controversial Historical Accounts. *Metacognition and Learning*, 15(3), 319–342. <https://doi.org/10.1007/s11409-020-09240-w>
- Iordanou, K., Muis, K. R. & Kendeou, P. (2019). Epistemic Perspective and Online Epistemic Processing of Evidence: Developmental and Domain Differences. *The Journal of Experimental Education*, 87(4), 531–551. <https://doi.org/10.1080/00220973.2018.1482857>
- Isberner, M.-B. & Richter, T. (2014). Comprehension and Validation: Separable Stages of Information Processing? A Case for Epistemic Monitoring in Language Comprehension. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 245–276). MIT Press.

- Jacob, R. & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement: A Review. *Review of Educational Research*, 85(4), 512–552. <https://doi.org/10.3102/0034654314561338>
- Jambor-Fahlen, S. & Philipp, M. (2022). Prozess- und Produktperspektiven des Lesens von der Wortebene bis zu multiplen Texten. In M. Philipp & S. Jambor-Fahlen (Hrsg.), *Lesen: Prozess- und Produktperspektiven von der Wortebene bis zu multiplen Texten* (S. 9–33). Beltz Juventa. <https://doi.org/10.5281/zenodo.5882162>
- Järvenoja, H., Järvelä, S. & Malmberg, J. (2015). Understanding Regulated Learning in Situative and Contextual Frameworks. *Educational Psychologist*, 50(3), 204–219. <https://doi.org/10.1080/00461520.2015.1075400>
- Jay, L. (2021). Revisiting Lexington Green: Implications for Teaching Historical Thinking. *Cognition and Instruction*, 39(3), 306–327. <https://doi.org/10.1080/07370008.2021.1880410>
- Jeong, S.-H. & Hwang, Y. (2016). Media Multitasking Effects on Cognitive vs. Attitudinal Outcomes: A Meta-Analysis. *Human Communication Research*, 42(4), 599–618. <https://doi.org/10.1111/hcre.12089>
- Jian, Y.-C. & Ko, H.-W. (2017). Influences of Text Difficulty and Reading Ability on Learning Illustrated Science Texts for Children: An Eye Movement Study. *Computers & Education*, 113, 263–279. <https://doi.org/10.1016/j.compedu.2017.06.002>
- Jiang, L. & Elen, J. (2011). Instructional Effectiveness of Higher-Order Questions: The Devil Is in the Detail of Students' Use of Questions. *Learning Environments Research*, 14(3), 279–298. <https://doi.org/10.1007/s10984-011-9095-x>
- Johnson, A. M., Azevedo, R. & D'Mello, S. K. (2011). The Temporal and Dynamic Nature of Self-Regulatory Processes during Independent and Externally Assisted Hypermedia Learning. *Cognition and Instruction*, 29(4), 471–504. <https://doi.org/10.1080/07370008.2011.610244>
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press. <https://books.google.ch/books?id=FS3zSKAflGMC>
- Jong, T. de (2010). Cognitive Load Theory, Educational Research, and Instructional Design: Some Food for Thought. *Instructional Science*, 38(2), 105–134. <https://doi.org/10.1007/s11251-009-9110-0>
- Jordão, M., Ferreira-Santos, F., Pinho, M. S. & St Jacques, P. L. (2019). Meta-Analysis of Aging Effects in Mind Wandering: Methodological and Sociodemographic Factors. *Psychology and aging*, 34(4), 531–544. <https://doi.org/10.1037/pag0000356>
- Jurado, M. B. & Rosselli, M. (2007). The Elusive Nature of Executive Functions: A Review of Our Current Understanding. *Neuropsychology Review*, 17(3), 213–233. <https://doi.org/10.1007/s11065-007-9040-z>
- Kaakinen, J. K. & Hyönä, J. (2005). Perspective Effects on Expository Text Comprehension: Evidence from Think-Aloud Protocols, Eyetracking, and Recall. *Discourse Processes*, 40(3), 239–257. https://doi.org/10.1207/s15326950dp4003_4
- Kaakinen, J. K. & Hyönä, J. (2007). Perspective Effects in Repeated Reading: An Eye Movement Study. *Memory & Cognition*, 35(6), 1323–1336. <https://doi.org/10.3758/BF03193604>
- Kaakinen, J. K. & Hyönä, J. (2008). Perspective-Driven Text Comprehension. *Applied Cognitive Psychology*, 22(3), 319–334. <https://doi.org/10.1002/acp.1412>
- Kaakinen, J. K. & Hyönä, J. (2011). Online Processing of and Memory for Perspective Relevant and Irrelevant Text Information. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 223–242). Information Age.
- Kaakinen, J. K. & Hyönä, J. (2014). Task Relevance Induces Momentary Changes in the Functional Visual Field during Reading. *Psychological Science*, 25(2), 626–632. <https://doi.org/10.1177/0956797613512332>
- Kaakinen, J. K., Hyönä, J. & Keenan, J. M. (2001). Individual Differences in Perspective Effects on Text Memory. *Current Psychology Letters: Behaviour, Brain & Cognition*, 5, 21–32.
- Kaakinen, J. K., Hyönä, J. & Keenan, J. M. (2002). Perspective Effects on Online Text Processing. *Discourse Processes*, 33(2), 159–173. https://doi.org/10.1207/S15326950DP3302_03
- Kaakinen, J. K., Hyönä, J. & Keenan, J. M. (2003). How Prior Knowledge, WMC, and Relevance of Information Affect Eye Fixations in Expository Text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(3), 447–457. <https://doi.org/10.1037/0278-7393.29.3.447>

- Kaakinen, J. K., Lehtola, A. & Paattilampi, S. (2015). The Influence of a Reading Task on Children's Eye Movements during Reading. *Journal of Cognitive Psychology*, 27(5), 640–656. <https://doi.org/10.1080/20445911.2015.1005623>
- Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, 23(1), 1–19. <https://doi.org/10.1007/s10648-010-9150-7>
- Kalyuga, S. & Singh, A.-M. (2016). Rethinking the Boundaries of Cognitive Load Theory in Complex Learning. *Educational Psychology Review*, 28(4), 831–852. <https://doi.org/10.1007/s10648-015-9352-0>
- Kammerer, Y., Amann, D. G. & Gerjets, P. (2015). When Adults without University Education Search the Internet for Health Information: The Roles of Internet-Specific Epistemic Beliefs and a Source Evaluation Intervention. *Computers in Human Behavior*, 48, 297–309. <https://doi.org/10.1016/j.chb.2015.01.045>
- Kammerer, Y., Bråten, I., Gerjets, P. & Strømsø, H. I. (2013). The Role of Internet-Specific Epistemic Beliefs in Laypersons' Source Evaluations and Decisions during Web Search on a Medical Issue. *Computers in Human Behavior*, 29(3), 1193–1203. <https://doi.org/10.1016/j.chb.2012.10.012>
- Kammerer, Y. & Gerjets, P. (2012a). Effects of Search Interface and Internet-Specific Epistemic Beliefs on Source Evaluations during Web Search for Medical Information: An Eye-Tracking Study. *Behaviour & Information Technology*, 31(1), 83–97. <https://doi.org/10.1080/0144929X.2011.599040>
- Kammerer, Y. & Gerjets, P. (2012b). The Impact of Discrepancies across Web Pages on High School Students' Trustworthiness Evaluations. In E. Vries & K. Scheiter (Hrsg.), *Staging Knowledge and Experience: How to Take Advantage of Representational Technologies in Education and Training?* (S. 97–99). Université Pierre-Mendès-France.
- Kammerer, Y. & Gerjets, P. (2014a). Quellenbewertungen und Quellenverweise bei Lesen und Zusammenfassen wissensbezogener Informationen aus multiplen Webseiten. *Unterrichtswissenschaft*, 42(1), 7–23.
- Kammerer, Y. & Gerjets, P. (2014b). The Role of Search Result Position and Source Trustworthiness in the Selection of Web Search Results When Using a List or a Grid Interface. *International Journal of Human – Computer Interaction*, 30(3), 177–191. <https://doi.org/10.1080/10447318.2013.846790>
- Kammerer, Y., Gottschling, S. & Bråten, I. (2021). The Role of Internet-Specific Justification Beliefs in Source Evaluation and Corroboration during Web Search on an Unsettled Socio-Scientific Issue. *Journal of Educational Computing Research*, 59(2), 342–378. <https://doi.org/10.1177/0735633120952731>
- Kammerer, Y., Kalbfell, E. & Gerjets, P. (2016). Is This Information Source Commercially Biased? How Contradictions between Web Pages Stimulate the Consideration of Source Information. *Discourse Processes*, 53(5–6), 430–456. <https://doi.org/10.1080/0163853X.2016.1169968>
- Kammerer, Y., Keck, D. & Starauschek, E. (2019). Effects of a Short Classroom Intervention on Students' Identification of Contradictions in an Internet Forum Text: Interaction with Reading Comprehension Skills. *Computers & Education*, 138, 46–56. <https://doi.org/10.1016/j.compedu.2019.04.015>
- Kammerer, Y., Meier, N. & Stahl, E. (2016). Fostering Secondary-School Students' Intertext Model Formation when Reading a Set of Websites: The Effectiveness of Source Prompts. *Computers & Education*, 102, 52–64. <https://doi.org/10.1016/j.compedu.2016.07.001>
- Kämpfe, J., Sedlmeier, P. & Renkewitz, F. (2011). The Impact of Background Music on Adult Listeners: A Meta-Analysis. *Psychology of Music*, 39(4), 424–448. <https://doi.org/10.1177/0305735610376261>
- Kanninen, L., Kiili, C., Tolvanen, A., Aro, M. & Leppänen, P. H. T. (2019). Literacy Skills and Online Research and Comprehension: Struggling Readers Face Difficulties Online. *Reading and Writing*, 32(9), 2201–2222. <https://doi.org/10.1007/s11145-019-09944-9>
- Kaplan, A. (2008). Clarifying Metacognition, Self-Regulation, and Self-Regulated Learning: What's the Purpose? *Educational Psychology Review*, 20(4), 477–484. <https://doi.org/10.1007/s10648-008-9087-2>
- Kardash, C. M. & Howell, K. L. (2000). Effects of Epistemological Beliefs and Topic-Specific Beliefs on Undergraduates' Cognitive and Strategic Processing of Dual-Positional Text. *Journal of Educational Psychology*, 92(3), 524–535. <https://doi.org/10.1037/0022-0663.92.3.524>

- Kardash, C. M. & Scholes, R. J. (1995). Effects of Preexisting Beliefs and Repeated Readings on Belief Change, Comprehension, and Recall of Persuasive Text. *Contemporary Educational Psychology*, 20(2), 201–221. <https://doi.org/10.1006/ceps.1995.1013>
- Kardash, C. M. & Scholes, R. J. (1996). Effects of Preexisting Beliefs, Epistemological Beliefs, and Need for Cognition on Interpretation of Controversial Issues. *Journal of Educational Psychology*, 88(2), 260–271. <https://doi.org/10.1037//0022-0663.88.2.260>
- Karimi, M. N. (2015). L2 Multiple-Documents Comprehension: Exploring the Contributions of L1 Reading Ability and Strategic Processing. *System*, 52, 14–25. <https://doi.org/10.1016/j.system.2015.04.019>
- Karimi, M. N. (2018). The Mediated/Unmediated Contributions of Language Proficiency and Prior Knowledge to L2 Multiple-Texts Comprehension: A Structural Equation Modelling Analysis. *Applied Linguistics*, 39(6), 912–932. <https://doi.org/10.1093/applin/amw059>
- Karimi, M. N. (2024). Text-Belief Consistency Effect in L2 Students' Integrated Written Representations Based on Multiple Conflicting Sources. *Contemporary Educational Psychology*, 76, Artikel 102239, 1–11. <https://doi.org/10.1016/j.cedpsych.2023.102239>
- Karimi, M. N. & Ashkani, P. (2023). Situation-Model Representations of Conflicting Textual Information in L2 Readers: The Effects of Prior Beliefs and L2 Proficiency. *Language Awareness*, 32(2), 323–341. <https://doi.org/10.1080/09658416.2022.2091583>
- Karimi, M. N. & Atai, M. R. (2014). ESAP Students' Comprehension of Multiple Technical Reading Texts: Insights from Personal Epistemological Beliefs. *Reading Psychology*, 35(8), 736–761. <https://doi.org/10.1080/02702711.2013.802753>
- Karimi, M. N. & Mousavi-Moghadam, T. (2022). Integrated Writing across Conflicting L2 Texts: The Effects of Prior Beliefs on Perspectivism and Propositional Content. *Assessing Writing*, 53, Artikel 100648, 1–10. <https://doi.org/10.1016/j.asw.2022.100648>
- Karimi, M. N. & Richter, T. (2021a). Biased Representations of Controversial Information: Certainty and Justification Beliefs as Moderators. *Contemporary Educational Psychology*, 66, Artikel 101995, 1–12. <https://doi.org/10.1016/j.cedpsych.2021.101995>
- Karimi, M. N. & Richter, T. (2021b). Text-Belief Consistency Effects in L2 Readers. *Discourse Processes*, 58(8), 726–742. <https://doi.org/10.1080/0163853X.2021.1913935>
- Karimi, M. N. & Richter, T. (2023a). Belief-Biased Representations of Textual Information in Bilinguals: Language as a Source Characteristic. *Current Psychology*, 42(12), 9852–9866. <https://doi.org/10.1007/s12144-021-02239-9>
- Karimi, M. N. & Richter, T. (2023b). Thinking Dispositions Moderate Adolescent Readers' Mental Models of Conflicting Sport Information. *Thinking Skills and Creativity*, 47, Artikel 101233, 1–14. <https://doi.org/10.1016/j.tsc.2023.101233>
- Karimi, M. N. & Shabani, M. B. (2013). Comparing the Strategic Behavior of More Successful vs. Less Successful Readers of Multiple Technical Reading Texts. *Innovation in Language Learning and Teaching*, 7(2), 125–138. <https://doi.org/10.1080/17501229.2012.726223>
- Karoly, P. (1993). Mechanisms of Self-Regulation: A Systems View. *Annual Review of Psychology*, 44(1), 23–52. <https://doi.org/10.1146/annurev.ps.44.020193.000323>
- Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L. & Garcia-Barrera, M. A. (2018). The Unity and Diversity of Executive Functions: A Systematic Review and Re-Analysis of Latent Variable Studies. *Psychological Bulletin*, 144(11), 1147–1185. <https://doi.org/10.1037/bul0000160>
- Kasneji, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneji, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*, 103, Artikel 102274, 1–9. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kates, A. W., Wu, H. & Coryn, C. L. (2018). The Effects of Mobile Phone Use on Academic Performance: A Meta-Analysis. *Computers & Education*, 127, 107–112. <https://doi.org/10.1016/j.compedu.2018.08.012>
- Kawashima, I., Hinuma, T. & Tanaka, S. C. (2023). Ecological Momentary Assessment of Mind-Wandering: Meta-Analysis and Systematic Review. *Scientific Reports*, 13(1), Artikel 2873, 1–11. <https://doi.org/10.1038/s41598-023-29854-9>

- Keck, D., Kammerer, Y. & Starauschek, E. (2015). Reading Science Texts Online: Does Source Information Influence the Identification of Contradictions within Texts? *Computers & Education*, 82, 442–449. <https://doi.org/10.1016/j.compedu.2014.12.005>
- Kendeou, P. (2015). A General Inference Skill. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Hrsg.), *Inferences during Reading* (S. 160–181). Cambridge University Press.
- Kendeou, P. (2024). A Theory of Knowledge Revision: The Development of the KReC Framework. *Educational Psychology Review*, 36(2), Artikel 44, 1–21. <https://doi.org/10.1007/s10648-024-09885-y>
- Kendeou, P., McMaster, K. L. & Christ, T. J. (2016). Reading Comprehension: Core Components and Processes. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 62–69. <https://doi.org/10.1177/2372732215624707>
- Kendeou, P., Muis, K. R. & Fulton, S. (2011). Reader and Text Factors in Reading Comprehension Processes. *Journal of Research in Reading*, 34(4), 365–383. <https://doi.org/10.1111/j.1467-9817.2010.01436.x>
- Kendeou, P. & O'Brien, E. J. (2014). The Knowledge Revision Components (KReC) Framework: Processes and Mechanisms. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 353–377). MIT Press.
- Kendeou, P., Robinson, D. H. & McCrudden, M. T. (2019). Misinformation and Disinformation in Education: An Introduction. In P. Kendeou, D. H. Robinson & M. T. McCrudden (Hrsg.), *Misinformation and Fake News in Education* (S. 1–4). Information Age.
- Kennedy, M. L. (1985). The Composing Process of College Students Writing from Sources. *Written Communication*, 2(4), 434–456. <https://doi.org/10.1177/0741088385002004006>
- Kent, S. C. & Wanzek, J. (2016). The Relationship between Component Skills and Writing Quality and Production across Developmental Levels: A Meta-Analysis of the Last 25 Years. *Review of Educational Research*, 86(2), 570–601. <https://doi.org/10.3102/0034654315619491>
- Keren, A. (2018). The Public Understanding of What? Laypersons' Epistemic Needs, the Division of Cognitive Labor, and the Demarcation of Science. *Philosophy of Science*, 85(5), 781–792. <https://doi.org/10.1086/699690>
- Kessler, E. D., Braasch, J. L. G. & Kardash, C. M. (2019). Individual Differences in Revising (and Maintaining) Accurate and Inaccurate Beliefs about Childhood Vaccinations. *Discourse Processes*, 56(5-6), 415–428. <https://doi.org/10.1080/0163853X.2019.1596709>
- Kessler, E. D., Braasch, J. L. G. & Kardash, C. M. (2021). The Contributions of Childhood Vaccination Misconceptions to the Evaluation and Sharing of Information from Multiple Internet Texts. *Reading Psychology*, 42(3), 281–301. <https://doi.org/10.1080/02702711.2021.1888357>
- Kienhues, D., Jucks, R. & Bromme, R. (2020). Sealing the Gateways for Post-Truthism: Reestablishing the Epistemic Authority of Science. *Educational Psychologist*, 55(3), 144–154. <https://doi.org/10.1080/00461520.2020.1784012>
- Kiili, C., Brante, E. W., Räikkönen, E. M. & Coiro, J. (2020). Citing as a Sourcing Practice: Students' Citing Self-Selected Online Sources in Their Essays. *Journal for the Study of Education and Development*, 43(1), 174–209.
- Kiili, C., Bråten, I., Kullberg, N. & Leppänen, P. H. T. (2020). Investigating Elementary School Students' Text-Based Argumentation with Multiple Online Information Resources. *Computers & Education*, 147(Article 103785), 1–15. <https://doi.org/10.1016/j.compedu.2019.103785>
- Kiili, C., Bråten, I., Strömso, H. I., Hagerman, M. S., Räikkönen, E. & Jyrkiäinen, A. (2022). Adolescents' Credibility Justifications when Evaluating Online Texts. *Education and Information Technologies*, 27(6), 7421–7450. <https://doi.org/10.1007/s10639-022-10907-x>
- Kiili, C., Coiro, J. & Raikkonen, E. (2019). Students' Evaluation of Information during Online Inquiry: Working Individually or in Pairs. *Australian Journal of Language and Literacy*, 42(3), 167–183.
- Kiili, C., Forzani, E., Brante, E. W., Räikkönen, E. & Marttunen, M. (2021). Sourcing on the Internet: Examining the Relations among Different Phases of Online Inquiry. *Computers and Education Open*, 2, Artikel 100037, 1–11. <https://doi.org/10.1016/j.caeo.2021.100037>
- Kiili, C., Laurinen, L. & Marttunen, M. (2008). Students Evaluating Internet Sources: From Versatile Evaluators to Uncritical Readers. *Journal of Educational Computing Research*, 39(1), 75–95. <https://doi.org/10.2190/EC.39.1.e>

- Kiili, C., Leu, D. J., Marttunen, M., Hautala, J. & Leppänen, P. H. T. (2018). Exploring Early Adolescents' Evaluation of Academic and Commercial Online Resources Related to Health. *Reading and Writing*, 31(3), 533–557. <https://doi.org/10.1007/s11145-017-9797-2>
- Kiili, C., Leu, D. J., Utriainen, J., Coiro, J., Kanninen, L., Tolvanen, A., Lohvansuu, K. & Leppänen, P. H. T. (2018). Reading to Learn from Online Information: Modeling the Factor Structure. *Journal of Literacy Research*, 50(3), 304–334. <https://doi.org/10.1177/1086296X18784640>
- Kiili, C., Räikkönen, E., Bråten, I., Strømso, H. I. & Hagerman, M. S. (2023). Examining the Structure of Credibility Evaluation when Sixth Graders Read Online Texts. *Journal of Computer Assisted Learning*, 39(3), 954–969. <https://doi.org/10.1111/jcal.12779>
- Kiili, C., Strømso, H. I., Bråten, I., Ruotsalainen, J. & Räikkönen, E. (in Druck). Reading Comprehension Skills and Prior Topic Knowledge Serve as Resources when Adolescents Justify the Credibility of Multiple Online Texts. *Reading Psychology*. <https://doi.org/10.1080/02702711.2024.2351485>
- Kim, H. (2020). Profiles of Undergraduate Student Writers: Differences in Writing Strategy and Impacts on Text Quality. *Learning and Individual Differences*, 78(Article 101823), 1–12. <https://doi.org/10.1016/j.lindif.2020.101823>
- Kim, Y.-S. G. (2020). Interactive Dynamic Literacy Model: An Integrative Theoretical Framework for Reading-Writing Relations. In R. A. Alves, T. Limpo & R. M. Joshi (Hrsg.), *Reading-Writing Connections: Towards Integrative Literacy Science* (S. 11–34). Springer. https://doi.org/10.1007/978-3-030-38811-9_2
- Kim, Y.-S. G., Wolters, A. & Lee, J. w. (2024). Reading and Writing Relations Are Not Uniform: They Differ by the Linguistic Grain Size, Developmental Phase, and Measurement. *Review of Educational Research*, 94(3), 311–342. <https://doi.org/10.3102/00346543231178830>
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Kintsch, W. (2018). Revisiting the Construction – Integration Model of Text Comprehension and Its Implications for Instruction. In D. E. Alvermann, N. J. Unrau, M. Sailors & R. B. Ruddell (Hrsg.), *Theoretical Models and Processes of Literacy* (7. Aufl., S. 178–203). Routledge.
- Kintsch, W. & van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kirkpatrick, L. C. & Klein, P. D. (2009). Planning Text Structure as a Way to Improve Students' Writing from Sources in the Compare – Contrast Genre. *Learning and Instruction*, 19(4), 309–321. <https://doi.org/10.1016/j.learninstruc.2008.06.001>
- Kirkpatrick, L. C. & Klein, P. D. (2016). High-Achieving High School Students' Strategies for Writing from Internet-Based Sources of Information. *Journal of Writing Research*, 8(1), 299–344.
- Kirschner, P. A. & Bruyckere, P. de (2017). The Myths of the Digital Native and the Multitasker. *Teaching and Teacher Education*, 67, 135–142. <https://doi.org/10.1016/j.tate.2017.06.001>
- Kitcher, P. (2011). Public Knowledge and Its Discontents. *Theory and Research in Education*, 9(2), 103–124. <https://doi.org/10.1177/1477878511409618>
- Klein, P. D. & van Dijk, A. (2019). Writing as a Learning Activity. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 266–291). Cambridge University Press. <https://doi.org/10.1017/9781108235631.012>
- Knoch, U. & Sitajalabhorn, W. (2013). A Closer Look at Integrated Writing Tasks: Towards a More Focused Definition for Assessment Purposes. *Assessing Writing*, 18(4), 300–308. <https://doi.org/10.1016/j.asw.2013.09.003>
- Kobayashi, K. (2007). The Influence of Critical Reading Orientation on External Strategy Use during Expository Text Reading. *Educational Psychology*, 27(3), 363–375. <https://doi.org/10.1080/01443410601104171>
- Kobayashi, K. (2009). The Influence of Topic Knowledge, External Strategy Use, and College Experience on Students' Comprehension of Controversial Texts. *Learning and Individual Differences*, 19(1), 130–134. <https://doi.org/10.1016/j.lindif.2008.06.001>
- Kobayashi, K. (2014). Students' Consideration of Source Information during the Reading of Multiple Texts and Its Effect on Intertextual Conflict Resolution. *Instructional Science*, 42(2), 183–205. <https://doi.org/10.1007/s11251-013-9276-3>
- Kobayashi, K. (2015). Learning from Conflicting Texts: The Role of Intertextual Conflict Resolution in Between-Text Integration. *Reading Psychology*, 36(6), 519–544. <https://doi.org/10.1080/02702711.2014.926304>

