

Trauten, Florian

Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie

Berlin : Logos Verlag 2024, 141 S. - (Studien zum Physik- und Chemielernen; 380) - (Dissertation, Universität Duisburg-Essen, 2024)



Quellenangabe/ Reference:

Trauten, Florian: Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie. Berlin : Logos Verlag 2024, 141 S. - (Studien zum Physik- und Chemielernen; 380) - (Dissertation, Universität Duisburg-Essen, 2024) - URN: urn:nbn:de:0111-pedocs-334760 - DOI: 10.25656/01:33476; 10.30819/5859

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-334760>

<https://doi.org/10.25656/01:33476>

in Kooperation mit / in cooperation with:



<https://www.logos-verlag.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc-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 make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

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

Studien zum Physik- und Chemielernen

M. Hopf und M. Ropohl [Hrsg.]

380

Florian Trauten

Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie



λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Diese Reihe im Logos Verlag Berlin lädt Forscherinnen und Forscher ein, ihre neuen wissenschaftlichen Studien zum Physik- und Chemielernen im Kontext einer Vielzahl von bereits erschienenen Arbeiten zu quantitativen und qualitativen empirischen Untersuchungen sowie evaluativ begleiteten Konzeptionsentwicklungen zu veröffentlichen. Die in den bisherigen Studien erfassten Themen und Inhalte spiegeln das breite Spektrum der Einflussfaktoren wider, die in den Lehr- und Lernprozessen in Schule und Hochschule wirksam sind.

Die Herausgeber hoffen, mit der Förderung von Publikationen, die sich mit dem Physik- und Chemielernen befassen, einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Verbesserung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Martin Hopf und Mathias Ropohl

Studien zum Physik- und Chemielernen

Band 380

Florian Trauten

**Entwicklung und Evaluation
von automatisierten Feedbackschleifen
in Online-Aufgaben im Fach Chemie**

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