- Koch, I., Poljac, E., Muller, H. & Kiesel, A. (2018). Cognitive Structure, Flexibility, and Plasticity in Human Multitasking – An Integrative Review of Dual-Task and Task-Switching Research. *Psychological Bulletin*, 144(6), 557–583. <https://doi.org/10.1037/bul0000144>
- Kohnen, A. M., Mertens, G. E. & Boehm, S. M. (2020). Can Middle Schoolers Learn to Read the Web Like Experts? Possibilities and Limits of a Strategy-Based Intervention. *Journal of Media Literacy Education*, 12(2), 64–79. <https://doi.org/10.23860/JMLE-2020-12-2-6>
- Kohnen, A. M., Mertens, G. E., Dawson, K., Hampton, J. & Fu, D. (2021). A Study of Middle School Students' Online Credibility Assessments: Challenges and Possibilities. *Research in the Teaching of English*, 56(1), 33–59. <https://doi.org/10.58680/rte202131342>
- Kong, F., Meng, S., Deng, H., Wang, M. & Sun, X. (2023). Cognitive Control in Adolescents and Young Adults with Media Multitasking Experience: A Three-Level Meta-Analysis. *Educational Psychology Review*, 35(1), Artikel 22, 1–37. <https://doi.org/10.1007/s10648-023-09746-0>
- Kong, Y., Seo, Y. S. & Zhai, L. (2018). Comparison of Reading Performance on Screen and on Paper: A Meta-Analysis. *Computers & Education*, 123, 138–149. <https://doi.org/10.1016/j.compedu.2018.05.005>
- Kozyreva, A., Lewandowsky, S. & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges with Cognitive Tools. *Psychological Science in the Public Interest*, 21(3), 103–156. <https://doi.org/10.1177/1529100620946707>
- Kozyreva, A., Wineburg, S. S., Lewandowsky, S. & Hertwig, R. (2023). Critical Ignoring as a Core Competence for Digital Citizens. *Current Directions in Psychological Science*, 32(1), 81–88. <https://doi.org/10.1177/09637214221121570>
- Kriegelstein, F., Beege, M., Rey, G. D., Ginns, P., Krell, M. & Schneider, S. (2022). A Systematic Meta-Analysis of the Reliability and Validity of Subjective Cognitive Load Questionnaires in Experimental Multimedia Learning Research. *Educational Psychology Review*, 34(4), 2485–2541. <https://doi.org/10.1007/s10648-022-09683-4>
- Kuhn, D. (2020). Why Is Reconciling Divergent Views a Challenge? *Current Directions in Psychological Science*, 29(1), 27–32. <https://doi.org/10.1177/0963721419885996>
- Kuhn, D. & Weinstock, M. (2002). What Is Epistemological Thinking and Why Does It Matter? In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (S. 121–144). Lawrence Erlbaum.
- Kullberg, N., Kiili, C., Bråten, I., González-Ibáñez, R. & Leppänen, P. H. T. (2023). Sixth Graders' Selection and Integration when Writing from Multiple Online Texts. *Instructional Science*, 51(1), 39–64. <https://doi.org/10.1007/s11251-022-09613-5>
- Kurby, C. A., Britt, M. A. & Magliano, J. P. (2005). The Role of Top-Down and Bottom-Up Processes in Between-Text Integration. *Reading Psychology*, 26(4-5), 335–362. <https://doi.org/10.1080/02702710500285870>
- Laing, S. P. & Kamhi, A. G. (2002). The Use of Think-Aloud Protocols to Compare Inferencing Abilities in Average and Below-Average Readers. *Journal of Learning Disabilities*, 35(5), 437–448. <https://doi.org/10.1177/00222194020350050401>
- Lang, A. & Chrzan, J. (2015). Media Multitasking: Good, Bad, or Ugly? *Annals of the International Communication Association*, 39(1), 99–128. <https://doi.org/10.1080/23808985.2015.11679173>
- Lang, F., Kammerer, Y., Oschatz, K., Stürmer, K. & Gerjets, P. (2020). The Role of Beliefs Regarding the Uncertainty of Knowledge and Mental Effort as Indicated by Pupil Dilation in Evaluating Scientific Controversies. *International Journal of Science Education*, 42(3), 350–371. <https://doi.org/10.1080/09500693.2019.1710875>
- Lang, F., Kammerer, Y., Stürmer, K. & Gerjets, P. (2021). Investigating Professed and Enacted Epistemic Beliefs about the Uncertainty of Scientific Knowledge When Students Evaluate Scientific Controversies. *European Journal of Psychology of Education*, 36(1), 125–146. <https://doi.org/10.1007/s10212-020-00471-8>
- Lapan, R. & Reynolds, R. E. (1994). The Selective Attention Strategy as a Time-Dependent Phenomenon. *Contemporary Educational Psychology*, 19(4), 379–398. <https://doi.org/10.1006/ceps.1994.1028>

- Latini, N., Bråten, I., Anmarkrud, Ø. & Salmerón, L. (2019). Investigating Effects of Reading Medium and Reading Purpose on Behavioral Engagement and Textual Integration in a Multiple Text Context. *Contemporary Educational Psychology*, 59, Artikel 101797, 1–16. <https://doi.org/10.1016/j.cedpsych.2019.101797>
- Latini, N., Bråten, I. & Haverkamp, Y. E. (2021). Breadth and Depth of Strategic Processing during Text Comprehension. *Learning and Individual Differences*, 91, Artikel 102058, 1–10. <https://doi.org/10.1016/j.lindif.2021.102058>
- Lawless, K. A. & Kulikowich, J. M. (1996). Understanding Hypertext Navigation through Cluster Analysis. *Journal of Educational Computing Research*, 14(4), 385–399. <https://doi.org/10.2190/DVAP-DE23-3XMV-9MXH>
- Lawless, K. A., Mills, R. & Brown, S. W. (2002). Children's Hypertext Navigation Strategies. *Journal of Research on Technology in Education*, 34(3), 274–284. <https://doi.org/10.1080/15391523.2002.10782349>
- Le Bigot, L. & Rouet, J.-F. (2007). The Impact of Presentation Format, Task Assignment, and Prior Knowledge on Students' Comprehension of Multiple Online Documents. *Journal of Literacy Research*, 39(4), 445–470. <https://doi.org/10.1080/10862960701675317>
- Lee, E., List, A., Campos Oaxaca, G. S., Lee, H. Y. & Du, H. (in Druck). Using Rhetorical Devices to Improve Integration in Writing Based on Multiple Texts. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-023-00778-2>
- Lee, H. Y. & List, A. (2019). Processing of Texts and Videos: A Strategy-Focused Analysis. *Journal of Computer Assisted Learning*, 35(2), 268–282. <https://doi.org/10.1111/jcal.12328>
- Lee, H. Y. & List, A. (2021). Examining Students' Self-Efficacy and Perceptions of Task Difficulty in Learning from Multiple Texts. *Learning and Individual Differences*, 90, Artikel 102052, 1–15. <https://doi.org/10.1016/j.lindif.2021.102052>
- Lee, H. Y. & List, A. (2023a). Examining Students' Help-Seeking when Learning from Multiple Texts. *Contemporary Educational Psychology*, 75, Artikel 102232, 1–17. <https://doi.org/10.1016/j.cedpsych.2023.102232>
- Lee, H. Y. & List, A. (2023b). The Role of Relevance Determinations in Multiple Text Reading and Writing: An Investigation of the MD-TRACE. *Discourse Processes*, 60(1), 42–72. <https://doi.org/10.1080/0163853X.2022.2159741>
- Lee, Y.-H., Jhang, J.-Y. & Hong, H.-Y. (2024). Becoming Epistemically Active in Online Reading: Facilitating Elementary School Students' Multimodal Multiple Document Reading via Sourcing Organizers. *Computers & Education*, 216, Artikel 105048, 1–12. <https://doi.org/10.1016/j.compedu.2024.105048>
- Lee, Y.-H. (2022). Beyond Online Search Strategies: The Effects of Internet Epistemic Beliefs and Different Note-Taking Formats on Online Multiple Document Reading Comprehension. *Journal of Computer Assisted Learning*, 38(4), 1102–1114. <https://doi.org/10.1111/jcal.12668>
- Lee-Sammons, W. H. & Whitney, P. (1991). Reading Perspectives and Memory for Text: An Individual Differences Analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(6), 1074–1081. <https://doi.org/10.1037/0278-7393.17.6.1074>
- Lehman, S. & Schraw, G. (2002). Effects of Coherence and Relevance on Shallow and Deep Text Processing. *Journal of Educational Psychology*, 94(4), 738–750. <https://doi.org/10.1037/0022-0663.94.4.738>
- Lehmann, T., Rott, B. & Schmidt-Borchering, F. (2019). Promoting Pre-Service Teachers' Integration of Professional Knowledge: Effects of Writing Tasks and Prompts on Learning from Multiple Documents. *Instructional Science*, 47(1), 99–126. <https://doi.org/10.1007/s11251-018-9472-2>
- Leijten, M., van Waes, L., Schrijver, I., Bernolet, S. & Vanghechuchten, L. (2019). Mapping Master's Students' Use of External Sources in Source-Based Writing in L1 and L2. *Studies in Second Language Acquisition*, 41(3), 555–582. <https://doi.org/10.1017/S0272263119000251>
- Lenhard, W. (2019). *Leseverständnis und Lesekompetenz: Grundlagen – Diagnostik – Förderung* (2., aktual. Auflage). Kohlhammer.
- Lenski, S. D. (1998). Strategic Knowledge when Reading in Order to Write. *Reading Psychology*, 19(3), 287–315. <https://doi.org/10.1080/0270271980190303>
- Lenski, S. D. & Johns, J. L. (1997). Patterns of Reading-to-Write. *Reading Research and Instruction*, 37(1), 15–38. <https://doi.org/10.1080/19388079709558252>

- León, J. A. & Escudero, I. (2017). Causal Inferences in Reading Comprehension: State of the Art. In J. A. León & I. Escudero (Hrsg.), *Reading Comprehension in Educational Settings* (S. 63–99). John Benjamins.
- León, J. A., Moreno, J. D., Escudero, I. & Kaakinen, J. K. (2019). Selective Attention to Question-Relevant Text Information Precedes High-Quality Summaries: Evidence from Eye Movements. *Journal of Eye Movement Research*, 12(1), 1–16. <https://doi.org/10.16910/JEMR.12.1.6>
- León, J. A., Moreno, J. D., Escudero, I., Olmos, R., Ruiz, M. & Lorch, R. F. (2019). Specific Relevance Instructions Promote Selective Reading Strategies: Evidences from Eye Tracking and Oral Summaries. *Journal of Research in Reading*, 42(2), 432–453. <https://doi.org/10.1111/1467-9817.12276>
- Leroy, C., Gerjets, P., Oestermeier, U. & Kammerer, Y. (2021). Simultaneous Presentation of Multiple Documents and Text-Highlighting: Online Integrative Processes and Offline Integrated Understanding. *Scientific Studies of Reading*, 25(2), 179–192. <https://doi.org/10.1080/10888438.2020.1784903>
- Leroy, C., Gerjets, P., Oestermeier, U. & Kammerer, Y. (2023). Investigating the Roles of Document Presentation and Reading Interactions on Different Aspects of Multiple Document Comprehension. *International Journal of Human – Computer Interaction*, 39(6), 1327–1340. <https://doi.org/10.1080/10447318.2022.2062854>
- Leroy, C. & Kammerer, Y. (2023). Reading Multiple Documents on a Health-Related Issue: The Roles of a Text-Highlighting Tool and Re-Reading Behaviour in Integrated Understanding. *Behaviour & Information Technology*, 42(14), 2331–2352. <https://doi.org/10.1080/0144929X.2022.2118077>
- Lescarret, C., Le Floch, V., Sakdavong, J.-C., Boucheix, J.-M., Tricot, A. & Amadiou, F. (2023). The Impact of Students' Prior Attitude on the Processing of Conflicting Videos: A Comparison between Middle-School and Undergraduate Students. *European Journal of Psychology of Education*, 38(2), 519–544. <https://doi.org/10.1007/s10212-022-00634-9>
- Lescarret, C., Magnier, J., Le Floch, V., Sakdavong, J.-C., Boucheix, J.-M., Tricot, A. & Amadiou, F. (2023). Expert But Not Persuasive: Middle School Students' Consideration of Source in the Processing of Conflicting Videos. *Infancia y Aprendizaje*, 46(2), 285–321. <https://doi.org/10.1080/02103702.2022.2159616>
- Leu, D. J., Kinzer, C. K., Coiro, J. & Cammack, D. (2004). Toward a Theory of New Literacies Emerging from the Internet and other Information and Communication Technologies. In R. B. Ruddell & N. Unrau (Hrsg.), *Theoretical Models and Processes of Reading* (5. Aufl., S. 1570–1613). International Reading Association.
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J. & Henry, L. A. (2013). New Literacies: A Dual-Level Theory of the Changing Nature of Literacy, Instruction, and Assessment. In D. E. Alvermann, N. Unrau & R. B. Ruddell (Hrsg.), *Theoretical Models and Processes of Reading* (6. Aufl., S. 1150–1181). International Reading Association.
- Levin, J. R., Anglin, G. J. & Carney, R. N. (1987). On Empirically Validating Functions of Pictures in Prose. In D. M. Willows & H. A. Houghton (Hrsg.), *The Psychology of Illustration: Volume 1: Basic Research* (S. 51–85). Springer. https://doi.org/10.1007/978-1-4612-4674-9_2
- Li, J., Ye, H., Tang, Y., Zhou, Z. & Hu, X. (2018). What Are the Effects of Self-Regulation Phases and Strategies for Chinese Students? A Meta-Analysis of Two Decades Research of the Association between Self-Regulation and Academic Performance. *Frontiers in Psychology*, 9, 2434. <https://doi.org/10.3389/fpsyg.2018.02434>
- Li, L.-Y., Tseng, S.-T. & Chen, G.-D. (2016). Effect of Hypertext Highlighting on Browsing, Reading, and Navigational Performance. *Computers in Human Behavior*, 54, 318–325. <https://doi.org/10.1016/j.chb.2015.08.012>
- Li, S. & Lajoie, S. P. (2022). Cognitive Engagement in Self-Regulated Learning: An Integrative Model. *European Journal of Psychology of Education*, 37(3), 833–852. <https://doi.org/10.1007/s10212-021-00565-x>
- Li, Y. & Yan, L. (2024). Which Reading Comprehension Is Better? A Meta-Analysis of the Effect of Paper versus Digital Reading in Recent 20 Years. *Telematics and Informatics Reports*, Artikel 100142. <https://doi.org/10.1016/j.teler.2024.100142>
- Liao, X. & Zhao, P. (2024). The Role of Executive Function in an Integrated Writing Task. *Written Communication*, 41(2), 230–259. <https://doi.org/10.1177/07410883231222950>

- Liao, X., Zhu, X. & Cheong, C. M. (2021). Direct and Indirect Effects of Independent Language Skills on the Integrated Writing Performance of Chinese-Speaking Students with Low Proficiency. *Reading and Writing*, 34(10), 2529–2557. <https://doi.org/10.1007/s11145-021-10152-7>
- Lim, L., Bannert, M., van der Graaf, J., Molenaar, I., Fan, Y., Kilgour, J., Moore, J. & Gašević, D. (2021). Temporal Assessment of Self-Regulated Learning by Mining Students' Think-Aloud Protocols. *Frontiers in Psychology*, 12, Artikel 749749, 1–18. <https://doi.org/10.3389/fpsyg.2021.749749>
- Lin, L.-M. & Zabracky, K. M. (1998). Calibration of Comprehension: Research and Implications for Education and Instruction. *Contemporary Educational Psychology*, 23(4), 345–391. <https://doi.org/10.1006/ceps.1998.0972>
- Linderholm, T., Cong, X. & Zhao, Q. (2008). Differences in Low and High Working-Memory Capacity Readers' Cognitive and Metacognitive Processing Patterns as a Function of Reading for Different Purposes. *Reading Psychology*, 29(1), 61–85. <https://doi.org/10.1080/02702710701568587>
- Linderholm, T., Dobson, J. & Yarbrough, M. B. (2016). The Benefit of Self-Testing and Interleaving for Synthesizing Concepts across Multiple Physiology Texts. *Advances in Physiology Education*, 40(3), 329–334. <https://doi.org/10.1152/advan.00157.2015>
- Linderholm, T., Theriault, D. J. & Kwon, H. (2014). Multiple Science Text Processing: Building Comprehension Skills for College Student Readers. *Reading Psychology*, 35(4), 332–356. <https://doi.org/10.1080/02702711.2012.726696>
- Linderholm, T. & van den Broek, P. W. (2002). The Effects of Reading Purpose and Working Memory Capacity on the Processing of Expository Text. *Journal of Educational Psychology*, 94(4), 778–784. <https://doi.org/10.1037/0022-0663.94.4.778>
- Linderholm, T. & Wilde, A. (2010). College Students' Beliefs about Comprehension when Reading for Different Purposes. *Journal of College Reading and Learning*, 40(2), 7–19. <https://doi.org/10.1080/10790195.2010.10850327>
- List, A. (in Druck). Critique Generation when Reading Two Opposing Newspaper Texts. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-023-00748-8>
- List, A. (2018). Strategies for Comprehending and Integrating Texts and Videos. *Learning and Instruction*, 57, 34–46. <https://doi.org/10.1016/j.learninstruc.2018.01.008>
- List, A. (2019). Drawing Is Integrating: An Examination of Students' Graphic Representations of Multiple Texts. *Reading Psychology*, 40(6), 491–524. <https://doi.org/10.1080/02702711.2019.1629517>
- List, A. (2020). Six Questions Regarding Strategy Use When Learning from Multiple Texts. In D. L. Dinsmore, L. K. Fryer & M. M. Parkinson (Hrsg.), *Handbook of Strategies and Strategic Processing* (S. 119–140). Routledge.
- List, A. (2021). Investigating the Cognitive Affective Engagement Model of Learning from Multiple Texts: A Structural Equation Modeling Approach. *Reading Research Quarterly*, 56(4), 781–817. <https://doi.org/10.1002/rrq.361>
- List, A. (2023a). Demonstrating the Effectiveness of Two Scaffolds for Fostering Students' Domain Perspective Reasoning. *European Journal of Psychology of Education*, 38(4), 1343–1376. <https://doi.org/10.1007/s10212-022-00643-8>
- List, A. (2023b). Social Justice Reasoning when Students Learn about Social Issues Using Multiple Texts. *Discourse Processes*, 60(4-5), 244–272. <https://doi.org/10.1080/0163853X.2023.2197692>
- List, A. & Alexander, P. A. (2017a). Analyzing and Integrating Models of Multiple Text Comprehension. *Educational Psychologist*, 52(3), 143–147. <https://doi.org/10.1080/00461520.2017.1328309>
- List, A. & Alexander, P. A. (2017b). Cognitive Affective Engagement Model of Multiple Source Use. *Educational Psychologist*, 52(3), 182–199. <https://doi.org/10.1080/00461520.2017.1329014>
- List, A. & Alexander, P. A. (2017c). Text Navigation in Multiple Source Use. *Computers in Human Behavior*, 75, 364–375. <https://doi.org/10.1016/j.chb.2017.05.024>
- List, A. & Alexander, P. A. (2018a). Cold and Warm Perspectives on the Cognitive Affective Engagement Model of Multiple Source Use. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 34–54). Routledge.
- List, A. & Alexander, P. A. (2018b). Corroborating Students' Self-Reports of Source Evaluation. *Behavior & Information Technology*, 37(3), 198–216. <https://doi.org/10.1080/0144929X.2018.1430849>
- List, A. & Alexander, P. A. (2019). Toward an Integrated Framework of Multiple Text Use. *Educational Psychologist*, 54(1), 20–39. <https://doi.org/10.1080/00461520.2018.1505514>

- List, A. & Alexander, P. A. (2020). Strategy Use in Learning from Multiple Texts: An Investigation of the Integrative Framework of Learning from Multiple Texts. *Frontiers in Education*, 5(Article 578062), 1–16. <https://doi.org/10.3389/feduc.2020.578062>
- List, A., Alexander, P. A. & Stephens, L. A. (2017). Trust but Verify: Examining the Association between Students' Sourcing Behaviors and Ratings of Text Trustworthiness. *Discourse Processes*, 54(2), 83–104. <https://doi.org/10.1080/0163853X.2016.1174654>
- List, A. & Du, H. (2021). Reasoning beyond History: Examining Students' Strategy Use when Completing a Multiple Text Task Addressing a Controversial Topic in Education. *Reading and Writing*, 34(4), 1003–1048. <https://doi.org/10.1007/s11145-020-10095-5>
- List, A., Du, H. & Lee, H. Y. (2021a). Examining Relation Formation across Consistent and Conflicting Texts. *Discourse Processes*, 58(2), 134–154. <https://doi.org/10.1080/0163853X.2020.1834328>
- List, A., Du, H. & Lee, H. Y. (2021b). How Do Students Integrate Multiple Texts? An Investigation of Top-Down Processing. *European Journal of Psychology of Education*, 46(3), 599–626. <https://doi.org/10.1007/s10212-020-00497-y>
- List, A., Du, H. & Lyu, B. (2022). Examining Undergraduates' Text-Based Evidence Identification, Evaluation, and Use. *Reading and Writing*, 35(5), 1059–1089. <https://doi.org/10.1007/s11145-021-10219-5>
- List, A., Du, H. & Wang, Y. (2019). Understanding Students' Perceptions of Task Assignments. *Contemporary Educational Psychology*, 59, Artikel 101801, 1–16. <https://doi.org/10.1016/j.cedpsych.2019.101801>
- List, A., Du, H., Wang, Y. & Lee, H. Y. (2019). Toward a Typology of Integration: Examining the Documents Model Framework. *Contemporary Educational Psychology*, 58, 228–242. <https://doi.org/10.1016/j.cedpsych.2019.03.003>
- List, A., Grossnickle, E. M. & Alexander, P. A. (2016a). Profiling Students' Multiple Source Use by Question Type. *Reading Psychology*, 37(5), 753–797. <https://doi.org/10.1080/02702711.2015.1111962>
- List, A., Grossnickle, E. M. & Alexander, P. A. (2016b). Undergraduate Students' Justifications for Source Selection in a Digital Academic Context. *Journal of Educational Computing Research*, 54(1), 22–61. <https://doi.org/10.1177/0735633115606659>
- List, A. & Lin, C.-J. (2023). Content and Quantity of Highlights and Annotations Predict Learning from Multiple Digital Texts. *Computers & Education*, 199, Artikel 104791, 1–23. <https://doi.org/10.1016/j.compedu.2023.104791>
- List, A. & Oaxaca, G. S. C. (2024). Comprehension and Critique: An Examination of Students' Evaluations of Information in Texts. *Reading and Writing*, 37(3), 641–671. <https://doi.org/10.1007/s11145-023-10417-3>
- List, A., Oaxaca, G. S. C., Lee, E., Du, H. & Lee, H. Y. (2021). Examining Perceptions, Selections, and Products in Undergraduates' Learning from Multiple Resources. *British Journal of Educational Psychology*, 91(4), 1555–1584. <https://doi.org/10.1111/bjep.12435>
- List, A., Stephens, L. A. & Alexander, P. A. (2019). Examining Interest throughout Multiple Text Use. *Reading and Writing*, 32(2), 307–333. <https://doi.org/10.1007/s11145-018-9863-4>
- List, A. & Sun, Y. (2023). To Clarity and Beyond: Situating Higher-Order, Critical, and Critical-Analytic Thinking in the Literature on Learning from Multiple Texts. *Educational Psychology Review*, 35(2), Artikel 40, 1–63. <https://doi.org/10.1007/s10648-023-09756-y>
- Liu, D., Kirschner, P. A. & Karpinski, A. C. (2017). A Meta-Analysis of the Relationship of Academic Performance and Social Network Site Use among Adolescents and Young Adults. *Computers in Human Behavior*, 77, 148–157. <https://doi.org/10.1016/j.chb.2017.08.039>
- Llorens, A. C. & Cerdán, R. (2012). Assessing the Comprehension of Questions in Task-Oriented Reading. *Revista de Psicodidáctica*, 17(2), 233–251. <https://doi.org/10.1387/Rev.Psicodidact.4496>
- Loderer, K., Pekrun, R. & Lester, J. C. (2020). Beyond Cold Technology: A Systematic Review and Meta-Analysis on Emotions in Technology-Based Learning Environments. *Learning and Instruction*, 70, Artikel 101162, 1–15. <https://doi.org/10.1016/j.learninstruc.2018.08.002>
- Loh, K. K. & Kanai, R. (2015). How Has the Internet Reshaped Human Cognition? *The Neuroscientist*, 22(5), 506–520. <https://doi.org/10.1177/1073858415595005>

- Lombardi, D., Nussbaum, E. M. & Sinatra, G. M. (2016). Plausibility Judgments in Conceptual Change and Epistemic Cognition. *Educational Psychologist*, 51(1), 35–56. <https://doi.org/10.1080/00461520.2015.1113134>
- Londra, F. & Saux, G. (2023). The Effect of Document Source Trustworthiness on the Evaluation and Strategic Use of Embedded Sources when Reading Health Information Online. *Reading Psychology*, 44(6), 623–648. <https://doi.org/10.1080/02702711.2023.2179144>
- Lorch, R. F., Lorch, E. & Klusewitz, M. (1993). College Students' Conditional Knowledge about Reading. *Journal of Educational Psychology*, 85(2), 239–252. <https://doi.org/10.1037/0022-0663.85.2.239>
- Lucassen, T., Muilwijk, R., Noordzij, M. L. & Schraagen, J. M. (2013). Topic Familiarity and Information Skills in Online Credibility Evaluation. *Journal of the American Society for Information Science and Technology*, 64(2), 254–264. <https://doi.org/10.1002/asi.22743>
- Lucassen, T. & Schraagen, J. M. (2011). Factual Accuracy and Trust in Information: The Role of Expertise. *Journal of the American Society for Information Science and Technology*, 62(7), 1232–1242. <https://doi.org/10.1002/asi.21545>
- Luna, M., Villalón, R., Martínez-Álvarez, I. & Mateos, M. (2023). Online Interventions to Help College Students to Improve the Degree of Integration of Their Argumentative Synthesis. *Reading and Writing*, 36(4), 937–963. <https://doi.org/10.1007/s11145-021-10248-0>
- Luna, M., Villalón, R., Mateos, M. & Martín, E. (2020). Improving University Argumentative Writing through an Online Training. *Journal of Writing Research*, 12(1), 233–262. <https://doi.org/10.17239/jowr-2020.12.01.08>
- Luo, L. & Kiewra, K. A. (2019). Soaring to Successful Synthesis Writing. *Journal of Writing Research*, 11(1), 163–209. <https://sig-writing.publication-archive.com/publication/2/1479>
- Luyten, L., Lowyck, J. & Tuerlinckx, F. (2001). Task Perception as a Mediating Variable: A Contribution to the Validation of Instructional Knowledge. *British Journal of Educational Psychology*, 71(2), 203–223. <https://doi.org/10.1348/000709901158488>
- Lyu, B., Peterson, E. M. G. & List, A. (2023). Using PowerPoints to Assess Students' Learning from Multiple Resources. *Contemporary Educational Psychology*, 74, Artikel 102204, 1–16. <https://doi.org/10.1016/j.cedpsych.2023.102204>
- MacArthur, C. A., Traga Philippakos, Z. A., May, H., Potter, A., van Horne, S. & Compello, J. (2023). The Challenges of Writing from Sources in College Developmental Courses: Self-Regulated Strategy Instruction. *Journal of Educational Psychology*, 115(5), 715–731. <https://doi.org/10.1037/edu0000805>
- Macedo-Rouet, M., Braasch, J. L. G., Britt, M. A. & Rouet, J.-F. (2013). Teaching Fourth and Fifth Graders to Evaluate Information Sources during Text Comprehension. *Cognition and Instruction*, 31(2), 204–226. <https://doi.org/10.1080/07370008.2013.769995>
- Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadler, M., Salmerón, L. & Rouet, J.-F. (2019). How Good Is This Page? Benefits and Limits of Prompting on Adolescents' Evaluation of Web Information Quality. *Reading Research Quarterly*, 54(3), 299–321. <https://doi.org/10.1002/rrq.241>
- Macedo-Rouet, M., Salmerón, L., Ros, C., Pérez, A., Stadler, M. & Rouet, J.-F. (2019). Are Frequent Users of Social Network Sites Good Information Evaluators? An Investigation of Adolescents' Sourcing Abilities. *Infancia y Aprendizaje*, 50(5), 1–38. <https://doi.org/10.1080/02103702.2019.1690849>
- Magliano, J. P., Hastings, P. M., Kopp, K., Blaum, D. & Hughes, S. (2018). Computer-Based Assessment of Essays Based on Multiple Documents: Evaluating the Use of Sources. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 502–526). Routledge.
- Magliano, J. P., McCrudden, M. T., Rouet, J.-F. & Sabatini, J. P. (2017). The Modern Reader: Should Changes to How We Read Affect Research and Theory? In M. F. Schober, D. N. Rapp & M. A. Britt (Hrsg.), *The Routledge Handbook of Discourse Processes* (2. Aufl., S. 343–361). Routledge. <https://doi.org/10.4324/9781315687384-18>
- Mahlow, N., Hahnel, C., Kroehne, U., Artelt, C., Goldhammer, F. & Schoor, C. (2022). The Role of Domain-Related Epistemic Beliefs for Mastering Cognitive Requirements in Multiple Document Comprehension. *Learning and Individual Differences*, 94, Artikel 102116, 1–12. <https://doi.org/10.1016/j.lindif.2022.102116>

- Maier, J. & Richter, T. (2013). Text Belief Consistency Effects in the Comprehension of Multiple Texts with Conflicting Information. *Cognition and Instruction*, 31(2), 151–175. <https://doi.org/10.1080/007370008.2013.769997>
- Maier, J. & Richter, T. (2014). Fostering Multiple Text Comprehension: How Metacognitive Strategies and Motivation Moderate the Text-Belief Consistency Effect. *Metacognition and Learning*, 9(1), 51–74. <https://doi.org/10.1007/s11409-013-9111-x>
- Maier, J. & Richter, T. (2016). Effects of Text-Belief Consistency and Reading Task on the Strategic Validation of Multiple Texts. *European Journal of Psychology of Education*, 31(4), 479–497. <https://doi.org/10.1007/s10212-015-0270-9>
- Maier, J., Richter, T. & Britt, M. A. (2018). Cognitive Processes Underlying the Text-Belief Consistency Effect: An Eye-Movement Study. *Applied Cognitive Psychology*, 32(2), 171–185. <https://doi.org/10.1002/acp.3391>
- Maki, R. H. (1998). Test Predictions Over Text Material. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in Educational Theory and Practice* (S. 117–144). Erlbaum.
- Malmberg, J., Järvelä, S. & Kirschner, P. A. (2014). Elementary School Students' Strategic Learning: Does Task-Type Matter? *Metacognition and Learning*, 9(2), 113–136. <https://doi.org/10.1007/s11409-013-9108-5>
- Malmberg, J., Järvenoja, H. & Järvelä, S. (2013). Patterns in Elementary School Students' Strategic Actions in Varying Learning Situations. *Instructional Science*, 41(5), 933–954. <https://doi.org/10.1007/s11251-012-9262-1>
- Mar, R. A., Li, J., Nguyen, A. T. P. & Ta, C. P. (2021). Memory and Comprehension of Narrative versus Expository Texts: A Meta-Analysis. *Psychonomic Bulletin & Review*, 28(3), 732–749. <https://doi.org/10.3758/s13423-020-01853-1>
- Marchionini, G. & White, R. (2007). Find What You Need, Understand What You Find. *International Journal of Human – Computer Interaction*, 23(3), 205–237. <https://doi.org/10.1080/10447310701702352>
- Marker, C., Gnamb, T. & Appel, M. (2018). Active on Facebook and Failing at School? Meta-Analytic Findings on the Relationship Between Online Social Networking Activities and Academic Achievement. *Educational Psychology Review*, 30(3), 651–677. <https://doi.org/10.1007/s10648-017-9430-6>
- Marton, F. & Säljö, R. (1976). On Qualitative Differences in Learning: I. Outcome and Process. *British Journal of Educational Psychology*, 46(1), 4–11. <https://doi.org/10.1111/j.2044-8279.1976.tb02980.x>
- Marttunen, M. & Kiili, C. (2022). Supporting University Students' Argumentative Source-Based Writing. *Written Language & Literacy*, 25(2), 228–252. <https://doi.org/10.1075/wll.00068.mar>
- Marttunen, M., Salminen, T. & Utriainen, J. (2021). Student Evaluations of the Credibility and Argumentation of Online Sources. *The Journal of Educational Research*, 114(3), 294–305. <https://doi.org/10.1080/00220671.2021.1929052>
- Mason, A. E., Braasch, J. L. G., Greenberg, D., Kessler, E. D., Allen, L. K. & McNamara, D. S. (2023). Comprehending Multiple Controversial Texts about Childhood Vaccinations: Topic Beliefs and Integration Instructions. *Reading Psychology*, 44(4), 436–462. <https://doi.org/10.1080/02702711.2022.2156952>
- Mason, L., Ariasi, N. & Boldrin, A. (2011). Epistemic Beliefs in Action: Spontaneous Reflections about Knowledge and Knowing during Online Information Searching and Their Influence on Learning. *Learning and Instruction*, 21(1), 137–151. <https://doi.org/10.1016/j.learninstruc.2010.01.001>
- Mason, L., Boldrin, A. & Ariasi, N. (2010a). Epistemic Metacognition in Context: Evaluating and Learning Online Information. *Metacognition and Learning*, 5(1), 67–90. <https://doi.org/10.1007/s11409-009-9048-2>
- Mason, L., Boldrin, A. & Ariasi, N. (2010b). Searching the Web to Learn about a Controversial Topic: Are Students Epistemically Active? *Instructional Science*, 38(6), 607–633. <https://doi.org/10.1007/s11251-008-9089-y>
- Mason, L. & Florit, E. (2018). Complementary Methods for Assessing Online Processing of Multiple Sources. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 425–446). Routledge.

- Mason, L., Junyent, A. A. & Tornatora, M. C. (2014). Epistemic Evaluation and Comprehension of Web-Source Information on Controversial Science-Related Topics: Effects of a Short-Term Instructional Intervention. *Computers & Education*, 76, 143–157. <https://doi.org/10.1016/j.compedu.2014.03.016>
- Mason, L., Moè, A., Tornatora, M. C. & Ronconi, A. (2022). Promoting Web-Source Evaluation and Comprehension of Conflicting Online Documents: Effects of Classroom Interventions. In P. Limone, R. Di Fuccio & G. A. Toto (Hrsg.), *Psychology, Learning, Technology: First International Conference, PLT 2022* (S. 3–21). Springer. https://doi.org/10.1007/978-3-031-15845-2_1
- Mason, L., Pluchino, P. & Ariasi, N. (2014). Reading Information about a Scientific Phenomenon on Webpages Varying for Reliability: An Eye-Movement Analysis. *Educational Technology Research and Development*, 62(6), 663–685. <https://doi.org/10.1007/s11423-014-9356-3>
- Mason, L., Ronconi, A., Carretti, B., Nardin, S. & Tarchi, C. (2024). Highlighting and Highlighted Information in Text Comprehension and Learning from Digital Reading. *Journal of Computer Assisted Learning*, 40(2), 637–653. <https://doi.org/10.1111/jcal.12903>
- Mason, L., Scirica, F. & Salvi, L. (2006). Effects of Beliefs about Meaning Construction and Task Instructions on Interpretation of Narrative Text. *Contemporary Educational Psychology*, 31(4), 411–437. <https://doi.org/10.1016/j.cedpsych.2005.12.001>
- Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C. & Moè, A. (2018). Internet Source Evaluation: The Role of Implicit Associations and Psychophysiological Self-Regulation. *Computers & Education*, 119, 59–75. <https://doi.org/10.1016/j.compedu.2017.12.009>
- Mason, L., Scrimin, S., Tornatora, M. C. & Zaccoletti, S. (2017). Emotional Reactivity and Comprehension of Multiple Online Texts. *Learning and Individual Differences*, 58, 10–21. <https://doi.org/10.1016/j.lindif.2017.07.002>
- Mason, L., Scrimin, S., Zaccoletti, S., Tornatora, M. C. & Goetz, T. (2018). Webpage Reading: Psychophysiological Correlates of Emotional Arousal and Regulation Predict Multiple-Text Comprehension. *Computers in Human Behavior*, 87, 317–326. <https://doi.org/10.1016/j.chb.2018.05.020>
- Mason, L., Tarchi, C., Ronconi, A., Manzione, L., Latini, N. & Bråten, I. (2022). Do Medium and Context Matter when Learning from Multiple Complementary Digital Texts and Videos? *Instructional Science*, 50(5), 653–679. <https://doi.org/10.1007/s11251-022-09591-8>
- Mason, L., Tornatora, M. C. & Pluchino, P. (2013). Do Fourth Graders Integrate Text and Picture in Processing and Learning from an Illustrated Science Text? Evidence from Eye-Movement Patterns. *Computers & Education*, 60(1), 95–109. <https://doi.org/10.1016/j.compedu.2012.07.011>
- Mason, L., Tornatora, M. C. & Pluchino, P. (2015). Integrative Processing of Verbal and Graphical Information during Re-Reading Predicts Learning from Illustrated Text: An Eye-Movement Study. *Reading and Writing*, 28(6), 851–872. <https://doi.org/10.1007/s11145-015-9552-5>
- Mason, L. & Zaccoletti, S. (2021). Inhibition and Conceptual Learning in Science: a Review of Studies. *Educational Psychology Review*, 33(1), 181–212. <https://doi.org/10.1007/s10648-020-09529-x>
- Mason, L., Zaccoletti, S., Scrimin, S., Tornatora, M. C., Florit, E. & Goetz, T. (2020). Reading with the Eyes and under the Skin: Comprehending Conflicting Digital Texts. *Journal of Computer Assisted Learning*, 36(1), 89–101. <https://doi.org/10.1111/jcal.12399>
- Mateos, M., Cuevas, I., Martín, E., Martín, A., Echeita, G. & Luna, M. (2011). Reading to Write an Argumentation: The Role of Epistemological, Reading and Writing Beliefs. *Journal of Research in Reading*, 34(3), 281–297. <https://doi.org/10.1111/j.1467-9817.2010.01437.x>
- Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I. & González-Lamas, J. (2018). Improving Written Argumentative Synthesis by Teaching the Integration of Conflicting Information from Multiple Sources. *Cognition and Instruction*, 36(2), 119–138. <https://doi.org/10.1080/07370008.2018.1425300>
- Mateos, M., Martín, E., Villalón, R. & Luna, M. (2008). Reading and Writing to Learn in Secondary Education: Online Processing Activity and Written Products in Summarizing and Synthesizing Tasks. *Reading and Writing*, 21(7), 675–697. <https://doi.org/10.1007/s11145-007-9086-6>
- Mateos, M. & Rijlaarsdam, G. (in Druck). Synthesis in Source-Based Writing. In C. A. MacArthur, St. Graham & J. Fitzgerald (Hrsg.), *Handbook of Writing Research* (3. Aufl.). Guilford.
- Mateos, M., Rijlaarsdam, G., Martín, E., Cuevas, I., van den Bergh, H. & Solari, M. (2020). Learning Paths in Synthesis Writing: Which Learning Path Contributes Most to Which Learning Outcome? *Instructional Science*, 48(2), 137–157. <https://doi.org/10.1007/s11251-020-09508-3>

- Mateos, M. & Solé, I. (2009). Synthesising Information from Various Texts: A Study of Procedures and Products at Different Educational Levels. *European Journal of Psychology of Education*, 24(4), 435–451. <https://doi.org/10.1007/BF03178760>
- Mateos, M., Solé, I., Martín, E., Castells, N., Cuevas, I. & González-Lamas, J. (2016). Epistemological and Reading Beliefs Profiles and Their Role in Multiple Text Comprehension. *Electronic Journal of Research in Educational Psychology*, 14(2), 226–252. <https://doi.org/10.25115/ejrep.39.15058>
- Mateos, M., Solé, I., Martín, E., Cuevas, I., Miras, M. & Castells, N. (2014). Writing a Synthesis from Multiple Sources as a Learning Activity. In P. D. Klein, P. Boscolo, L. C. Kirkpatrick & C. Gelati (Hrsg.), *Writing as a Learning Activity* (S. 169–190). Brill.
- May, K. E. & Elder, A. D. (2018). Efficient, Helpful, or Distracting? A Literature Review of Media Multitasking in Relation to Academic Performance. *International Journal of Educational Technology in Higher Education*, 15(1), 1–17. <https://doi.org/10.1186/s41239-018-0096-z>
- Mayer, R. E. (1996). Learning Strategies for Making Sense Out of Expository Text: The SOI Model for Guiding Three Cognitive Processes in Knowledge Construction. *Educational Psychology Review*, 8(4), 357–371. <https://doi.org/10.1007/BF01463939>
- Mayer, R. E. (2002a). Multimedia Learning. In B. H. Ross (Hrsg.), *Psychology of Learning and Motivation: Volume 41* (Bd. 41, S. 85–139). Academic Press. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- Mayer, R. E. (2002b). Rote versus Meaningful Learning. *Theory into Practice*, 41(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4
- Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 43–71). Cambridge University Press.
- Mayer, R. E. (2019a). How Multimedia Can Improve Learning and Instruction. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 460–479). Cambridge University Press. <https://doi.org/10.1017/9781108235631.019>
- Mayer, R. E. (2019b). Taking a New Look at Seductive Details. *Applied Cognitive Psychology*, 33(1), 139–141. <https://doi.org/10.1002/acp.3503>
- Mayer, R. E. (2021). *Multimedia Learning* (3. Aufl.). Cambridge University Press.
- Mayer, R. E. (2022a). Cognitive Theory of Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 57–72). Cambridge University Press. <https://doi.org/10.1017/9781108894333.008>
- Mayer, R. E. (2022b). The Multimedia Principle. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 145–157). Cambridge University Press. <https://doi.org/10.1017/9781108894333.015>
- Mayer, R. E. & Fiorella, L. (2022). Principles for Managing Essential Processing in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 243–260). Cambridge University Press. <https://doi.org/10.1017/9781108894333.025>
- Mayer, R. E. & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6
- McCarthy, K. S. & McNamara, D. S. (2021). The Multidimensional Knowledge in Text Comprehension Framework. *Educational Psychologist*, 56(3), 196–214. <https://doi.org/10.1080/00461520.2021.1872379>
- McCarthy, K. S., Yan, E. F., Allen, L. K., Sonia, A. N., Magliano, J. P. & McNamara, D. S. (2022). On the Basis of Source: Impacts of Individual Differences on Multiple-Document Integrated Reading and Writing Tasks. *Learning and Instruction*, 79, Artikel 101599, 1–20. <https://doi.org/10.1016/j.learninstruc.2022.101599>
- McCrudden, M. T. (2011). Reading Goals and Text Relevance for Assigned Course Readings. *Journal of Education Research*, 4(3), 205–222.
- McCrudden, M. T. (2018). Text Relevance and Multiple Source Use. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 168–183). Routledge.
- McCrudden, M. T. (2020). Processes and Products of Encountering Belief-Related Information. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 191–204). Routledge.

- McCrudden, M. T., Bråten, I. & Salmerón, L. (2023). Learning from Multiple Texts. In R. J. Tierney, F. Rizvi, K. Erkican, I. A. G. Wilkinson & J. M. Parr (Hrsg.), *International Encyclopedia of Education: Volume 6: Learning, Cognition and Human Development* (4. Aufl., S. 353–363). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.14046-1>
- McCrudden, M. T., Huynh, L., Lyu, B. & Kulikowich, J. M. (2021). Bridging Inferences and Learning from Multiple Complementary Texts. *Discourse Processes*, 58(5–6), 529–548. <https://doi.org/10.1080/0163853X.2021.1924586>
- McCrudden, M. T., Huynh, L., Lyu, B., Kulikowich, J. M. & McNamara, D. S. (2024). Coherence Building while Reading Multiple Complementary Documents. *Contemporary Educational Psychology*, 77, Artikel 102266, 1–10. <https://doi.org/10.1016/j.cedpsych.2024.102266>
- McCrudden, M. T., Kulikowich, J. M., Lyu, B. & Huynh, L. (2022). Promoting Integration and Learning from Multiple Complementary Texts. *Journal of Educational Psychology*, 114(8), 1832–1843. <https://doi.org/10.1037/edu0000746>
- McCrudden, M. T., Magliano, J. P. & Schraw, G. (2010). Exploring How Relevance Instructions Affect Personal Reading Intentions, Reading Goals and Text Processing: A Mixed Methods Study. *Contemporary Educational Psychology*, 35(4), 229–241. <https://doi.org/10.1016/j.cedpsych.2009.12.001>
- McCrudden, M. T., Magliano, J. P. & Schraw, G. (2011a). Relevance in Text Comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 1–17). Information Age.
- McCrudden, M. T., Magliano, J. P. & Schraw, G. (2011b). Toward an Integrated View of Relevance in Text Comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 395–414). Information Age.
- McCrudden, M. T. & Schraw, G. (2007). Relevance and Goal-Focusing in Text Processing. *Educational Psychology Review*, 19(2), 113–139. <https://doi.org/10.1007/s10648-006-9010-7>
- McCrudden, M. T. & Schraw, G. (2010). The Effects of Relevance Instructions and Verbal Ability on Text Processing. *Journal of Experimental Education*, 78(1), 96–117. <https://doi.org/10.1080/00220970903224529>
- McCrudden, M. T., Schraw, G. & Kambe, G. (2005). The Effect of Relevance Instructions on Reading Time and Learning. *Journal of Educational Psychology*, 97(1), 88–102. <https://doi.org/10.1037/0022-0663.97.1.88>
- McCrudden, M. T., Stenseth, T., Bråten, I. & Strømsø, H. I. (2016). The Effects of Topic Familiarity, Author Expertise, and Content Relevance on Norwegian Students' Document Selection: A Mixed Methods Study. *Journal of Educational Psychology*, 108(2), 147–162. <https://doi.org/10.1037/edu0000057>
- McDaniel, M. A. & Donnelly, C. M. (1996). Learning with Analogy and Elaborative Interrogation. *Journal of Educational Psychology*, 88(3), 508–519. <https://doi.org/10.1037/0022-0663.88.3.508>
- McGinley, W. (1992). The Role of Reading and Writing while Composing from Sources. *Reading Research Quarterly*, 27(3), 226–248. <https://doi.org/10.2307/747793>
- McGrew, S. (2020). Learning to Evaluate: An Intervention in Civic Online Reasoning. *Computers & Education*, 145, Artikel 103711, 1–13. <https://doi.org/10.1016/j.compedu.2019.103711>
- McGrew, S. (2021a). Challenging Approaches: Sharing and Responding to Weak Digital Heuristics in Class Discussions. *Teaching and Teacher Education*, 108, Artikel 103512, 1–12. <https://doi.org/10.1016/j.tate.2021.103512>
- McGrew, S. (2021b). Skipping the Source and Checking the Contents: An In-Depth Look at Students' Approaches to Web Evaluation. *Computers in the Schools*, 38(2), 75–97. <https://doi.org/10.1080/07380569.2021.1912541>
- McGrew, S. (2022). Internet or Archive? Expertise in Searching for Digital Sources on a Contentious Historical Question. *Cognition and Instruction*, 40(4), 488–516. <https://doi.org/10.1080/07370008.2021.1908288>
- McGrew, S. (2024). Teaching Lateral Reading: Interventions to Help People Read like Fact Checkers. *Current Opinion in Psychology*, 55, Artikel 101737, 1–5. <https://doi.org/10.1016/j.copsy.2023.101737>

- McGrew, S. & Breakstone, J. (2023). Civic Online Reasoning across the Curriculum: Developing and Testing the Efficacy of Digital Literacy Lessons. *AERA Open*, 9(1), 1-14. <https://doi.org/10.1177/23328584231176451>
- McGrew, S., Breakstone, J., Ortega, T., Smith, M. D. & Wineburg, S. S. (2018). Can Students Evaluate Online Sources? Learning from Assessments of Civic Online Reasoning. *Theory and Research in Social Education*, 46(2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>
- McGrew, S. & Byrne, V. L. (2021). Who Is behind This? Preparing High School Students to Evaluate Online Content. *Journal of Research on Technology in Education*, 53(4), 457–475. <https://doi.org/10.1080/15391523.2020.1795956>
- McGrew, S. & Byrne, V. L. (2022). Conversations after Lateral Reading: Supporting Teachers to Focus on Process, Not Content. *Computers & Education*, 185, Artikel 104519, 1–14. <https://doi.org/10.1016/j.compedu.2022.104519>
- McGrew, S. & Chinoy, I. (2022). Fighting Misinformation in College: Students Learn to Search and Evaluate Online Information through Flexible Modules. *Information and Learning Sciences*, 123(1/2), 45–64. <https://doi.org/10.1108/ILS-09-2021-0081>
- McGrew, S., Reynolds, E. C. & Glass, A. C. (in Druck). The Problem with Perspective: Students' and Teachers' Reasoning about Credibility During Discussions of Online Sources. *Cognition and Instruction*. <https://doi.org/10.1080/07370008.2024.2340981>
- McGrew, S., Smith, M. D., Breakstone, J., Ortega, T. & Wineburg, S. S. (2019). Improving University Students' Web Savvy: An Intervention Study. *British Journal of Educational Psychology*, 89(3), 485–500. <https://doi.org/10.1111/bjep.12279>
- McKoon, G. & Ratcliff, R. (1992). Inference during Reading. *Psychological Review*, 99(3), 440–466. <https://doi.org/10.1037/0033-295X.99.3.440>
- McNamara, D. S. (2021). If Integration Is the Keystone of Comprehension: Inferencing Is the Key. *Discourse Processes*, 58(1), 86–91. <https://doi.org/10.1080/0163853X.2020.1788323>
- McNamara, D. S., Jaconiva, M. E. & Allen, L. K. (2016). Higher Order Thinking in Comprehension. In P. Afflerbach (Hrsg.), *Handbook of Individual Differences in Reading: Reader, Text, and Context* (S. 164–176). Routledge.
- McNamara, D. S. & Magliano, J. P. (2009). Toward a Comprehensive Model of Comprehension. In B. H. Ross (Hrsg.), *The Psychology of Learning and Motivation: Volume 51* (S. 297–384). Elsevier. [https://doi.org/10.1016/S0079-7421\(09\)51009-2](https://doi.org/10.1016/S0079-7421(09)51009-2)
- McNamara, D. S., Watanabe, M., Huynh, L., McCarthy, K. S., Allen, L. K. & Magliano, J. P. (2024). Summarizing versus Rereading Multiple Documents. *Contemporary Educational Psychology*, 76, Artikel 102238, 1–11. <https://doi.org/10.1016/j.cedpsych.2023.102238>
- McVay, J. C. & Kane, M. J. (2010). Does Mind Wandering Reflect Executive Function or Executive Failure? Comment on Samuel Wood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, 136(2), 188–197. <https://doi.org/10.1037/a0018298>
- McVay, J. C. & Kane, M. J. (2012). Why Does Working Memory Capacity Predict Variation in Reading Comprehension? On the Influence of Mind Wandering and Executive Attention. *Journal of Experimental Psychology: General*, 141(2), 302–320. <https://doi.org/10.1037/a0025250>
- Meijer, J., Veenman, M. V. J. & van Hout-Wolters, B. H. A. M. (2006). Metacognitive Activities in Text-Studying and Problem-Solving: Development of a Taxonomy. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 12(3), 209–237. <https://doi.org/10.1080/13803610500479991>
- Meneses, L. F. S. (2021). Thinking Critically through Controversial Issues on Digital Media: Dispositions and Key Criteria for Content Evaluation. *Thinking Skills and Creativity*, 42, Artikel 100927, 1–19. <https://doi.org/10.1016/j.tsc.2021.100927>
- Mentel, H., Aust, L., Ehlert, M., Thomas, L. & Souvignier, E. (2022). Führt Lesefreude zu Lesekompetenz? Empirische Befunde zu unterschiedlichen Ansätzen der Leseförderung. In G. Steins, B. Spinath, S. Dutke, M. Roth & M. Limbourg (Hrsg.), *Mythen, Fehlvorstellungen, Fehlkonzepte und Irrtümer in Schule und Unterricht* (S. 135–155). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-36260-77>
- Merchie, E. & van Keer, H. (2014a). Learning from Text in Late Elementary Education. Comparing Think-Aloud Protocols with Self-Reports. *Procedia – Social and Behavioral Sciences*, 112, 489–496. <https://doi.org/10.1016/j.sbspro.2014.01.1193>

- Merchie, E. & van Keer, H. (2014b). Using On-Line and Off-Line Measures to Explore Fifth and Sixth Graders' Text-Learning Strategies and Schematizing Skills. *Learning and Individual Differences*, 32, 193–203. <https://doi.org/10.1016/j.lindif.2014.03.012>
- Merkt, M. & Huff, M. (2020). Does the Position of Source Information for Multiple Documents Matter? Insights from Two Experiments. *Contemporary Educational Psychology*, 62, Artikel 101900, 1–13. <https://doi.org/10.1016/j.cedpsych.2020.101900>
- Merkt, M., Werner, M. & Wagner, W. (2017). Historical Thinking Skills and Mastery of Multiple Document Tasks. *Learning and Individual Differences*, 54, 135–148. <https://doi.org/10.1016/j.lindif.2017.01.021>
- Metcalfe, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18(3), 159–163. <https://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Metzger, M. J. (2007). Making Sense of Credibility on the Web: Models for Evaluating Online Information and Recommendations for Future Research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091. <https://doi.org/10.1002/asi.20672>
- Metzger, M. J. & Flanagin, A. J. (2015). Psychological Approaches to Credibility Assessment Online. In S. S. Sundar (Hrsg.), *The Handbook of the Psychology of Communication Technology* (S. 445–466). John Wiley & Sons. <https://doi.org/10.1002/9781118426456.ch20>
- Miller, B. W. (2015). Using Reading Times and Eye-Movements to Measure Cognitive Engagement. *Educational Psychologist*, 50(1), 31–42. <https://doi.org/10.1080/00461520.2015.1004068>
- Millis, K., Long, D. L., Magliano, J. P. & Wiemer, K. (Hrsg.). (2018a). *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension*. Routledge.
- Millis, K., Long, D. L., Magliano, J. P. & Wiemer, K. (2018b). Introduction. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Hrsg.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (S. XIV–XXI). Routledge.
- Minguela, M., Solé, I. & Pieschl, S. (2015). Flexible Self-Regulated Reading as a Cue for Deep Comprehension: Evidence from Online and Offline Measures. *Reading and Writing*, 28(5), 721–744. <https://doi.org/10.1007/s11145-015-9547-2>
- Mirza, F., Agostinho, S., Tindall-Ford, S., Paas, F. G. W. C. & Chandler, P. (2019). Self-Management of Cognitive Load: Potential and Challenges. In S. Tindall-Ford, S. Agostinho & J. Sweller (Hrsg.), *Advances in Cognitive Load Theory: Rethinking Teaching* (S. 157–167). Routledge.
- Miyake, A. & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, 21(1), 8–14. <https://doi.org/10.1177/0963721411429458>
- Miyamoto, A., Pfost, M. & Artelt, C. (2019). The Relationship between Intrinsic Motivation and Reading Comprehension: Mediating Effects of Reading Amount and Metacognitive Knowledge of Strategy Use. *Scientific Studies of Reading*, 23(6), 445–460. <https://doi.org/10.1080/10888438.2019.1602836>
- Miyatsu, T., Nguyen, K. & McDaniel, M. A. (2018). Five Popular Study Strategies: Their Pitfalls and Optimal Implementations. *Perspectives on Psychological Science*, 13(3), 390–407. <https://doi.org/10.1177/1745691617710510>
- Molden, D. C. & Dweck, C. S. (2006). Finding „Meaning“ in Psychology: A Lay Theories Approach to Self-Regulation, Social Perception, and Social Development. *American Psychologist*, 61(3), 192–203. <https://doi.org/10.1037/0003-066X.61.3.192>
- Mooneyham, B. W. & Schooler, J. W. (2013). The Costs and Benefits of Mind-Wandering: A Review. *Canadian Journal of Experimental Psychology*, 67(1), 11–18. <https://doi.org/10.1037/a0031569>
- Moos, D. C. (2010). Self-Regulated Learning with Hypermedia: Too Much of a Good Thing? *Journal of Educational Multimedia and Hypermedia*, 19(1), 59–77.
- Moos, D. C. (2013). Examining Hypermedia Learning: The Role of Cognitive Load and Self-Regulated Learning. *Journal of Educational Multimedia and Hypermedia*, 22(1), 39–61.
- Moos, D. C. & Azevedo, R. (2006). The Role of Goal Structure in Undergraduates' Use of Self-Regulatory Variables in Two Hypermedia Learning Tasks. *Journal of Educational Multimedia and Hypermedia*, 15(1), 49–86.
- Moos, D. C. & Azevedo, R. (2008a). Exploring the Fluctuation of Motivation and Use of Self-Regulatory Processes during Learning with Hypermedia. *Instructional Science*, 36(3), 203–231. <https://doi.org/10.1007/s11251-007-9028-3>

- Moos, D. C. & Azevedo, R. (2008b). Self-Regulated Learning with Hypermedia: The Role of Prior Domain Knowledge. *Contemporary Educational Psychology*, 33(2), 270–298. <https://doi.org/10.1016/j.cedpsych.2007.03.001>
- Moos, D. C. & Azevedo, R. (2009). Self-Efficacy and Prior Domain Knowledge: To What Extent Does Monitoring Mediate Their Relationship with Hypermedia Learning? *Metacognition and Learning*, 4(3), 197–216. <https://doi.org/10.1007/s11409-009-9045-5>
- Moos, D. C. & Miller, A. (2015). The Cyclical Nature of Self-Regulated Learning Phases: Stable between Learning Tasks? *Journal of Cognitive Education and Psychology*, 14(2), 199–218. <https://doi.org/10.1891/1945-8959.14.2.199>
- Moreno, J. D., León, J. A., Kaakinen, J. K. & Hyönä, J. (2021). Relevance Instructions Combined with Elaborative Interrogation Facilitate Strategic Reading: Evidence from Eye Movements. *Psicologia Educativa*, 27(1), 51–65. <https://doi.org/10.5093/psed2020a20>
- Moreno, R. & Park, B. (2010). Cognitive Load Theory: Historical Development and Relation to Other Theories. In J. L. Plass, R. Moreno & R. Brünken (Hrsg.), *Cognitive Load Theory* (S. 9–28). Cambridge University Press.
- Muijselaar, M. M. L., Swart, N. M., Steenbeek-Planting, E. G., Droop, M., Verhoeven, L. & Jong, P. F. (2017). Developmental Relations between Reading Comprehension and Reading Strategies. *Scientific Studies of Reading*, 21(3), 194–209. <https://doi.org/10.1080/10888438.2017.1278763>
- Muis, K. R. (2007). The Role of Epistemic Beliefs in Self-Regulated Learning. *Educational Psychologist*, 42(3), 173–190. <https://doi.org/10.1080/00461520701416306>
- Muis, K. R., Bendixen, L. D. & Haerle, F. C. (2006). Domain-Generality and Domain-Specificity in Personal Epistemology Research: Philosophical and Empirical Reflections in the Development of a Theoretical Framework. *Educational Psychology Review*, 18(1), 3–54. <https://doi.org/10.1007/s10648-006-9003-6>
- Muis, K. R., Chevrier, M., Denton, C. A. & Losenno, K. M. (2021). Epistemic Emotions and Epistemic Cognition Predict Critical Thinking About Socio-Scientific Issues. *Frontiers in Education*, 6, Artikel 669908, 1–18. <https://doi.org/10.3389/feduc.2021.669908>
- Muis, K. R., Chevrier, M. & Singh, C. A. (2018). The Role of Epistemic Emotions in Personal Epistemology and Self-Regulated Learning. *Educational Psychologist*, 53(3), 165–184. <https://doi.org/10.1080/00461520.2017.1421465>
- Muis, K. R. & Franco, G. M. (2010). Epistemic Profiles and Metacognition: Support for the Consistency Hypothesis. *Metacognition and Learning*, 5(1), 27–45. <https://doi.org/10.1007/s11409-009-9041-9>
- Muis, K. R., Kendeou, P. & Franco, G. M. (2011). Consistent Results with the Consistency Hypothesis? The Effects of Epistemic Beliefs on Metacognitive Processing. *Metacognition and Learning*, 6(1), 45–63. <https://doi.org/10.1007/s11409-010-9066-0>
- Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G. J., Meier, E. & Heddy, B. C. (2015). The Curious Case of Climate Change: Testing a Theoretical Model of Epistemic Beliefs, Epistemic Emotions, and Complex Learning. *Learning and Instruction*, 39, 168–183. <https://doi.org/10.1016/j.learninstruc.2015.06.003>
- Muis, K. R. & Singh, C. A. (2018). The Three Facets of Epistemic Thinking in Self-Regulated Learning. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 434–456). Routledge.
- Müller, B. & Richter, T. (2014). Lesekompetenz. In J. Grabowski (Hrsg.), *Sinn und Unsinn von Kompetenzen: Fähigkeitskonzepte im Bereich von Sprache, Medien und Kultur* (S. 29–49). Barbara Budrich.
- Nagele, F., Katstaller, M. & Greiner, U. (2021). Einschätzungen von Textunterschieden durch Lehramtsstudent*innen: Eine Studie am Beispiel von Texten zum Umgang mit (Leistungs-)Heterogenität. *Zeitschrift für Erziehungswissenschaft*, 24(6), 1503–1524. <https://doi.org/10.1007/s11618-021-01048-x>
- Narvaez, D., van den Broek, P. W. & Ruiz, A. B. (1999). The Influence of Reading Purpose on Inference Generation and Comprehension in Reading. *Journal of Educational Psychology*, 91(3), 488–496. <https://doi.org/10.1037/0022-0663.91.3.488>
- Naumann, A. B., Wechsung, I. & Krems, J. F. (2009). How to Support Learning from Multiple Hyper-text Sources. *Behavior Research Methods*, 41(3), 639–646. <https://doi.org/10.3758/BRM.41.3.639>

- Neal, A., Ballard, T. & Vancouver, J. B. (2017). Dynamic Self-Regulation and Multiple-Goal Pursuit. *Annual Review of Organizational Psychology and Organizational Behavior*, 4(1), 401–423. <https://doi.org/10.1146/annurev-orgpsych-032516-113156>
- Nelson, N. (2001). Discourse Synthesis: The Process and the Product. In R. G. McInnis (Hrsg.), *Discourse Synthesis: Studies in Historical and Contemporary Social Epistemology* (S. 379–396). Praeger.
- Nelson, N. & King, J. R. (2023). Discourse Synthesis: Textual Transformations in Writing from Sources. *Reading and Writing*, 36(4), 769–808. <https://doi.org/10.1007/s11145-021-10243-5>
- Nelson, T. O. & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. In G. H. Bower (Hrsg.), *Psychology of Learning and Motivation: Vol. 26* (S. 125–173). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Neuenhaus, N., Grobe, F. B., Schoor, C. & Artelt, C. (2023). High-Achieving Ninth Grade Students' Self-Reported Strategy Use and Its Relation to Strategic Reading Behavior. *European Journal of Psychology of Education*, 38(4), 1571–1591. <https://doi.org/10.1007/s10212-022-00659-0>
- Neumann, H., Leu, S., McDonough, K., Gil, L. & Crawford, B. (2020). Does the Quality of Source Notes Matter? An Exploratory Study of Source-based Academic Writing. *TESL Canada Journal*, 37(1), 1–21. <https://doi.org/10.18806/tesl.v37i1.1333>
- Newsome, G. L. (1986). The Effects of Reader Perspective and Cognitive Style on Remembering Important Information from Texts. *Journal of Reading Behavior*, 18(2), 117–133. <https://doi.org/10.1080/10862968609547560>
- Nigg, J. T. (2000). On Inhibition/Disinhibition in Developmental Psychopathology: Views from Cognitive and Personality Psychology and a Working Inhibition Taxonomy. *Psychological Bulletin*, 126(2), 220–246. <https://doi.org/10.1037/0033-2909.126.2.220>
- Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., Del Pozo Cruz, B. & Lonsdale, C. (2022). Multimedia Design for Learning: An Overview of Reviews with Meta-Analysis. *Review of Educational Research*, 92(3), 413–454. <https://doi.org/10.3102/00346543211052329>
- Noordman, L., Vonk, W., Cozijn, R. & Frank, S. (2015). Causal Inferences and World Knowledge. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Hrsg.), *Inferences during Reading* (S. 260–289). Cambridge University Press.
- Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T. & Dahl, T. I. (2019). Metacognition in Psychology. *Review of General Psychology*, 23(4), 403–424. <https://doi.org/10.1177/1089268019883821>
- Nussbaum, E. M. (2021). Critical Integrative Argumentation: Toward Complexity in Students' Thinking. *Educational Psychologist*, 56(1), 1–17. <https://doi.org/10.1080/00461520.2020.1845173>
- Nygren, T. & Guath, M. (2019). Swedish Teenagers' Difficulties and Abilities to Determine Digital News Credibility. *Nordicom Review*, 40(1), 23–42. <https://doi.org/10.2478/nor-2019-0002>
- O'Brien, E. J. & Cook, A. E. (2016a). Coherence Threshold and the Continuity of Processing: The RI-Val Model of Comprehension. *Discourse Processes*, 53(5–6), 326–338. <https://doi.org/10.1080/0163853X.2015.1123341>
- O'Brien, E. J. & Cook, A. E. (2016b). Separating the Activation, Integration, and Validation Components of Reading. In B. H. Ross (Hrsg.), *Psychology of Learning and Motivation* (S. 249–276). Elsevier Academic Press.
- O'Reilly, T. P., Sabatini, J. & Wang, Z. (2018). Using Scenario-Based Assessment to Measure Deep Learning. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Hrsg.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (S. 197–208). Routledge.
- O'Reilly, T. P., Symons, S. & MacLachy-Gaudet, H. (1998). A Comparison of Self-Explanation and Elaborative Interrogation. *Contemporary Educational Psychology*, 23(4), 434–445. <https://doi.org/10.1006/ceps.1997.0977>
- Ober, T. M., Brooks, P. J., Homer, B. D. & Rindskopf, D. (2020). Executive Functions and Decoding in Children and Adolescents: A Meta-Analytic Investigation. *Educational Psychology Review*, 32(3), 735–763. <https://doi.org/10.1007/s10648-020-09526-0>
- Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, 2(1), Artikel 36, 1–23. <https://doi.org/10.5334/joc.58>

- Ohtani, K. & Hisasaka, T. (2018). Beyond Intelligence: A Meta-Analytic Review of the Relationship among Metacognition, Intelligence, and Academic Performance. *Metacognition and Learning*, 13(2), 179–212. <https://doi.org/10.1007/s11409-018-9183-8>
- Okkinga, M., van Steensel, R., van Gelderen, A., van Schooten, E., Slegers, P. J. C. & Arends, L. R. (2018). Effectiveness of Reading-Strategy Interventions in Whole Classrooms: A Meta-Analysis. *Educational Psychology Review*, 30(4), 1215–1239. <https://doi.org/10.1007/s10648-018-9445-7>
- Oudega, M. & van den Broek, P. W. (2018). Standards of Coherence in Reading: Variations in Processing and Comprehension of Text. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Hrsg.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (S. 41–51). Routledge.
- Ozgunogor, S. & Guthrie, J. T. (2004). Interactions among Elaborative Interrogation, Knowledge, and Interest in the Process of Constructing Knowledge from Text. *Journal of Educational Psychology*, 96(3), 437–443. <https://doi.org/10.1037/0022-0663.96.3.437>
- Paans, C., Molenaar, I., Segers, E. & Verhoeven, L. (2019). Temporal Variation in Children's Self-Regulated Hypermedia Learning. *Computers in Human Behavior*, 96, 246–258. <https://doi.org/10.1016/j.chb.2018.04.002>
- Paas, F. G. W. C. & Sweller, J. (2014). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 27–42). Cambridge University Press.
- Panadero, E. (2017). A Review of Self-Regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, 8, Artikel 422, 1–28. <https://doi.org/10.3389/fpsyg.2017.00422>
- Panadero, E. & Alonso-Tapia, J. (2014). How Do Students Self-Regulate? Review of Zimmerman's Cyclical Model of Self-Regulated Learning. *Anales de Psicología*, 30(2), 450–462. <https://doi.org/10.6018/analesps.30.2.167221>
- Panadero, E., Jonsson, A. & Botella, J. (2017). Effects of Self-Assessment on Self-Regulated Learning and Self-Efficacy: Four Meta-Analyses. *Educational Research Review*, 22, 74–98. <https://doi.org/10.1016/j.edurev.2017.08.004>
- Panadero, E., Klug, J. & Järvelä, S. (2016). Third Wave of Measurement in the Self-Regulated Learning Field: When Measurement and Intervention Come Hand in Hand. *Scandinavian Journal of Educational Research*, 60(6), 723–735. <https://doi.org/10.1080/00313831.2015.1066436>
- Panizza, F., Ronzani, P., Martini, C., Mattavelli, S., Morisseau, T. & Motterlini, M. (2022). Lateral Reading and Monetary Incentives to Spot Disinformation about Science. *Scientific Reports*, 12(1), Artikel 5678, 1–15. <https://doi.org/10.1038/s41598-022-09168-y>
- Parkinson, M. M. & Dinsmore, D. L. (2018). Multiple Aspects of High School Students' Strategic Processing on Reading Outcomes: The Role of Quantity, Quality, and Conjunctive Strategy Use. *British Journal of Educational Psychology*, 88(1), 42–62. <https://doi.org/10.1111/bjep.12176>
- Parry, D. A. & Le Roux, D. B. (2021). „Cognitive Control in Media Multitaskers“ Ten Years On: A Meta-Analysis. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 15(2), Artikel 7, 1–27. <https://doi.org/10.5817/CP2021-2-7>
- Paul, J., Cerdán, R., Rouet, J.-F. & Stadler, M. (2018). Exploring Fourth Graders' Sourcing Skills. *Infancia y Aprendizaje*, 41(3), 536–580. <https://doi.org/10.1080/02103702.2018.1480458>
- Paul, J., Stadler, M. & Bromme, R. (2019). Effects of a Sourcing Prompt and Conflicts in Reading Materials on Elementary Students' Use of Source Information. *Discourse Processes*, 56(2), 155–169. <https://doi.org/10.1080/0163853X.2017.1402165>
- Payant, C., McDonough, K., Uludag, P. & Lindberg, R. (2019). Predicting Integrated Writing Task Performance: Source Comprehension, Prewriting Planning, and Individual Differences. *Journal of English for Academic Purposes*, 40, 87–97. <https://doi.org/10.1016/j.jeap.2019.06.001>
- Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W. & Tao, S. (2018). A Meta-Analysis on the Relation between Reading and Working Memory. *Psychological Bulletin*, 144(1), 48–76. <https://doi.org/10.1037/bul0000124>
- Peng, P. & Kievit, R. A. (2020). The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. *Child Development Perspectives*, 14(1), 15–20. <https://doi.org/10.1111/cdep.12352>

- Penrose, A. M. (1992). To Write or Not to Write: Effects of Task and Task Interpretation on Learning through Writing. *Written Communication*, 9(4), 465–500. <https://doi.org/10.1177/0741088392009004002>
- Perfetti, C. A. (1997). Sentences, Individual Differences, and Multiple Texts: Three Issues in Text Comprehension. *Discourse Processes*, 23(3), 337–355. <https://doi.org/10.1080/01638539709544996>
- Perfetti, C. A., Rouet, J.-F. & Britt, M. A. (1999). Toward a Theory of Documents Representation. In H. van Oostendorp & S. R. Goldman (Hrsg.), *The Construction of Mental Representations during Reading* (S. 99–122). Lawrence Erlbaum.
- Perfetti, C. A. & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- Perin, D., Keselman, A. & Monopoli, M. (2003). The Academic Writing of Community College Remedial Students: Text and Learner Variables. *Higher Education*, 45(1), 19–42. <https://doi.org/10.1023/A:1021237532056>
- Perner, J. (1991). *Understanding the Representational Mind*. The MIT Press.
- Perry, V., Albeg, L. & Tung, C. (2012). Meta-Analysis of Single-Case Design Research on Self-Regulatory Interventions for Academic Performance. *Journal of Behavioral Education*, 21(3), 217–229. <https://doi.org/10.1007/s10864-012-9156-y>
- Peterson, E. M. G. & Alexander, P. A. (2020). Navigating Print and Digital Sources: Students' Selection, Use, and Integration of Multiple Sources across Mediums. *The Journal of Experimental Education*, 88(1), 27–46. <https://doi.org/10.1080/00220973.2018.1496058>
- Peverly, S. T. & Wolf, A. D. (2019). Note-Taking. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 320–355). Cambridge University Press. <https://doi.org/10.1017/9781108235631.014>
- Philipp, M. (2015). *Lesestrategien: Bedeutung, Formen und Vermittlung*. Beltz Juventa.
- Philipp, M. (2017). *Materialgestütztes Schreiben: Anforderungen, Grundlagen, Vermittlung*. Beltz Juventa.
- Philipp, M. (2018a). *Lesekompetenz bei multiplen Texten: Grundlagen, Prozesse, Didaktik*. Francke.
- Philipp, M. (2018b). *Multiple Modelle des Leseverstehens multipler Texte: Eine Synopse aktueller kognitiver Modellierungen aus lesedidaktischer Perspektive*. https://www.leseforum.ch/sysModules/obxLeseforum/Artikel/646/2018_3_de_philipp.pdf
- Philipp, M. (2019). *Multiple Wege führen nach Rom: Ergebnisse einer quantitativen Sekundäranalyse effektiver Fördermaßnahmen zur Verbesserung der Sourcing- und Integrationsprozesse in der Nutzung multipler Texte*. https://www.leseforum.ch/sysModules/obxLeseforum/Artikel/665/2019_1_de_philipp.pdf
- Philipp, M. (2020a). *Multiple Dokumente verstehen: Theoretische und empirische Perspektiven auf Prozesse und Produkte des Lesens mehrerer Dokumente*. Beltz Juventa.
- Philipp, M. (2020b). Reading into the Future?! Merkmale effektiver Lesefördermaßnahmen mit multiplen Dokumenten mit und ohne Einsatz digitaler Technologie – Ergebnisse eines quantitativen Reviews. In K. Kaspar, M. Becker-Mrotzek, S. Hoffues, J. König & D. Schmeinck (Hrsg.), *Bildung, Schule und Digitalisierung* (S. 83–89). Waxmann.
- Philipp, M. (2021a). *Lesen – Schreiben – Lernen: Prozesse, Strategien und Prinzipien des generativen Lernens*. Beltz. <https://doi.org/10.5281/zenodo.4463835>
- Philipp, M. (2021b). Liest du noch oder sourcst du schon? Sourcing als neuer Bestandteil kompetenten Lesens (multipler, teils digitaler Dokumente). In S. Gailberger & C. Sappok (Hrsg.), *Weiterführende Grundlagenforschung in Lesedidaktik und Leseförderung: Theorie – Empirie – Anwendung* (S. 214–230). Universitätsbibliothek der Ruhr-Universität Bochum.
- Philipp, M. (2021c). *Materialgestütztes Argumentieren – effektiver als andere Schreibenanlässe? Ein quantitativer Review experimenteller Studien*. In Budke, Alexandra, Schäbitz, Frank (Hrsg.), *Argumentieren und Vergleichen: Beiträge aus der Perspektive verschiedener Fachdidaktiken* (S. 153–172). Lit.
- Philipp, M. (2021d). *Schreiben lernen, schreibend lernen: Prinzipien des Aufbaus und der Nutzung von Schreibkompetenz*. Springer VS. <https://doi.org/10.1007/978-3-658-33253-2>

- Philipp, M. (2022a). Ist argumentatives materialgestütztes Schreiben dem informierenden überlegen? Eine Sekundäranalyse experimenteller Studien zu den Effekten der Zieltextsorte auf Verstehensleistungen. In M. Philipp & S. Jambor-Fahlen (Hrsg.), *Lesen: Prozess- und Produktperspektiven von der Wortebene bis zu multiplen Texten* (S. 128–154). Beltz Juventa. <https://doi.org/10.5281/zenodo.5882410>
- Philipp, M. (2022b). Lesen multipler Dokumente zum Zweck des Schreibens: Zusammenhänge von Leseprozessen und Schreibprodukten. In S. Alker-Windbichler, A. Kuhn, B. Lodes & G. Stocker (Hrsg.), *Akademisches Lesen: Medien, Praktiken, Bibliotheken* (S. 25–56). V&R unipress.
- Philipp, M. (2023). Lesekompetenz – neue prozessuale Herausforderungen im kritisch-evaluativen Lesen und ihre Folgen für das Kompetenzkonstrukt. *Leseräume*, 10(9), 1–25.
- Philipp, M. (2024a). *Digitales Lesen fördern*. Vandenhoeck & Ruprecht.
- Philipp, M. (2024b). Lesen 4.0: Drei Thesen zu den Veränderungen des Lesens und den sich daraus ergebenden Implikationen. In M.-O. Carl, M. Jörgens & T. Schulze (Hrsg.), *Literarische Texte lesen – Texte literarisch lesen: Festschrift für Cornelia Rosebrock* (S. 71–93). Metzler. https://doi.org/10.1007/978-3-662-67816-9_6
- Pichert, J. W. & Anderson, R. C. (1977). Taking Different Perspectives on a Story. *Journal of Educational Psychology*, 69(4), 309–315. <https://doi.org/10.1037/0022-0663.69.4.309>
- Pieschl, S. & Sivyer, D. (2021). Secondary Students' Epistemic Thinking and Year as Predictors of Critical Source Evaluation of Internet Blogs. *Computers & Education*, 160, Artikel 104038, 1–16. <https://doi.org/10.1016/j.compedu.2020.104038>
- Pieschl, S., Stahl, E. & Bromme, R. (2008). Epistemological Beliefs and Self-Regulated Learning with Hypertext. *Metacognition and Learning*, 3(1), 17–37. <https://doi.org/10.1007/s11409-007-9008-7>
- Pieschl, S., Stahl, E. & Bromme, R. (2013). Adaptation to Context as Core Component of Self-Regulated Learning: The Example of Complexity and Epistemic Beliefs. In R. Azevedo & V. Aleven (Hrsg.), *International Handbook of Metacognition and Learning Technologies* (Bd. 28, S. 53–65). Springer New York. https://doi.org/10.1007/978-1-4419-5546-3_4
- Pieschl, S., Stahl, E., Murray, T. & Bromme, R. (2012). Is Adaptation to Task Complexity Really Beneficial for Performance? *Learning and Instruction*, 22(4), 281–289. <https://doi.org/10.1016/j.learninstruc.2011.08.005>
- Pieschl, S., Stallmann, F. & Bromme, R. (2014). High School Students' Adaptation of Task Definitions, Goals and Plans to Task Complexity – The Impact of Epistemic Beliefs. *Psihologijske Teme*, 23(1), 31–52.
- Pilotek, L., Karimi, M. N. & Richter, T. (in Druck). Document Language Moderates the Text-Belief Consistency Effect. *Discourse Processes*. <https://doi.org/10.1080/0163853X.2024.2336408>
- Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Handbook of Self-Regulation* (S. 451–502). Academic Press.
- Plakans, L. (2009). The Role of Reading Strategies in Integrated L2 Writing Tasks. *Journal of English for Academic Purposes*, 8(4), 252–266. <https://doi.org/10.1016/j.jeap.2009.05.001>
- Plakans, L. (2010). Independent vs. Integrated Writing Tasks: A Comparison of Task Representation. *TESOL Quarterly*, 44(1), 185–194. <https://doi.org/10.5054/tq.2010.215251>
- Plonsky, L. (2011). The Effectiveness of Second Language Strategy Instruction: A Meta-Analysis. *Language Learning*, 61(4), 993–1038. <https://doi.org/10.1111/j.1467-9922.2011.00663.x>
- Ponce, H. R., Mayer, R. E. & Méndez, E. E. (2022). Effects of Learner-Generated Highlighting and Instructor-Provided Highlighting on Learning from Text: A Meta-Analysis. *Educational Psychology Review*, 34(2), 989–1024. <https://doi.org/10.1007/s10648-021-09654-1>
- Ponce, H. R., Mayer, R. E. & Méndez, E. E. (2023). Effects of Learner-Generated Outlining and Instructor-Provided Outlining on Learning from Text: A Meta-Analysis. *Educational Research Review*, 39, Artikel 100538, 1–17. <https://doi.org/10.1016/j.edurev.2023.100538>
- Popham, M., Counts, J., Ryan, J. B. & Katsiyannis, A. (2018). A Systematic Review of Self-Regulation Strategies to Improve Academic Outcomes of Students with EBD. *Journal of Research in Special Educational Needs*, 18(4), 239–253. <https://doi.org/10.1111/1471-3802.12408>
- Pornpitakpan, C. (2004). The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence. *Journal of Applied Social Psychology*, 34(2), 243–281. <https://doi.org/10.1111/j.1559-1816.2004.tb02547.x>

- Potocki, A., Pereyra, G., Ros, C., Macedo-Rouet, M., Stadtler, M., Salmerón, L. & Rouet, J.-F. (2020). The Development of Source Evaluation Skills during Adolescence: Exploring Different Levels of Source Processing and Their Relationships. *Infancia y Aprendizaje*, 43(1), 19–59. <https://doi.org/10.1080/02103702.2019.1690848>
- Pressley, M. (2000). Development of Grounded Theories of Complex Cognitive Processing: Exhaustive Within- and Between-Study Analyses of Think-Aloud Data. In G. Schraw & J. C. Impara (Hrsg.), *Issues in the Measurement of Metacognition* (S. 261–296). Buros Institute of Mental Measurements and Department of Educational Psychology.
- Pressley, M. & Afflerbach, P. (1995). *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. Lawrence Erlbaum.
- Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L. & Turnure, J. E. (1988). Elaborative Interrogation Facilitates Acquisition of Confusing Facts. *Journal of Educational Psychology*, 80(3), 268–278. <https://doi.org/10.1037/0022-0663.80.3.268>
- Primor, L. & Katzir, T. (2018). Measuring Multiple Text Integration: A Review. *Frontiers in Psychology*, 9, Artikel 2294, 1–16. <https://doi.org/10.3389/fpsyg.2018.02294>
- Primor, L., Yeari, M. & Katzir, T. (2021). Choosing the Right Question: The Effect of Different Question Types on Multiple Text Integration. *Reading and Writing*, 34(6), 1539–1567. <https://doi.org/10.1007/s11145-021-10127-8>
- Prinz, A., Golke, S. & Wittwer, J. (2020a). How Accurately Can Learners Discriminate Their Comprehension of Texts? A Comprehensive Meta-Analysis on Relative Metacomprehension Accuracy and Influencing Factors. *Educational Research Review*, 31, Artikel 100358, 1–31. <https://doi.org/10.1016/j.edurev.2020.100358>
- Prinz, A., Golke, S. & Wittwer, J. (2020b). To What Extent Do Situation-Model-Approach Interventions Improve Relative Metacomprehension Accuracy? Meta-Analytic Insights. *Educational Psychology Review*, 32(4), 917–949. <https://doi.org/10.1007/s10648-020-09558-6>
- Puustinen, M. & Pulkkinen, L. (2001). Models of Self-Regulated Learning: A Review. *Scandinavian Journal of Educational Research*, 45(3), 269–286. <https://doi.org/10.1080/00313830120074206>
- Raccanello, D., Florit, E., Brondino, M., Rodà, A. & Mason, L. (2022). Control and Value Appraisals and Online Multiple-Text Comprehension in Primary School: The Mediating Role of Boredom and the Moderating Role of Word-Reading Fluency. *British Journal of Educational Psychology*, 92(1), 258–279. <https://doi.org/10.1111/bjep.12448>
- Raković, M., Bernacki, M. L., Greene, J. A., Plumley, R. D., Hogan, K. A., Gates, K. M. & Panter, A. T. (2022). Examining the Critical Role of Evaluation and Adaptation in Self-Regulated Learning. *Contemporary Educational Psychology*, 68, Artikel 102027, 1–14. <https://doi.org/10.1016/j.cedpsych.2021.102027>
- Raković, M., Fan, Y., van der Graaf, J., Singh, S., Kilgour, J., Lim, L., Moore, J., Bannert, M., Molenaar, I. & Gašević, D. (2022). Using Learner Trace Data to Understand Metacognitive Processes in Writing from Multiple Sources. In Association for Computing Machinery (Hrsg.), *LAK22: 12th International Learning Analytics and Knowledge Conference* (S. 130–141). Association for Computing Machinery. <https://doi.org/10.1145/3506860.3506876>
- Raković, M., Iqbal, S., Li, T., Fan, Y., Singh, S., Surendrannair, S., Kilgour, J., van der Graaf, J., Lim, L., Molenaar, I., Bannert, M., Moore, J. & Gašević, D. (2023). Harnessing the Potential of Trace Data and Linguistic Analysis to Predict Learner Performance in a Multi-Text Writing Task. *Journal of Computer Assisted Learning*, 39(3), 703–718. <https://doi.org/10.1111/jcal.12769>
- Raković, M. & Winne, P. H. (2022). SR-WMS: A Typology of Self-Regulation in Writing from Multiple Sources. In Y. Wang, S. Joksimović, M. O. Z. San Pedro, J. D. Way & J. Whitmer (Hrsg.), *Social and Emotional Learning and Complex Skills Assessment* (S. 109–129). Springer International Publishing. https://doi.org/10.1007/978-3-031-06333-6_7
- Raković, M., Winne, P. H., Marzouk, Z. & Chang, D. (2021). Automatic Identification of Knowledge-Transforming Content in Argument Essays Developed from Multiple Sources. *Journal of Computer Assisted Learning*, 37(4), 903–924. <https://doi.org/10.1111/jcal.12531>
- Ramsay, C. M. & Sperling, R. A. (2010). Designating Reader Perspective to Increase Comprehension and Interest. *Contemporary Educational Psychology*, 35(3), 215–227. <https://doi.org/10.1016/j.cedpsych.2010.04.005>

- Ramsay, C. M. & Sperling, R. A. (2011). The Relevance of Purpose: Why Purpose Situates Relevance Instruction. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 243–265). Information Age.
- Ramsay, C. M. & Sperling, R. A. (2015). Reading Perspective: Can It Improve Middle School Students' Comprehension of Informational Text? *The Journal of Educational Research*, 108(2), 81–94. <https://doi.org/10.1080/00220671.2013.838538>
- RAND Reading Study Group. (2002). *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. Rand.
- Randall, J. G., Oswald, F. L. & Beier, M. E. (2014). Mind-Wandering, Cognition, and Performance: A Theory-Driven Meta-Analysis of Attention Regulation. *Psychological Bulletin*, 140(6), 1411–1431. <https://doi.org/10.1037/a0037428>
- Rau, M. A. (2020). Cognitive and Socio-Cultural Theories on Competencies and Practices Involved in Learning with Multiple External Representations. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 17–32). Routledge.
- Rayner, K., Chace, K. H., Slattery, T. J. & Ashby, J. (2006). Eye Movements as Reflections of Comprehension Processes in Reading. *Scientific Studies of Reading*, 10(3), 241–255. https://doi.org/10.1207/s1532799xssr1003_3
- Reed, S. K. (2006). Cognitive Architectures for Multimedia Learning. *Educational Psychologist*, 41(2), 87–98. https://doi.org/10.1207/s15326985ep4102_2
- Rey, G. D. (2012). A Review of Research and a Meta-Analysis of the Seductive Detail Effect. *Educational Research Review*, 7(3), 216–237. <https://doi.org/10.1016/j.edurev.2012.05.003>
- Reynolds, G. A. & Perin, D. (2009). A Comparison of Text Structure and Self-Regulated Writing Strategies for Composing from Sources by Middle School Students. *Reading Psychology*, 30(3), 265–300. <https://doi.org/10.1080/02702710802411547>
- Reynolds, R. E. (1992). Selective Attention and Prose Learning: Theoretical and Empirical Research. *Educational Psychology Review*, 4(4), 345–391. <https://doi.org/10.1007/BF01332144>
- Reynolds, R. E. (2000). Attentional Resource Emancipation: Toward Understanding the Interaction of Word Identification and Comprehension Processes in Reading. *Scientific Studies of Reading*, 4(3), 169–195. https://doi.org/10.1207/S1532799XSSR0403_1
- Reynolds, R. E. & Anderson, R. C. (1982). Influence of Questions on the Allocation of Attention during Reading. *Journal of Educational Psychology*, 74(5), 623–632. <https://doi.org/10.1037/0022-0663.74.5.623>
- Reynolds, R. E., Shepard, C., Lapan, R., Kreek, C. & Goetz, E. T. (1990). Differences in the Use of Selective Attention by More Successful and Less Successful Tenth-Grade Readers. *Journal of Educational Psychology*, 82(4), 749–759. <https://doi.org/10.1037/0022-0663.82.4.749>
- Reynolds, R. E., Standiford, S. N. & Anderson, R. C. (1979). Distribution of Reading Time When Questions are Asked About a Restricted Category of Text Information. *Journal of Educational Psychology*, 71(2), 183–190. <https://doi.org/10.1037/0022-0663.71.2.183>
- Reynolds, R. E., Trathen, W., Sawyer, M. L. & Shepard, C. R. (1993). Causal and Epiphenomenal Use of the Selective Attention Strategy in Prose Comprehension. *Contemporary Educational Psychology*, 18(2), 258–278. <https://doi.org/10.1006/ceps.1993.1020>
- Rhodes, M. G. & Tauber, S. K. (2011). The Influence of Delaying Judgments of Learning on Metacognitive Accuracy: A Meta-Analytic Review. *Psychological Bulletin*, 137(1), 131–148. <https://doi.org/10.1037/a0021705>
- Richardson, M., Abraham, C. & Bond, R. (2012). Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin*, 138(2), 353–387. <https://doi.org/10.1037/a0026838>
- Richter, J., Scheiter, K. & Eitel, A. (2016). Signaling Text-Picture Relations in Multimedia Learning: A Comprehensive Meta-Analysis. *Educational Research Review*, 17, 19–36. <https://doi.org/10.1016/j.edurev.2015.12.003>
- Richter, T. (2015). Validation and Comprehension of Text Information: Two Sides of the Same Coin. *Discourse Processes*, 52(5–6), 337–355. <https://doi.org/10.1080/0163853X.2015.1025665>

- Richter, T. & Maier, J. (2017). Comprehension of Multiple Documents with Conflicting Information: A Two-Step Model of Validation. *Educational Psychologist*, 52(3), 148–166. <https://doi.org/10.1080/00461520.2017.1322968>
- Richter, T. & Maier, J. (2018). The Role of Validation in Multiple Documents Comprehension. In J. L. G. Braasch, I. Bräten & M. T. McCrudden (Hrsg.), *Handbook of Multiple Source Use* (S. 151–167). Routledge.
- Richter, T., Münchow, H. & Abendroth, J. (2020). The Role of Validation in Integrating Multiple Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 259–275). Routledge.
- Richter, T. & Schnotz, W. (2018). Textverstehen. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5., überarb. und erw. Aufl., S. 826–837). Beltz.
- Richter, T. & Singer, M. (2017). Discourse Updating: Acquiring and Revising Knowledge through Discourse. In M. F. Schober, D. N. Rapp & M. A. Britt (Hrsg.), *The Routledge Handbook of Discourse Processes* (2. Aufl., S. 167–190). Routledge.
- Richter, T. & Tiffin-Richards, S. P. (2024). Learning as an Epistemic Activity. *Journal for the Study of Education and Development*, 47(1), 3–48. <https://doi.org/10.1177/02103702231224633>
- Rickert, A. & Philipp, M. (2022). Synthesen schreiben – eine textbasierte Längsschnittuntersuchung studentischer Lese- und Schreibperformanz. In M. Philipp & S. Jambor-Fahlen (Hrsg.), *Lesen: Prozess- und Produktperspektiven von der Wortebene bis zu multiplen Texten* (S. 100–127). Beltz Juventa. <https://doi.org/10.5281/zenodo.5882311>
- Rieh, S. Y. & Danielson, D. R. (2007). Credibility: A Multidisciplinary Framework. *Annual Review of Information Science and Technology*, 41(1), 307–364. <https://doi.org/10.1002/aris.2007.1440410114>
- Risemberg, R. (1996). Reading to Write: Self-Regulated Learning Strategies when Writing Essays from Sources. *Reading Research and Instruction*, 35(4), 365–383. <https://doi.org/10.1080/19388079609558221>
- Robson, D. A., Allen, M. S. & Howard, S. J. (2020). Self-Regulation in Childhood as a Predictor of Future Outcomes: A Meta-Analytic Review. *Psychological Bulletin*, 146(4), 324–354. <https://doi.org/10.1037/bul0000227>
- Rodicio, H. G. (2015). Students' Evaluation Strategies in a Web Research Task: Are They Sensitive to Relevance and Reliability? *Journal of Computing in Higher Education*, 27(2), 134–157. <https://doi.org/10.1007/s12528-015-9098-1>
- Rogiers, A., Merchie, E. & van Keer, H. (2019). Learner Profiles in Secondary Education: Occurrence and Relationship with Performance and Student Characteristics. *The Journal of Educational Research*, 112(3), 385–396. <https://doi.org/10.1080/00220671.2018.1538093>
- Rogiers, A., Merchie, E. & van Keer, H. (2020a). Opening the Black Box of Students' Text-Learning Processes: A Process Mining Perspective. *Frontline Learning Research*, 8(3), 40–62. <https://doi.org/10.14786/flr.v8i3.527>
- Rogiers, A., Merchie, E. & van Keer, H. (2020b). What They Say Is What They Do? Comparing Task-Specific Self-Reports, Think-Aloud Protocols, and Study Traces for Measuring Secondary School Students' Text-Learning Strategies. *European Journal of Psychology of Education*, 35(2), 315–332. <https://doi.org/10.1007/s10212-019-00429-5>
- Rothkopf, E. Z. & Billington, M. J. (1975a). Relevance and Similarity of Text Elements to Descriptions of Learning Goals. *Journal of Educational Psychology*, 67(6), 745–750. <https://doi.org/10.1037/0022-0663.67.6.745>
- Rothkopf, E. Z. & Billington, M. J. (1975b). A Two-Factor Model of the Effect of Goal-Descriptive Directions on Learning from Text. *Journal of Educational Psychology*, 67(5), 692–704. <https://doi.org/10.1037/0022-0663.67.5.692>
- Rothkopf, E. Z. & Koether, M. E. (1978). Instructional Effects of Discrepancies in Content and Organization between Study Goals and Information Sources. *Journal of Educational Psychology*, 70(1), 67–71. <https://doi.org/10.1037/0022-0663.70.1.67>
- Rott, S. & Gavin, B. (2015). Comprehending and Learning from Internet Sources: A Conceptual Replication Study of Goldman, Braasch, Wiley, Greasser and Brodowinska (2012). *CALICO Journal*, 32(2), 323–354. <https://doi.org/10.1558/cj.v32i2.25139>

- Rouet, J.-F. (2003). What Was I Looking For? The Influence of Task Specificity and Prior Knowledge on Students' Search Strategies in Hypertext. *Interacting with Computers*, 15(3), 409–428. [https://doi.org/10.1016/S0953-5438\(02\)00064-4](https://doi.org/10.1016/S0953-5438(02)00064-4)
- Rouet, J.-F. (2006). *The Skills of Document Use: From Text Comprehension to Web-Based Learning*. L. Erlbaum Associates.
- Rouet, J.-F., Ayroles, J., Macedo-Rouet, M. & Potocki, A. (2020). Children's Acquisition of Text Search Strategies: The Role of Task Models and Relevance Processes. In W. T. Fu & H. van Oostendorp (Hrsg.), *Understanding and Improving Information Search: A Cognitive Approach* (S. 185–212). Springer International Publishing.
- Rouet, J.-F. & Britt, M. A. (2011). Relevance Processes in Multiple Document Comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 19–52). Information Age.
- Rouet, J.-F. & Britt, M. A. (2022). Multimedia Learning from Multiple Documents. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 521–536). Cambridge University Press. <https://doi.org/10.1017/9781108894333.053>
- Rouet, J.-F., Britt, M. A. & Durik, A. M. (2017). RESOLV: Readers' Representation of Reading Contexts and Tasks. *Educational Psychologist*, 52(3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>
- Rouet, J.-F., Britt, M. A., Gabrielsen, E., Kaakinen, J., Richter, T. & Lennon, M. (2021). PIAAC Cycle 2 Assessment Framework: Literacy. In OECD (Hrsg.), *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies* (S. 39–64). OECD. <https://doi.org/10.1787/7b3bf33b-en>
- Rouet, J.-F., Britt, M. A., Mason, R. A. & Perfetti, C. A. (1996). Using Multiple Sources of Evidence to Reason about History. *Journal of Educational Psychology*, 88(3), 478–493. <https://doi.org/10.1037/0022-0663.88.3.478>
- Rouet, J.-F., Britt, M. A. & Potocki, A. (2019). Multiple-Text Comprehension. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 356–380). Cambridge University Press. <https://doi.org/10.1017/9781108235631.015>
- Rouet, J.-F., Favart, M., Britt, M. A. & Perfetti, C. A. (1997). Studying and Using Multiple Documents in History: Effects of Discipline Expertise. *Cognition and Instruction*, 15(1), 85–106.
- Rouet, J.-F., Le Bigot, L., Pereyra, G. & Britt, M. A. (2016). Whose Story Is This? Discrepancy Triggers Readers' Attention to Source Information in Short Narratives. *Reading and Writing*, 29(8), 1549–1570. <https://doi.org/10.1007/s11145-016-9625-0>
- Rouet, J.-F. & Potocki, A. (2018). From Reading Comprehension to Document Literacy: Learning to Search for, Evaluate and Integrate Information across Texts. *Infancia y Aprendizaje*, 41(3), 415–446. <https://doi.org/10.1080/02103702.2018.1480313>
- Rouet, J.-F., Saux, G., Ros, C., Stadler, M., Vibert, N. & Britt, M. A. (2021). Inside Document Models: Role of Source Attributes in Readers' Integration of Multiple Text Contents. *Discourse Processes*, 58(1), 60–79. <https://doi.org/10.1080/0163853X.2020.1750246>
- Rouet, J.-F., Vidal-Abarca, E., Erboul, A. B. & Millogo, V. (2001). Effects of Information Search Tasks on the Comprehension of Instructional Text. *Discourse Processes*, 31(2), 163–186. https://doi.org/10.1207/S15326950DP3102_03
- Rovers, S. F. E., Clarebout, G., Savelberg, Hans H. C. M., de Bruin, A. B. H. & van Merriënboer, J. J. G. (2019). Granularity Matters: Comparing Different Ways of Measuring Self-Regulated Learning. *Metacognition and Learning*, 14(1), 1–19. <https://doi.org/10.1007/s11409-019-09188-6>
- Royer, J. M., Carlo, M. S., Dufresne, R. & Mestre, J. (1996). The Assessment of Levels of Domain Expertise while Reading. *Cognition and Instruction*, 14(3), 373–408.
- Rukavina, I. & Daneman, M. (1996). Integration and Its Effect on Acquiring Knowledge about Competing Scientific Theories from Text. *Journal of Educational Psychology*, 88(2), 272–287. <https://doi.org/10.1037/0022-0663.88.2.272>
- Sadoski, M. & Paivio, A. (2007). Toward a Unified Theory of Reading. *Scientific Studies of Reading*, 11(4), 337–356. <https://doi.org/10.1080/1088430701530714>

- Salmerón, L., Altamura, L., Delgado, P., Karagiorgi, A. & Vargas, C. (2024). Reading Comprehension on Handheld Devices versus on Paper: A Narrative Review and Meta-Analysis of the Medium Effect and Its Moderators. *Journal of Educational Psychology*, 116(2), 153–172. <https://doi.org/10.1037/edu0000830>
- Salmerón, L., Delgado, P. & Mason, L. (2020). Using Eye-Movement Modelling Examples to Improve Critical Reading of Multiple Webpages on a Conflicting Topic. *Journal of Computer Assisted Learning*, 36(6), 1038–1051. <https://doi.org/10.1111/jcal.12458>
- Salmerón, L., Fajardo, I. & Gómez-Puerta, M. (2019). Selection and Evaluation of Internet Information by Adults with Intellectual Disabilities. *European Journal of Special Needs Education*, 34(3), 272–284. <https://doi.org/10.1080/08856257.2018.1468634>
- Salmerón, L., Gil, L. & Bråten, I. (2018a). Effects of Reading Real versus Print-Out Versions of Multiple Documents on Students' Sourcing and Integrated Understanding. *Contemporary Educational Psychology*, 52, 25–35. <https://doi.org/10.1016/j.cedpsych.2017.12.002>
- Salmerón, L., Gil, L. & Bråten, I. (2018b). Using Eye-Tracking to Assess Sourcing during Multiple Document Reading: A Critical Analysis. *Frontline Learning Research*, 6(3), 105–122.
- Salmerón, L., Gil, L., Bråten, I. & Strømso, H. I. (2010). Comprehension Effects of Signalling Relationships between Documents in Search Engines. *Computers in Human Behavior*, 26(3), 419–426. <https://doi.org/10.1016/j.chb.2009.11.013>
- Salmerón, L., Kammerer, Y. & García-Carrión, P. (2013). Searching the Web for Conflicting Topics: Page and User Factors. *Computers in Human Behavior*, 29(6), 2161–2171. <https://doi.org/10.1016/j.chb.2013.04.034>
- Salmerón, L., Macedo-Rouet, M. & Rouet, J.-F. (2016). Multiple Viewpoints Increase Students' Attention to Source Features in Social Question and Answer Forum Messages. *Journal of the Association for Information Science and Technology*, 67(10), 2404–2419. <https://doi.org/10.1002/asi.23585>
- Salmerón, L., Sampietro, A. & Delgado, P. (2020). Using Internet Videos to Learn about Controversies: Evaluation and Integration of Multiple and Multimodal Documents by Primary School Students. *Computers & Education*, 148, Artikel 103796, 1–14. <https://doi.org/10.1016/j.compedu.2019.103796>
- Salmerón, L., Strømso, H. I., Kammerer, Y., Stadtler, M. & van den Broek, P. W. (2018). Comprehension Processes in Digital Reading. In M. Barzillai, J. Thomson, S. Schroeder & P. W. van den Broek (Hrsg.), *Learning to Read in a Digital World* (S. 91–120). John Benjamins. <https://doi.org/10.1075/swll.17.04sal>
- Sanders, S., Losinski, M., Parks Ennis, R., White, W., Teagarden, J. & Lane, J. (2019). A Meta-Analysis of Self-Regulated Strategy Development Reading Interventions to Improve the Reading Comprehension of Students With Disabilities. *Reading & Writing Quarterly*, 35(4), 339–353. <https://doi.org/10.1080/10573569.2018.1545616>
- Sandoval, W. A., Greene, J. A. & Bråten, I. (2016). Understanding and Promoting Thinking about Knowledge: Origins, Issues, and Future Directions of Research on Epistemic Cognition. *Review of Research in Education*, 40, 457–496. <https://doi.org/10.3102/0091732X16669319>
- Saux, G., Britt, M. A., Le Bigot, L., Vibert, N., Burin, D. & Rouet, J.-F. (2017). Conflicting but Close: Readers' Integration of Information Sources as a Function of Their Disagreement. *Memory & Cognition*, 45(1), 151–167. <https://doi.org/10.3758/s13421-016-0644-5>
- Saux, G., Britt, M. A., Vibert, N. & Rouet, J.-F. (2021). Building Mental Models from Multiple Texts: How Readers Construct Coherence from Inconsistent Sources. *Language and Linguistics Compass*, 15(3), Artikel e12409, 1–19. <https://doi.org/10.1111/lnc3.12409>
- Saux, G., Ros, C., Britt, M. A., Stadtler, M., Burin, D. I. & Rouet, J.-F. (2018). Readers' Selective Recall of Source Features as a Function of Claim Discrepancy and Task Demands. *Discourse Processes*, 55(5-6), 525–544. <https://doi.org/10.1080/0163853X.2018.1463722>
- Saux, G., Vibert, N., Dampuré, J., Burin, D. I., Britt, M. A. & Rouet, J.-F. (2021). From Simple Agents to Information Sources: Readers' Differential Processing of Story Characters as a Function of Story Consistency. *Acta Psychologica*, 212, Artikel 103191, 1–16. <https://doi.org/10.1016/j.actpsy.2020.103191>
- Sbaffi, L. & Rowley, J. (2017). Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *Journal of Medical Internet Research*, 19(6), Artikel e218, 1–17. <https://doi.org/10.2196/jmir.7579>

- Scardamalia, M. & Bereiter, C. (1987). Knowledge Telling and Knowledge Transforming in Written Composition. In S. Rosenberg (Hrsg.), *Advances in Applied Psycholinguistics: Volume 2: Reading, Writing, and Language Learning* (S. 142–175). Cambridge University Press.
- Scharrer, L. & Salmerón, L. (2016). Sourcing in the Reading Process: Introduction to the Special Issue. *Reading and Writing*, 29(8), 1539–1548. <https://doi.org/10.1007/s11145-016-9676-2>
- Scharrer, L., Stadler, M. & Bromme, R. (2019). Judging Scientific Information: Does Source Evaluation Prevent the Seductive Effect of Text Easiness? *Learning and Instruction*, 63, Artikel 101215, 1–16. <https://doi.org/10.1016/j.learninstruc.2019.101215>
- Schellings, G. L. M. & Broekkamp, H. (2011). Signaling Task Awareness in Think-Aloud Protocols from Students Selecting Relevant Information from Text. *Metacognition and Learning*, 6(1), 65–82. <https://doi.org/10.1007/s11409-010-9067-z>
- Schellings, G. L. M. & van Hout-Wolters, B. H. A. M. (1995). Main Points in an Instructional Text, as Identified by Students and by Their Teachers. *Reading Research Quarterly*, 30(4), 742–756. <https://doi.org/10.2307/748196>
- Schellings, G. L. M., van Hout-Wolters, B. H. A. M., Veenman, M. V. J. & Meijer, J. (2013). Assessing Metacognitive Activities: The In-Depth Comparison of a Task-Specific Questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, 28(3), 963–990. <https://doi.org/10.1007/s10212-012-0149-y>
- Schellings, G. L. M., van Hout-Wolters, B. H. A. M. & Vermunt, J. D. (1996a). Individual Differences in Adapting to Three Different Tasks of Selecting Information from Texts. *Contemporary Educational Psychology*, 21(4), 423–446. <https://doi.org/10.1006/ceps.1996.0029>
- Schellings, G. L. M., van Hout-Wolters, B. H. A. M. & Vermunt, J. D. (1996b). Selection of Main Points in Instructional Texts: Influences of Task Demands. *Journal of Literacy Research*, 28(3), 355–378. <https://doi.org/10.1080/10862969609547930>
- Schmidt, S., Zlatkin-Troitschanskaia, O., Roeper, J., Klose, V., Weber, M., Bültmann, A.-K. & Brückner, S. (2020). Undergraduate Students' Critical Online Reasoning-Process Mining Analysis. *Frontiers in Psychology*, 11, Artikel 576273, 1–21. <https://doi.org/10.3389/fpsyg.2020.576273>
- Schneider, M. & Preckel, F. (2017). Variables Associated with Achievement in Higher Education: A Systematic Review of Meta-Analyses. *Psychological Bulletin*, 143(6), 565–600. <https://doi.org/10.1037/bul0000098>
- Schneider, S., Beege, M., Nebel, S. & Rey, G. D. (2018). A Meta-Analysis of How Signaling Affects Learning with Media. *Educational Research Review*, 23, 1–24. <https://doi.org/10.1016/j.edurev.2017.11.001>
- Schnotz, W. (2014). Integrated Model of Text and Picture Comprehension. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 72–103). Cambridge University Press.
- Schnotz, W. (2022). Integrated Model of Text and Picture Comprehension. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 82–99). Cambridge University Press. <https://doi.org/10.1017/9781108894333.010>
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für experimentelle Psychologie*, 46(3), 217–236. <https://doi.org/10.1026//0949-3964.46.3.217>
- Schnotz, W. & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz: Vertiefende Analysen im Rahmen von PISA 2000* (S. 61–99). VS Verlag für Sozialwissenschaften.
- Schnotz, W. & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, 19(4), 469–508. <https://doi.org/10.1007/s10648-007-9053-4>
- Schommer-Aikins, M. (2004). Explaining the Epistemological Belief System: Introducing the Embedded Systemic Model and Coordinated Research Approach. *Educational Psychologist*, 39(1), 19–29. https://doi.org/10.1207/s15326985ep3901_3
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D. & Sayette, M. A. (2011). Meta-Awareness, Perceptual Decoupling and the Wandering Mind. *Trends in Cognitive Sciences*, 15(7), 319–326. <https://doi.org/10.1016/j.tics.2011.05.006>

- Schoor, C., Hahnel, C., Artelt, C., Reimann, D., Kröhne, U. & Goldhammer, F. (2020). Entwicklung und Skalierung eines Tests zur Erfassung des Verständnisses multipler Dokumente von Studierenden. *Diagnostica*, 66(2), 123–135.
- Schoor, C., Rouet, J.-F., Artelt, C., Mahlow, N., Hahnel, C., Kröhne, U. & Goldhammer, F. (2021). Readers' Perceived Task Demands and their Relation to Multiple Document Comprehension Strategies and Outcome. *Learning and Individual Differences*, 88, Artikel 102018, 1–13. <https://doi.org/10.1016/j.lindif.2021.102018>
- Schoor, C., Rouet, J.-F. & Britt, M. A. (2023). Effects of Context and Discrepancy when Reading Multiple Documents. *Reading and Writing*, 36(5), 1111–1143. <https://doi.org/10.1007/s11145-022-10321-2>
- Schoor, C., Rouet, J.-F., Britt, M. A. & Britt, M. A. (2024). Reading for University or for Myself? Effects of Context and Beliefs about Science on College Students' Document Selection. *Journal of Educational Psychology*, 116(3), 317–345. <https://doi.org/10.1037/edu0000849>
- Schraw, G. (2009). Measuring Metacognitive Judgments. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Handbook of Metacognition in Education* (S. 415–429). Routledge.
- Schraw, G. (2013). Conceptual Integration and Measurement of Epistemological and Ontological Beliefs in Educational Research. *ISRN Education*, Artikel 327680. <https://doi.org/10.1155/2013/327680>
- Schraw, G. & Bruning, R. H. (1996). Readers' Implicit Models of Reading. *Reading Research Quarterly*, 31(3), 290–305. <https://doi.org/10.1598/RRQ.31.3.4>
- Schraw, G. & Dennison, R. S. (1994). The Effect of Reader Purpose on Interest and Recall. *Journal of Reading Behavior*, 26(1), 1–18. <https://doi.org/10.1080/10862969409547834>
- Schraw, G. & Moshman, D. (1995). Metacognitive Theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>
- Schraw, G., Wade, S. & Kardash, C. (1993). Interactive Effects of Text-Based and Task-Based Importance on Learning from Text. *Journal of Educational Psychology*, 85(4), 652–661.
- Schroeder, N. L. & Cenkci, A. T. (2018). Spatial Contiguity and Spatial Split-Attention Effects in Multimedia Learning Environments: A Meta-Analysis. *Educational Psychology Review*, 30(3), 679–701. <https://doi.org/10.1007/s10648-018-9435-9>
- Schroeder, N. L. & Cenkci, A. T. (2020). Do Measures of Cognitive Load Explain the Spatial Split-Attention Principle in Multimedia Learning Environments? A Systematic Review. *Journal of Educational Psychology*, 112(2), 254–270. <https://doi.org/10.1037/edu0000372>
- Schroeder, N. L. & Kucera, A. C. (2022). Refutation Text Facilitates Learning: A Meta-Analysis of Between-Subjects Experiments. *Educational Psychology Review*, 34(2), 957–987. <https://doi.org/10.1007/s10648-021-09656-z>
- Chunk, D. H. & DiBenedetto, M. K. (2020). Motivation and Social Cognitive Theory. *Contemporary Educational Psychology*, 60, Artikel 101832, 1–10. <https://doi.org/10.1016/j.cedpsych.2019.101832>
- Schuster, C., Alef, M., Mierwald, M., Brauch, N. & Stadtler, M. (2023). Metacomprehension in Multiple Document Reading: Can Judgment Accuracy Be Fostered by Generating Twitter-Like Summaries? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 55(2-3), 123–135. <https://doi.org/10.1026/0049-8637/a000278>
- Schwabe, A., Lind, F., Kosch, L. & Boomgaarden, H. G. (2022). No Negative Effects of Reading on Screen on Comprehension of Narrative Texts Compared to Print: A Meta-Analysis. *Media Psychology*, 25(6), 779–796. <https://doi.org/10.1080/15213269.2022.2070216>
- Schwinger, M., Trautner, M., Pütz, N., Fabianek, S., Lemmer, G., Lauermann, F. & Wirthwein, L. (2022). Why Do Students Use Strategies That Hurt Their Chances of Academic Success? A Meta-Analysis of Antecedents of Academic Self-Handicapping. *Journal of Educational Psychology*, 114(3), 576–596. <https://doi.org/10.1037/edu0000706>
- Scrimin, S., Patron, E., Ruli, E., Pagui, C. E. K., Altoè, G. & Mason, L. (2018). Dynamic Psychophysiological Correlates of a Learning from Text Episode in Relation to Reading Goals. *Learning and Instruction*, 54, 1–10. <https://doi.org/10.1016/j.learninstruc.2018.01.007>
- Segev-Miller, R. (2007). Cognitive Processes in Discourse Synthesis: The Case of Intertextual Processing Strategies. In M. Torrance, L. van Waes & D. Galbraith (Hrsg.), *Writing and Cognition: Research and Applications* (S. 231–250). Elsevier. [https://doi.org/10.1108/S1572-6304\(2007\)0000020016](https://doi.org/10.1108/S1572-6304(2007)0000020016)

- Seifert, T. L. (1993). Effects of Elaborative Interrogation with Prose Passages. *Journal of Educational Psychology*, 85(4), 642–651. <https://doi.org/10.1037/0022-0663.85.4.642>
- Seifert, T. L. (1994). Enhancing Memory for Main Ideas Using Elaborative Interrogation. *Contemporary Educational Psychology*, 19(3), 360–366. <https://doi.org/10.1006/ceps.1994.1026>
- Seipel, B., Carlson, S. E. & Clinton, V. E. (2017). When Do Comprehender Groups Differ? A Moment-by-Moment Analysis of Think-Aloud Protocols of Good and Poor Comprehenders. *Reading Psychology*, 38(1), 39–70. <https://doi.org/10.1080/02702711.2016.1216489>
- Sencibaugh, J. M. (2007). Meta-Analysis of Reading Comprehension Interventions for Students with Learning Disabilities: Strategies and Implications. *Reading Improvement*, 44(1), 6–22.
- Seufert, T. (2018). The Interplay between Self-Regulation in Learning and Cognitive Load. *Educational Research Review*, 24, 116–129. <https://doi.org/10.1016/j.edurev.2018.03.004>
- Seufert, T. (2020). Building Bridges between Self-Regulation and Cognitive Load – an Invitation for a Broad and Differentiated Attempt. *Educational Psychology Review*, 32(4), 1151–1162. <https://doi.org/10.1007/s10648-020-09574-6>
- Shah, A. K. & Oppenheimer, D. M. (2008). Heuristics Made Easy: An Effort-Reduction Framework. *Psychological Bulletin*, 134(2), 207–222. <https://doi.org/10.1037/0033-2909.134.2.207>
- Sharon, A. J. & Baram-Tsabari, A. (2020). Can Science Literacy Help Individuals Identify Misinformation in Everyday Life? *Science Education*, 104(5), 873–894. <https://doi.org/10.1002/sce.21581>
- Shin, S.-Y. & Ewert, D. (2015). What Accounts for Integrated Reading-to-Write Task Scores? *Language Testing*, 32(2), 259–281. <https://doi.org/10.1177/0265532214560257>
- Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L. & Schneider, M. (2022). Domain-Specific Prior Knowledge and Learning: A Meta-Analysis. *Educational Psychologist*, 57(1), 31–54. <https://doi.org/10.1080/00461520.2021.1939700>
- Sinatra, G. M. & Lombardi, D. (2020). Evaluating Sources of Scientific Evidence and Claims in the Post-Truth Era May Require Reappraising Plausibility Judgments. *Educational Psychologist*, 55(3), 120–131. <https://doi.org/10.1080/00461520.2020.1730181>
- Singer, L. M. & Alexander, P. A. (2017). Reading on Paper and Digitally: What the Past Decades of Empirical Research Reveal. *Review of Educational Research*, 87(6), 1007–1041. <https://doi.org/10.3102/0034654317722961>
- Singer, M. (2013). Validation in Reading Comprehension. *Current Directions in Psychological Science*, 22(5), 361–366. <https://doi.org/10.1177/0963721413495236>
- Singer, M. (2019). Challenges in Processes of Validation and Comprehension. *Discourse Processes*, 56(5–6), 465–483. <https://doi.org/10.1080/0163853X.2019.1598167>
- Singer Trakhman, L. M., Alexander, P. A. & Silverman, A. B. (2018). Profiling Reading in Print and Digital Mediums. *Learning and Instruction*, 57, 5–17. <https://doi.org/10.1016/j.learninstruc.2018.04.001>
- Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L. & Macnamara, B. N. (2018). To What Extent and under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses. *Psychological Science*, 29(4), 549–571. <https://doi.org/10.1177/0956797617739704>
- Skulmowski, A. & Xu, K. M. (2022). Understanding Cognitive Load in Digital and Online Learning: A New Perspective on Extraneous Cognitive Load. *Educational Psychology Review*, 34(1), 171–196. <https://doi.org/10.1007/s10648-021-09624-7>
- Slotte, V. & Lonka, K. (1999). Spontaneous Concept Maps Aiding the Understanding of Scientific Concepts. *International Journal of Science Education*, 21(5), 515–531. <https://doi.org/10.1080/095006999290552>
- Smallwood, J. (2011). Mind-Wandering while Reading: Attentional Decoupling, Mindless Reading and the Cascade Model of Inattention. *Language and Linguistics Compass*, 5(2), 63–77. <https://doi.org/10.1111/j.1749-818X.2010.00263.x>
- Smallwood, J. (2013). Distinguishing How from Why the Mind Wanders: A Process – Occurrence Framework for Self-Generated Mental Activity. *Psychological Bulletin*, 139(3), 519–535. <https://doi.org/10.1037/a0030010>
- Smallwood, J. & Schooler, J. W. (2006). The Restless Mind. *Psychological Bulletin*, 132(6), 946–958. <https://doi.org/10.1037/0033-2909.132.6.946>

- Smallwood, J. & Schooler, J. W. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology*, 66(1), 487–518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Smith, B. L., Holliday, W. G. & Austin, H. W. (2010). Students' Comprehension of Science Textbooks Using a Question-Based Reading Strategy. *Journal of Research in Science Teaching*, 47(4), 363–379. <https://doi.org/10.1002/tea.20378>
- Smith, R., Snow, P., Serry, T. & Hammond, L. (2021). The Role of Background Knowledge in Reading Comprehension: A Critical Review. *Reading Psychology*, 42(3), 214–240. <https://doi.org/10.1080/02702711.2021.1888348>
- Solé, I., Miras, M., Castells, N., Espino, S. & Minguela, M. (2013). Integrating Information: An Analysis of the Processes Involved and the Products Generated in a Written Synthesis Task. *Written Communication*, 30(1), 63–90. <https://doi.org/10.1177/0741088312466532>
- Sonnenberg, C. & Bannert, M. (2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100. <https://doi.org/10.18608/jla.2015.21.5>
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G. & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, 25(4), 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M. & Lonigan, C. J. (2021). Relations between Executive Functions and Academic Outcomes in Elementary School Children: A Meta-Analysis. *Psychological Bulletin*, 147(4), 329–351. <https://doi.org/10.1037/bul0000322>
- Spivey, N. N. (1990). Transforming Texts: Constructive Processes in Reading and Writing. *Written Communication*, 7(2), 256–287. <https://doi.org/10.1177/0741088390007002004>
- Spivey, N. N. (1991). The Shaping of Meaning: Options in Writing the Comparison. *Research in the Teaching of English*, 25(4), 390–418.
- Spivey, N. N. (1992). Discourse Synthesis: Creating Texts from Texts. In J. R. Hayes (Hrsg.), *Reading Empirical Research Studies: The Rhetoric of Research* (S. 469–512). Erlbaum.
- Spivey, N. N. & King, J. R. (1989). Readers as Writers Composing from Sources. *Reading Research Quarterly*, 24(1), 7–26. <http://www.jstor.org/stable/748008>
- Spörer, N. & Brunstein, J. C. (2006). Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren: Ein Überblick zum Stand der Forschung. *Zeitschrift für Pädagogische Psychologie*, 20(3), 147–160. <https://doi.org/10.1024/1010-0652.20.3.147>
- Stadtler, M. (2017). The Art of Reading in a Knowledge Society: Commentary on the Special Issue on Models of Multiple Text Comprehension. *Educational Psychologist*, 52(3), 225–231. <https://doi.org/10.1080/00461520.2017.1322969>
- Stadtler, M. & Bromme, R. (2014). The Content – Source Integration Model: A Taxonomic Description of How Readers Comprehend Conflicting Scientific Information. In D. N. Rapp & J. L. G. Braasch (Hrsg.), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (S. 379–402). MIT Press.
- Stadtler, M., Scharrer, L. & Bromme, R. (2013). How Do Readers Explain the Occurrence of Conflicts in Science Texts? Effects of Presentation Format and Source Expertise. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35, 3448–3453.
- Stadtler, M., Scharrer, L. & Bromme, R. (2020). How Relevance Affects Understanding of Conflicts between Multiple Documents: An Eye-Tracking Study. *Reading Research Quarterly*, 55(4), 625–641.
- Stadtler, M., Scharrer, L., Brummernhenrich, B. & Bromme, R. (2013). Dealing with Uncertainty: Readers' Memory for and Use of Conflicting Information from Science Texts as Function of Presentation Format and Source Expertise. *Cognition and Instruction*, 31(2), 130–150. <https://doi.org/10.1080/07370008.2013.769996>
- Stadtler, M., Scharrer, L., Macedo-Rouet, M., Rouet, J.-F. & Bromme, R. (2016). Improving Vocational Students' Consideration of Source Information when Deciding about Science Controversies. *Reading and Writing*, 29(4), 705–729. <https://doi.org/10.1007/s11145-016-9623-2>
- Stadtler, M., Scharrer, L., Skodzik, T. & Bromme, R. (2014). Comprehending Multiple Documents on Scientific Controversies: Effects of Reading Goals and Signaling Rhetorical Relationships. *Discourse Processes*, 51(1–2), 93–116. <https://doi.org/10.1080/0163853X.2013.855535>

- Stahl, E., Pieschl, S. & Bromme, R. (2006). Task Complexity, Epistemological Beliefs and Metacognitive Calibration: An Exploratory Study. *Journal of Educational Computing Research*, 35(4), 319–338. <https://doi.org/10.2190/1266-0413-387K-7J51>
- Stahl, S. A., Hynd, C. R., Britton, B. K., McNish, M. M. & Bosquet, D. (1996). What Happens when Students Read Multiple Source Documents in History? *Reading Research Quarterly*, 31(4), 430–456.
- Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W. & Strømso, H. I. (2019). Direct and Indirect Effects of Textual and Individual Factors on Source-Content Integration when Reading about a Socio-Scientific Issue. *Reading and Writing*, 32(2), 335–356. <https://doi.org/10.1007/s11145-018-9868-z>
- Stang Lund, E., Bråten, I., Brante, E. W. & Strømso, H. I. (2017). Memory for Textual Conflicts Predicts Sourcing when Adolescents Read Multiple Expository Texts. *Reading Psychology*, 38(4), 417–437. <https://doi.org/10.1080/02702711.2016.1278417>
- Steffens, B., Britt, M. A., Braasch, J. L. G., Strømso, H. I. & Bråten, I. (2014). Memory for Scientific Arguments and Their Sources: Claim-Evidence Consistency Matters. *Discourse Processes*, 51(1–2), 117–142. <https://doi.org/10.1080/0163853X.2013.855868>
- Stenseth, T., Strømso, H. I. & Strømso, H. I. (2019). To Read or Not to Read: A Qualitative Study of Students' Justifications for Document Selection in Task-Oriented Reading. *Scandinavian Journal of Educational Research*, 63(5), 771–788. <https://doi.org/10.1080/00313831.2018.1434828>
- Stine-Morrow, E. A. L., Miller, L. M. S. & Hertzog, C. (2006). Aging and Self-Regulated Language Processing. *Psychological Bulletin*, 132(4), 582–606. <https://doi.org/10.1037/0033-2909.132.4.582>
- Strømso, H. I. & Bråten, I. (2009). Beliefs about Knowledge and Knowing and Multiple-Text Comprehension among Upper Secondary Students. *Educational Psychology*, 29(4), 425–445. <https://doi.org/10.1080/01443410903046864>
- Strømso, H. I. & Bråten, I. (2010). Learning from Multiple Information Sources. In P. Peterson, E. Baker & B. McGaw (Hrsg.), *International Encyclopedia of Education* (3. Aufl., S. 191–196). Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00496-6>
- Strømso, H. I. & Bråten, I. (2014). Students' Sourcing while Reading and Writing from Multiple Web Documents. *Nordic Journal of Digital Literacy*, 9(2), 92–111.
- Strømso, H. I. & Bråten, I. (2017). Students' Trust in Research-Based Results about Potential Health Risks Presented in Popular Media. *Bulletin of Science, Technology & Society*, 37(1), 3–14. <https://doi.org/10.1177/0270467617740495>
- Strømso, H. I., Bråten, I., Anmarkrud, Ø. & Ferguson, L. E. (2016). Beliefs about Justification for Knowing when Ethnic Majority and Ethnic Minority Students Read Multiple Conflicting Documents. *Educational Psychology*, 36(4), 638–657. <https://doi.org/10.1080/01443410.2014.920080>
- Strømso, H. I., Bråten, I. & Brante, E. W. (2020). Profiles of Warm Engagement and Cold Evaluation in Multiple-Document Comprehension. *Reading and Writing*, 33(9), 2337–2359. <https://doi.org/10.1007/s11145-020-10041-5>
- Strømso, H. I., Bråten, I. & Britt, M. A. (2010). Reading Multiple Texts about Climate Change: The Relationship between Memory for Sources and Text Comprehension. *Learning and Instruction*, 20(3), 192–204. <https://doi.org/10.1016/j.learninstruc.2009.02.001>
- Strømso, H. I., Bråten, I. & Britt, M. A. (2011). Do Students' Beliefs about Knowledge and Knowing Predict Their Judgement of Texts' Trustworthiness? *Educational Psychology*, 31(2), 177–206. <https://doi.org/10.1080/01443410.2010.538039>
- Strømso, H. I., Bråten, I., Britt, M. A. & Ferguson, L. E. (2013). Spontaneous Sourcing among Students Reading Multiple Documents. *Cognition and Instruction*, 31(2), 176–203. <https://doi.org/10.1080/07370008.2013.769994>
- Strømso, H. I., Bråten, I. & Samuelstuen, M. S. (2008). Dimensions of Topic-Specific Epistemological Beliefs as Predictors of Multiple Text Understanding. *Learning and Instruction*, 18(6), 513–527. <https://doi.org/10.1016/j.learninstruc.2007.11.001>
- Strømso, H. I., Bråten, I. & Stenseth, T. (2017). The Role of Students' Prior Topic Beliefs in Recall and Evaluation of Information from Texts on Socio-Scientific Issues. *Nordic Psychology*, 69(3), 127–142. <https://doi.org/10.1080/19012276.2016.1198270>

- Strømso, H. I. & Kammerer, Y. (2016). Epistemic Cognition and Reading for Understanding in the Internet Age. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of Epistemic Cognition* (S. 230–246). Routledge.
- Sun, Y., Zhang, Y., Gwizdka, J. & Trace, C. B. (2019). Consumer Evaluation of the Quality of Online Health Information: Systematic Literature Review of Relevant Criteria and Indicators. *Journal of Medical Internet Research*, 21(5), Artikel e12522, 1–22. <https://doi.org/10.2196/12522>
- Sun, Y., Wang, J., Dong, Y., Zheng, H., Yang, J., Zhao, Y. & Dong, W. (2021). The Relationship between Reading Strategy and Reading Comprehension: A Meta-Analysis. *Frontiers in Psychology*, 12, Artikel 2494, 1–11. <https://doi.org/10.3389/fpsyg.2021.635289>
- Sundararajan, N. & Adesope, O. (2020). Keep it Coherent: A Meta-Analysis of the Seductive Details Effect. *Educational Psychology Review*, 32(3), 707–734. <https://doi.org/10.1007/s10648-020-09522-4>
- Sutcliffe, A. & Ennis, M. (1998). Towards a Cognitive Theory of Information Retrieval. *Interacting with Computers*, 10(3), 321–351. [https://doi.org/10.1016/S0953-5438\(98\)00013-7](https://doi.org/10.1016/S0953-5438(98)00013-7)
- Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Sweller, J. (2020). Cognitive Load Theory and Educational Technology. *Educational Technology Research and Development*, 68(1), 1–16. <https://doi.org/10.1007/s11423-019-09701-3>
- Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer. <https://doi.org/10.1007/978-1-4419-8126-4>
- Sweller, J. & Paas, F. G. W. C. (2017). Should Self-Regulated Learning Be Integrated with Cognitive Load Theory? A Commentary. *Learning and Instruction*, 51, 85–89. <https://doi.org/10.1016/j.learninstruc.2017.05.005>
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Tarchi, C. (2019). Identifying Fake News through Trustworthiness Judgements of Documents. *Cultura y Educación*, 65(3), 1–38. <https://doi.org/10.1080/11356405.2019.1597442>
- Tarchi, C. (2021a). Effects of Think-Aloud on Students' Multiple-Documents Comprehension. *Applied Cognitive Psychology*, 35(2), 526–537. <https://doi.org/10.1002/acp.3782>
- Tarchi, C. (2021b). Prompting Readers to Plan Might Negatively Affect Their Comprehension of Multiple Documents. *Journal of College Reading and Learning*, 51(2), 131–152. <https://doi.org/10.1080/10790195.2020.1823910>
- Tarchi, C. (in Druck). Promoting Multiple-Text Comprehension through Thinking Dispositions: A Randomized Controlled Trial. *Discourse Processes*. <https://doi.org/10.1080/0163853X.2024.2355059>
- Tarchi, C., Casado-Ledesma, L., Sanna, G. & Conti, M. (in Druck). The Relationship between Theory of Mind and Multiple-Documents Comprehension in University Students: Preliminary Research Findings. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-023-00755-9>
- Tarchi, C. & Mason, L. (2020). Effects of Critical Thinking on Multiple-Document Comprehension. *European Journal of Psychology of Education*, 35(2), 289–313. <https://doi.org/10.1007/s10212-019-00426-8>
- Tarchi, C. & Villalón, R. (2021). The Influence of Thinking Dispositions on Integration and Recall of Multiple Texts. *British Journal of Educational Psychology*, 91(4), 1498–1516. <https://doi.org/10.1111/bjep.12432>
- Tarchi, C. & Villalón, R. (2022). Fostering University Students' Written Argumentation via Recursive Reading: A Randomized Controlled Trial. *Journal of College Reading and Learning*, 52(1), 42–63. <https://doi.org/10.1080/10790195.2022.2021771>
- Tarchi, C., Villalón, R. & Lamanda, E. (2022). Investigating the Effect of Actively-Openminded Thinking on Source-Based Writing: A Randomized Control Trial. *Thinking Skills and Creativity*, 46, Artikel 101128, 1–11. <https://doi.org/10.1016/j.tsc.2022.101128>
- Tarchi, C., Villalón, R., Vandermeulen, N., Casado-Ledesma, L. & Fallaci, A. P. (in Druck). Recursivity in Source-Based Writing: A Process Analysis. *Reading and Writing*. <https://doi.org/10.1007/s11145-023-10482-8>

- Tarricone, P. (2011). *The Taxonomy of Metacognition*. Psychology Press. <https://doi.org/10.4324/9780203830529>
- Taub, M. & Azevedo, R. (2019). How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? *International Journal of Artificial Intelligence in Education*, 29(1), 1–28. <https://doi.org/10.1007/s40593-018-0165-4>
- Taub, M., Azevedo, R., Bouchet, F. & Khosravifar, B. (2014). Can the Use of Cognitive and Metacognitive Self-Regulated Learning Strategies Be Predicted by Learners' Levels of Prior Knowledge in Hypermedia-Learning Environments? *Computers in Human Behavior*, 39, 356–367. <https://doi.org/10.1016/j.chb.2014.07.018>
- Theobald, M. (2021). Self-Regulated Learning Training Programs Enhance University Students' Academic Performance, Self-Regulated Learning Strategies, and Motivation: A Meta-Analysis. *Contemporary Educational Psychology*, 66, Artikel 101976, 1–19. <https://doi.org/10.1016/j.cedpsych.2021.101976>
- Thiede, K. W. & de Bruin, A. B. H. (2018). Self-Regulated Learning in Reading. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 124–137). Routledge.
- Thiede, K. W., Griffin, T. D., Wiley, J. & Redford, J. S. (2009). Metacognitive Monitoring during and after Reading. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Handbook of Metacognition in Education* (S. 85–106). Routledge.
- Thomm, E., Barzilai, S. & Bromme, R. (2017). Why Do Experts Disagree? The Role of Conflict Topics and Epistemic Perspectives in Conflict Explanations. *Learning and Instruction*, 52, 15–26. <https://doi.org/10.1016/j.learninstruc.2017.03.008>
- Thomm, E. & Bromme, R. (2016). How Source Information Shapes Lay Interpretations of Science Conflicts: Interplay between Sourcing, Conflict Explanation, Source Evaluation, and Claim Evaluation. *Reading and Writing*, 29(8), 1629–1652. <https://doi.org/10.1007/s11145-016-9638-8>
- Thomm, E., Hentschke, J. & Bromme, R. (2015). The Explaining Conflicting Scientific Claims (ECSC) Questionnaire: Measuring Laypersons' Explanations for Conflicts in Science. *Learning and Individual Differences*, 37, 139–152. <https://doi.org/10.1016/j.lindif.2014.12.001>
- Thon, F. M. & Jucks, R. (2017). Believing in Expertise: How Authors' Credentials and Language Use Influence the Credibility of Online Health Information. *Health Communication*, 32(7), 828–836. <https://doi.org/10.1080/10410236.2016.1172296>
- Tise, J., Follmer, D. J. & Sperling, R. A. (2019). A Review of the Self-Regulation Strategy Inventory – Self-Report (SRSI-SR). *Psychology*, 10(3), 305–319. <https://doi.org/10.4236/psych.2019.103022>
- Toste, J., Didion, L., Peng, P., Filderman, M. & McClelland, A. (2020). A Meta-Analytic Review of the Relations between Motivation and Reading Achievement for K – 12 Students. *Review of Educational Research*, 90(3), 420–456. <https://doi.org/10.3102/0034654320919352>
- Trevors, G. J., Duffy, M. & Azevedo, R. (2014). Note-Taking within MetaTutor: Interactions between an Intelligent Tutoring System and Prior Knowledge on Note-Taking and Learning. *Educational Technology Research and Development*, 62(5), 507–528. <https://doi.org/10.1007/s11423-014-9343-8>
- Trevors, G. J., Kendeou, P., Bråten, I. & Braasch, J. L. G. (2017). Adolescents' Epistemic Profiles in the Service of Knowledge Revision. *Contemporary Educational Psychology*, 49, 107–120. <https://doi.org/10.1016/j.cedpsych.2017.01.005>
- Trevors, G. J. & Muis, K. R. (2015). Effects of Text Structure, Reading Goals and Epistemic Beliefs on Conceptual Change. *Journal of Research in Reading*, 38(4), 361–386. <https://doi.org/10.1111/1467-9817.12031>
- Trites, L. & McGroarty, M. (2005). Reading to Learn and Reading to Integrate: New Tasks for Reading Comprehension Tests? *Language Testing*, 22(2), 174–210. <https://doi.org/10.1191/0265532205lt2990a>
- Tze, V. M. C., Daniels, L. M. & Klassen, R. M. (2016). Evaluating the Relationship between Boredom and Academic Outcomes: A Meta-Analysis. *Educational Psychology Review*, 28(1), 119–144. <https://doi.org/10.1007/s10648-015-9301-y>

- Ulyshen, T. Z., Koehler, M. J. & Gao, F. (2015). Understanding the Connection between Epistemic Beliefs and Internet Searching. *Journal of Educational Computing Research*, 53(3), 345–383. <https://doi.org/10.1177/0735633115599604>
- Urduan, T. & Kaplan, A. (2020). The Origins, Evolution, and Future Directions of Achievement Goal Theory. *Contemporary Educational Psychology*, 61, Artikel 101862, 1–10. <https://doi.org/10.1016/j.cedpsych.2020.101862>
- Vakkari, P. (2016). Searching as Learning: A Systematization Based on Literature. *Journal of Information Science*, 42(1), 7–18. <https://doi.org/10.1177/0165551515615833>
- Valcke, M. (2002). Cognitive Load: Updating the Theory? *Learning and Instruction*, 12(1), 147–154. [https://doi.org/10.1016/S0959-4752\(01\)00022-6](https://doi.org/10.1016/S0959-4752(01)00022-6)
- van den Broek, P. & Kendeou, P. (2015). Building Coherence in Web-Based and Other Non-Traditional Reading Environments: Cognitive Opportunities and Challenges. In R. J. Spiro, M. Deschryver, M. S. Hagerman, P. M. Morsink & P. Thompson (Hrsg.), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices* (S. 104–114). Routledge.
- van den Broek, P. W., Beker, K. & Oudega, M. (2015). Inference Generation in Text Comprehension: Automatic and Strategic Processes in the Construction of a Mental Representation. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Hrsg.), *Inferences during Reading* (S. 94–121). Cambridge University Press.
- van den Broek, P. W., Bohn-Gettler, C. M., Kendeou, P., Carlson, S. & White, M. J. (2011). When a Reader Meets a Text: The Role of Standards of Coherence in Reading Comprehension. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 123–139). Information Age.
- van den Broek, P. W., Fletcher, C. & Risden, K. (1993). Investigations of Inferential Processes in Reading: A Theoretical and Methodological Integration. *Discourse Processes*, 16(1–2), 169–180. <https://doi.org/10.1080/01638539309544835>
- van den Broek, P. W. & Helder, A. (2017). Cognitive Processes in Discourse Comprehension: Passive Processes, Reader-Initiated Processes, and Evolving Mental Representations. *Discourse Processes*, 54(5–6), 360–372. <https://doi.org/10.1080/0163853X.2017.1306677>
- van den Broek, P. W. & Kendeou, P. (2022). Reading Comprehension I: Discourse. In M. J. Snowling, C. Hulme & K. Nation (Hrsg.), *The Science of Reading* (2. Aufl., S. 239–260). Wiley. <https://doi.org/10.1002/9781119705116.ch11>
- van den Broek, P. W., Lorch, R. F., Linderholm, T. & Gustafson, M. (2001). The Effects of Readers' Goals on Inference Generation and Memory for Texts. *Memory & Cognition*, 29(8), 1081–1087. <https://doi.org/10.3758/BF03206376>
- van den Broek, P. W., Risden, K. & Husebye-Hartmann, E. (1995). The Role of Readers' Standards for Coherence in the Generation of Inferences during Reading. In R. F. Lorch & E. J. O'Brien (Hrsg.), *Sources of Coherence in Reading* (S. 353–373). Laurence Erlbaum Associates.
- van den Broek, P. W., Tzeng, Y., Risden, K., Trabasso, T. & Basche, P. (2001). Inferential Questioning: Effects on Comprehension of Narrative Texts as a Function of Grade and Timing. *Journal of Educational Psychology*, 93(3), 521–529. <https://doi.org/10.1037/0022-0663.93.3.521>
- van den Broek, P. W., Virtue, S., Everson, M. G., Tzeng, Y. & Sung, Y. (2002). Comprehension and Memory of Science Texts: Inferential Processes and the Construction of a Mental Representation. In J. C. Otero, J. A. León & A. C. Graesser (Hrsg.), *The Psychology of Science Text Comprehension* (S. 131–154). Routledge.
- van der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Gašević, D., Bannert, M. & Molenaar, I. (2022). The Dynamics between Self-Regulated Learning and Learning Outcomes: An Exploratory Approach and Implications. *Metacognition and Learning*, 17(3), 745–771. <https://doi.org/10.1007/s11409-022-09308-9>
- van der Heide, B. & Lim, Y. (2016). On the Conditional Cueing of Credibility Heuristics: The Case of Online Influence. *Communication Research*, 43(5), 672–693. <https://doi.org/10.1177/0093650214565915>
- van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R. & Valkenburg, P. M. (2015). The Consequences of Media Multitasking for Youth: A Review. *Computers in Human Behavior*, 53, 204–215. <https://doi.org/10.1016/j.chb.2015.06.035>

- van Dijk, T. A. (1979). Relevance Assignment in Discourse Comprehension. *Discourse Processes*, 2(2), 113–126. <https://doi.org/10.1080/01638537909544458>
- van Dijk, T. A. (1980). *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition*. Erlbaum.
- van Gog, T. (2022). The Signaling (or Cueing) Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 221–230). Cambridge University Press. <https://doi.org/10.1017/9781108894333.022>
- van Kraayenoord, C. E. (2010). The Role of Metacognition in Reading Comprehension. In H.-P. Trolldenier, W. Lenhard & P. Marx (Hrsg.), *Brennpunkte der Gedächtnisforschung: Entwicklungs- und pädagogisch-psychologische Perspektiven* (S. 277–302). Hogrefe.
- van Meter, P. N. & Stepanik, N. (2020). Interventions to Support Learning from Multiple External Representations. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Hrsg.), *Handbook of Learning from Multiple Representations and Perspectives* (S. 76–91). Routledge.
- van Ockenburg, L., van Weijen, D. & Rijlaarsdam, G. (2019). Learning to Write Synthesis Texts: A Review of Intervention Studies. *Journal of Writing Research*, 10(3), 401–428. <https://doi.org/10.17239/jowr-2019.10.03.01>
- van Steendam, E., Vandermeulen, N., Maeyer, S. de, Lesterhuis, M., van den Bergh, H. & Rijlaarsdam, G. (2022). How Students Perform Synthesis Tasks: An Empirical Study into Dynamic Process Configurations. *Journal of Educational Psychology*, 114(8), 1773–1800. <https://doi.org/10.1037/edu0000755>
- van Strien, J. L. H., Brand-Gruwel, S. & Boshuizen, H. P. A. (2014). Dealing with Conflicting Information from Multiple Nonlinear Texts: Effects of Prior Attitudes. *Computers in Human Behavior*, 32, 101–111. <https://doi.org/10.1016/j.chb.2013.11.021>
- van Strien, J. L. H., Kammerer, Y., Brand-Gruwel, S. & Boshuizen, H. P. A. (2016). How Attitude Strength Biases Information Processing and Evaluation on the Web. *Computers in Human Behavior*, 60, 245–252. <https://doi.org/10.1016/j.chb.2016.02.057>
- Vandermeulen, N., Maeyer, S., Steendam, E., Lesterhuis, M., van den Bergh, H. & Rijlaarsdam, G. (2020). Mapping Synthesis Writing in Various Levels of Dutch Upper-Secondary Education: A National Baseline Study on Text Quality, Writing Process and Students' Perspectives on Writing. *Pedagogische Studien*, 97(3), 187–236.
- Vandermeulen, N., van den Broek, B., van Steendam, E. & Rijlaarsdam, G. (2020). In Search of an Effective Source Use Pattern for Writing Argumentative and Informative Synthesis Texts. *Reading and Writing*, 33(2), 239–266. <https://doi.org/10.1007/s11145-019-09958-3>
- Vandermeulen, N., van Steendam, E. & Rijlaarsdam, G. (2023). Introduction to the Special Issue on Synthesis Tasks: Where Reading and Writing Meet. *Reading and Writing*, 36(4), 747–768. <https://doi.org/10.1007/s11145-022-10394-z>
- Vandevelde, S., van Keer, H., Schellings, G. & van Hout-Wolters, B. H. A. M. (2015). Using Think-Aloud Protocol Analysis to Gain In-Depth Insights into Upper Primary School Children's Self-Regulated Learning. *Learning and Individual Differences*, 43, 11–30. <https://doi.org/10.1016/j.lindif.2015.08.027>
- Vasilev, M. R., Kirkby, J. A. & Angele, B. (2018). Auditory Distraction during Reading: A Bayesian Meta-Analysis of a Continuing Controversy. *Perspectives on Psychological Science*, 13(5), 567–597. <https://doi.org/10.1177/1745691617747398>
- Veenman, M. V. J. (2011). Alternative Assessment of Strategy Use with Self-Report Instruments: A Discussion. *Metacognition and Learning*, 6(2), 205–211. <https://doi.org/10.1007/s11409-011-9080-x>
- Veenman, M. V. J. & Beishuizen, J. J. (2004). Intellectual and Metacognitive Skills of Novices while Studying Texts under Conditions of Text Difficulty and Time Constraint. *Learning and Instruction*, 14(6), 621–640. <https://doi.org/10.1016/j.learninstruc.2004.09.004>
- Veenman, M. V. J., van Hout-Wolters, B. H. A. M. & Afflerbach, P. (2006). Metacognition and Learning: Conceptual and Methodological Considerations. *Metacognition and Learning*, 1(1), 3–14. <https://doi.org/10.1007/s11409-006-6893-0>
- Vidal-Abarca, E., Mañá, A. & Gil, L. (2010). Individual Differences for Self-Regulating Task-Oriented Reading Activities. *Journal of Educational Psychology*, 102(4), 817–826.

- Vidal-Abarca, E., Salmerón, L. & Mañá, A. (2011). Individual Differences in Task-Oriented Reading. In M. T. McCrudden, J. P. Magliano & G. Schraw (Hrsg.), *Text Relevance and Learning from Text* (S. 267–293). Information Age.
- von der Mühlen, S., Richter, T., Schmid, S., Schmidt, E. M. & Berthold, K. (2015). Judging the Plausibility of Arguments in Scientific Texts: A Student – Scientist Comparison. *Thinking & Reasoning*, 22(2), 221–249. <https://doi.org/10.1080/13546783.2015.1127289>
- von der Mühlen, S., Richter, T., Schmid, S., Schmidt, E. M. & Berthold, K. (2016). The Use of Source-Related Strategies in Evaluating Multiple Psychology Texts: A Student – Scientist Comparison. *Reading and Writing*, 29(8), 1677–1698. <https://doi.org/10.1007/s11145-015-9601-0>
- Vongkulluksn, V. W., Nishiyama, C., Rodriguez, M. C. & Nussbaum, E. M. (2023). Critical Reading of Informational Texts (CRIT) Scaffold: Evaluating the Efficacy of an Instructional Scaffold for Reading Multiple Scientific Texts. *Contemporary Educational Psychology*, 75, Artikel 102229, 1–12. <https://doi.org/10.1016/j.cedpsych.2023.102229>
- Voss, J. F. & Wiley, J. (1997). Developing Understanding while Writing Essays in History. *International Journal of Educational Research*, 27(3), 255–265. [https://doi.org/10.1016/S0883-0355\(97\)89733-9](https://doi.org/10.1016/S0883-0355(97)89733-9)
- Vu, N. N., Hung, B. P., Van, N. T. T. & Lien, N. T. H. (2022). Theoretical and Instructional Aspects of Using Multimedia Resources in Language Education: A Cognitive View. In R. Kumar, R. Sharma & P. K. Pattnaik (Hrsg.), *Multimedia Technologies in the Internet of Things Environment*, Volume 2 (S. 165–194). Springer Singapore. https://doi.org/10.1007/978-981-16-3828-2_9
- Wade, S. E., Trathen, W. & Schraw, G. (1990). An Analysis of Spontaneous Study Strategies. *Reading Research Quarterly*, 25(2), 147–166. <https://doi.org/10.2307/747599>
- Walraven, A., Brand-Gruwel, S. & Boshuizen, H. P. A. (2008). Information-Problem Solving: A Review of Problems Students Encounter and Instructional Solutions. *Computers in Human Behavior*, 24(3), 623–648. <https://doi.org/10.1016/j.chb.2007.01.030>
- Walraven, A., Brand-Gruwel, S. & Boshuizen, H. P. A. (2009). How Students Evaluate Information and Sources when Searching the World Wide Web for Information. *Computers & Education*, 52(1), 234–246. <https://doi.org/10.1016/j.compedu.2008.08.003>
- Walsh-Moorman, B. & Pytash, K. E. (2022). Guiding their Thinking: A Formative Study of Digital Source Evaluation. *Journal of Adolescent & Adult Literacy*, 65(6), 469–479. <https://doi.org/10.1002/jaal.1225>
- Wang, Y. & List, A. (2019). Calibration in Multiple Text Use. *Metacognition and Learning*, 14(2), 131–166. <https://doi.org/10.1007/s11409-019-09201-y>
- Wei, Z., Zhang, Y., Clariana, R. B. & Chen, X. (2024). The Effects of Reading Prompts and of Post-Reading Generative Learning Tasks on Multiple Document Integration: Evidence from Concept Network Analysis. *Educational Technology Research and Development*, 72(2), 661–685. <https://doi.org/10.1007/s11423-023-10326-w>
- Weinstein, Y. (2018). Mind-Wandering, How Do I Measure Thee with Probes? Let Me Count the Ways. *Behavior Research Methods*, 50(2), 642–661. <https://doi.org/10.3758/s13428-017-0891-9>
- Wertgen, A. G. & Richter, T. (2020). Source Credibility Modulates the Validation of Implausible Information. *Memory & Cognition*, 48(8), 1359–1375. <https://doi.org/10.3758/s13421-020-01067-9>
- Wertgen, A. G. & Richter, T. (2023). Source Credibility and Plausibility Are Considered in the Validation of Textual Information: Evidence from a Social Media Context. *Journal of Cognitive Psychology*, 35(2), 183–204. <https://doi.org/10.1080/20445911.2022.2149757>
- Westerwick, A. (2013). Effects of Sponsorship, Web Site Design, and Google Ranking on the Credibility of Online Information. *Journal of Computer-Mediated Communication*, 18(2), 194–211. <https://doi.org/10.1111/jcc4.12006>
- Weston-Sementelli, J. L., Allen, L. K. & McNamara, D. S. (2018). Comprehension and Writing Strategy Training Improves Performance on Content-Specific Source-Based Writing Tasks. *International Journal of Artificial Intelligence in Education*, 28(1), 106–137. <https://doi.org/10.1007/s40593-016-0127-7>
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K. & Hemmerich, J. A. (2009). Source Evaluation, Comprehension, and Learning in Internet Science Inquiry Tasks. *American Educational Research Journal*, 46(4), 1060–1106. <https://doi.org/10.3102/0002831209333183>

- Wiley, J., Griffin, T. D., Steffens, B. & Britt, M. A. (2020). Epistemic Beliefs about the Value of Integrating Information across Multiple Documents in History. *Learning and Instruction*, 65(Article 101266), 1–16. <https://doi.org/10.1016/j.learninstruc.2019.101266>
- Wiley, J. & Voss, J. F. (1996). The Effects of 'Playing Historian' on Learning in History. *Applied Cognitive Psychology*, 10(7), 63–72. [https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1099-0720\(199611\)10:7%3C63::AID-ACP438%3E3.0.CO;2-5](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1099-0720(199611)10:7%3C63::AID-ACP438%3E3.0.CO;2-5)
- Wiley, J. & Voss, J. F. (1999). Constructing Arguments from Multiple Sources: Tasks That Promote Understanding and Not Just Memory for Text. *Journal of Educational Psychology*, 91(2), 301–311. <https://doi.org/10.1037/0022-0663.91.2.301>
- Willhite, S. C. (1985). Differential Effects of High-Level and Low-Level Postpassage Questions. *The American Journal of Psychology*, 98(1), 41–58. <https://doi.org/10.2307/1422766>
- Willoughby, T., Wood, E. & Khan, M. (1994). Isolating Variables That Impact on or Detract from the Effectiveness of Elaboration Strategies. *Journal of Educational Psychology*, 86(2), 279–289. <https://doi.org/10.1037/0022-0663.86.2.279>
- Wilson, E. J. & Sherrell, D. L. (1993). Source Effects in Communication and Persuasion Research: A Meta-Analysis of Effect Size. *Journal of the Academy of Marketing Science*, 21(2), 101–112. <https://doi.org/10.1007/BF02894421>
- Wineburg, S. S. (1991). Historical Problem Solving: A Study of the Cognitive Processes Used in the Evaluation of Documentary and Pictorial Evidence. *Journal of Educational Psychology*, 83(1), 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>
- Wineburg, S. S., Breakstone, J., McGrew, S., Smith, M. D. & Ortega, T. (2022). Lateral Reading on the Open Internet: A District-Wide Field Study in High School Government Classes. *Journal of Educational Psychology*, 114(5), 893–909. <https://doi.org/10.1037/edu0000740>
- Wineburg, S. S. & McGrew, S. (2019). Lateral Reading and the Nature of Expertise: Reading Less and Learning More when Evaluating Digital Information. *Teachers College Record*, 121(11), Artikel 110302, 1–40. <https://doi.org/10.1177/016146811912101102>
- Winne, P. H. (2001). Self-Regulated Learning Viewed from Models of Information Processing. In B. J. Zimmerman & D. H. Schunk (Hrsg.), *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives* (2. Aufl., S. 153–189). Routledge.
- Winne, P. H. (2010). Improving Measurements of Self-Regulated Learning. *Educational Psychologist*, 45(4), 267–276. <https://doi.org/10.1080/00461520.2010.517150>
- Winne, P. H. (2015). Self-Regulated Learning. In J. D. Wright (Hrsg.), *International Encyclopedia of the Social & Behavioral Sciences* (2. Aufl., S. 535–540). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.25091-5>
- Winne, P. H. (2018a). Cognition and Metacognition within Self-Regulated Learning. In D. H. Schunk & J. A. Greene (Hrsg.), *Handbook of Self-Regulation of Learning and Performance* (2. Aufl., S. 36–48). Routledge.
- Winne, P. H. (2018b). Theorizing and Researching Levels of Processing in Self-Regulated Learning. *British Journal of Educational Psychology*, 88(1), 9–20. <https://doi.org/10.1111/bjep.12173>
- Winne, P. H. & Hadwin, A. F. (1998). Studying as Self-Regulated Learning. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in Educational Theory and Practice* (S. 277–304). Erlbaum.
- Winne, P. H. & Hadwin, A. F. (2008). The Weave of Motivation and Self-Regulated Learning. In D. H. Schunk & B. J. Zimmerman (Hrsg.), *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (S. 297–314). Routledge.
- Winne, P. H. & Jamieson-Noel, D. (2002). Exploring Students' Calibration of Self Reports about Study Tactics and Achievement. *Contemporary Educational Psychology*, 27(4), 551–572. [https://doi.org/10.1016/S0361-476X\(02\)00006-1](https://doi.org/10.1016/S0361-476X(02)00006-1)
- Winne, P. H. & Marzouk, Z. (2019). Learning Strategies and Self-Regulated Learning. In J. Dunlosky & K. A. Rawson (Hrsg.), *The Cambridge Handbook of Cognition and Education* (S. 696–715). Cambridge University Press. <https://doi.org/10.1017/9781108235631.028>
- Winter, S. & Krämer, N. C. (2012). Selecting Science Information in Web 2.0: How Source Cues, Message Sidedness, and Need for Cognition Influence Users' Exposure to Blog Posts. *Journal of Computer-Mediated Communication*, 18(1), 80–96. <https://doi.org/10.1111/j.1083-6101.2012.01596.x>

- Winters, F. I. & Azevedo, R. (2005). High-School Students' Regulation of Learning during Computer-Based Science Inquiry. *Journal of Educational Computing Research*, 33(2), 189–217. <https://doi.org/10.2190/F7HM-9JN5-JUX8-4BM9>
- Wiradhany, W. & Koerts, J. (2021). Everyday Functioning-Related Cognitive Correlates of Media Multitasking: A Mini Meta-Analysis. *Media Psychology*, 24(2), 276–303. <https://doi.org/10.1080/015213269.2019.1685393>
- Wiradhany, W. & Nieuwenstein, M. R. (2017). Cognitive Control in Media Multitaskers: Two Replication Studies and a Meta-Analysis. *Attention, perception & psychophysics*, 79(8), 2620–2641. <https://doi.org/10.3758/s13414-017-1408-4>
- Wirth, J., Stebner, F., Trypke, M., Schuster, C. & Leutner, D. (2020). An Interactive Layers Model of Self-Regulated Learning and Cognitive Load. *Educational Psychology Review*, 32(4), 1127–1149. <https://doi.org/10.1007/s10648-020-09568-4>
- Witherspoon, A. M., Azevedo, R. & D'Mello, S. (2008). The Dynamics of Self-Regulatory Processes within Self- and Externally Regulated Learning Episodes During Complex Science Learning with Hypermedia. In B. P. Woolf, E. Aïmeur, R. Nkambou & S. Lajoie (Hrsg.), *Intelligent Tutoring Systems: 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008 Proceedings* (S. 260–269). Springer Berlin Heidelberg.
- Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, 24(4), 345–376. https://doi.org/10.1207/s15326985ep2404_2
- Wolfe, M. B. W. & Goldman, S. R. (2005). Relations between Adolescents' Text Processing and Reasoning. *Cognition and Instruction*, 23(4), 467–502. https://doi.org/10.1207/s1532690xc2304_2
- Wolfe, M. B. W. & Griffin, T. D. (2017). Beliefs and Discourse Processing. In M. F. Schober, D. N. Rapp & M. A. Britt (Hrsg.), *The Routledge Handbook of Discourse Processes* (2. Aufl., S. 295–314). Routledge.
- Woloshyn, V. E., Willoughby, T., Wood, E. & Pressley, M. (1990). Elaborative Interrogation Facilitates Adult Learning of Factual Paragraphs. *Journal of Educational Psychology*, 82(3), 513–524. <https://doi.org/10.1037/0022-0663.82.3.513>
- Wong, A. Y., Smith, S. L., McGrath, C. A., Flynn, L. E. & Mills, C. (2022). Task-Unrelated Thought during Educational Activities: A Meta-Analysis of Its Occurrence and Relationship with Learning. *Contemporary Educational Psychology*, 71, Artikel 102098, 1–18. <https://doi.org/10.1016/j.cedpsych.2022.102098>
- Wong, B. Y. L., Wong, R. & LeMare, L. (1982). The Effects of Knowledge of Criterion Task on Comprehension and Recall in Normally Achieving and Learning Disabled Children. *The Journal of Educational Research*, 76(2), 119–126. <https://doi.org/10.1080/00220671.1982.10885434>
- Wylie, J., Thomson, J., Leppänen, P. H. T., Ackerman, R., Kanninen, L. & Prieler, T. (2018). Cognitive Processes and Digital Reading. In M. Barzillai, J. Thomson, S. Schroeder & P. W. van den Broek (Hrsg.), *Learning to Read in a Digital World* (S. 57–90). John Benjamins. <https://doi.org/10.1075/swll.17.03wyl>
- Xie, H., Wang, F., Hao, Y., Chen, J., An, J., Wang, Y. & Liu, H. (2017). The More Total Cognitive Load Is Reduced by Cues, the Better Retention and Transfer of Multimedia Learning: A Meta-Analysis and Two Meta-Regression Analyses. *PLoS ONE*, 12(8), Artikel e0183884, 1–20. <https://doi.org/10.1371/journal.pone.0183884>
- Yang, C., Zhao, W., Yuan, B., Luo, L. & Shanks, D. R. (2023). Mind the Gap between Comprehension and Metacomprehension: Meta-Analysis of Metacomprehension Accuracy and Intervention Effectiveness. *Review of Educational Research*, 93(2), 143–194. <https://doi.org/10.3102/00346543221094083>
- Yang, H.-C. (2014). Toward a Model of Strategies and Summary Writing Performance. *Language Assessment Quarterly*, 11(4), 403–431. <https://doi.org/10.1080/15434303.2014.957381>
- Yang, H.-C. & Plakans, L. (2012). Second Language Writers' Strategy Use and Performance on an Integrated Reading-Listening-Writing Task. *TESOL Quarterly*, 46(1), 80–103. <https://doi.org/10.1002/tesq.6>
- Yapp, D. J., Graaff, R. de & van den Bergh, H. (2021). Improving Second Language Reading Comprehension through Reading Strategies: A Meta-Analysis of L2 Reading Strategy Interventions. *Journal of Second Language Studies*, 4(1), 154–192. <https://doi.org/10.1075/jsls.19013.yap>

- Yeager, D. S. & Dweck, C. S. (2020). What Can Be Learned from Growth Mindset Controversies? *American Psychologist*, 75(9), 1269–1284. <https://doi.org/10.1037/amp0000794>
- Yeari, M., van den Broek, P. W. & Oudega, M. (2015). Processing and Memory of Central versus Peripheral Information as a Function of Reading Goals: Evidence from Eye-Movements. *Reading and Writing*, 28(8), 1071–1097. <https://doi.org/10.1007/s11145-015-9561-4>
- Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H. & Pieper, S. (2013). Shifting Ability Predicts Math and Reading Performance in Children: A Meta-Analytical Study. *Learning and Individual Differences*, 23(1), 1–9. <https://doi.org/10.1016/j.lindif.2012.10.004>
- Yukhymenko-Lescroart, M. A., Goldman, S. R., Lawless, K. A., Pellegrino, J. W. & Shanahan, C. R. (2022). Assessing Information Synthesis within and across Multiple Texts with Verification Tasks: A Signal Detection Theory Approach. *Educational Psychology*, 42(5), 1–18. <https://doi.org/10.1080/01443410.2020.1811840>
- ZanESCO, A. P., Denkova, E. & Jha, A. P. (in Druck). Mind-Wandering Increases in Frequency over Time during Task Performance: An Individual-Participant Meta-Analytic Review. *Psychological Bulletin*. <https://doi.org/10.1037/bul0000424>
- Zengilowski, A., Schuetze, B. A., Nash, B. L. & Schallert, D. L. (2021). A Critical Review of the Refutation Text Literature: Methodological Confounds, Theoretical Problems, and Possible Solutions. *Educational Psychologist*, 56(3), 175–195. <https://doi.org/10.1080/00461520.2020.1861948>
- Zhang, D., Hyönä, J., Cui, L., Zhu, Z. & Li, S. (2019). Effects of Task Instructions and Topic Signaling on Text Processing among Adult Readers with Different Reading Styles: An Eye-Tracking Study. *Learning and Instruction*, 64, Artikel 101246, 1–15. <https://doi.org/10.1016/j.learninstruc.2019.101246>
- Zhang, S., Koning, B. de, Agostinho, S., Tindall-Ford, S., Chandler, P. & Paas, F. (2022). The Cognitive Load Self-Management Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (3. Aufl., S. 430–436). Cambridge University Press. <https://doi.org/10.1017/9781108894333.044>
- Zhao, P., Zhu, X., Yao, Y. & Liao, X. (2023). Ideal L2 Self, Enjoyment, and Strategy Use in L2 Integrated Writing: A Self-Regulatory Learning Perspective. *System*, 115, Artikel 103033, 1–10. <https://doi.org/10.1016/j.system.2023.103033>
- Zheng, L. (2016). The Effectiveness of Self-Regulated Learning Scaffolds on Academic Performance in Computer-Based Learning Environments: A Meta-Analysis. *Asia Pacific Education Review*, 17(2), 187–202. <https://doi.org/10.1007/s12564-016-9426-9>
- Zhu, X., Li, G. Y., Cheong, C. M. & Wen, H. (2021). Effects of L1 Single-Text and Multiple-Text Comprehension on L2 Integrated Writing. *Assessing Writing*, 49, Artikel 100546, 1–13. <https://doi.org/10.1016/j.asw.2021.100546>
- Zimmerman, B. J. (2000). Attaining Self-Regulation. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Handbook of Self-Regulation* (S. 13–39). Academic Press.
- Zimmerman, B. J. (2013). From Cognitive Modeling to Self-Regulation: A Social Cognitive Career Path. *Educational Psychologist*, 48(3), 135–147. <https://doi.org/10.1080/00461520.2013.794676>
- Zimmerman, B. J. & Moylan, A. R. (2009). Self-Regulation: Where Metacognition and Motivation Intersect. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Handbook of Metacognition in Education* (S. 299–315). Routledge.
- Zimmerman, B. J. & Risemberg, R. (1997). Becoming a Self-Regulated Writer: A Social Cognitive Perspective. *Contemporary Educational Psychology*, 22(1), 73–101. <https://doi.org/10.1006/ceps.1997.0919>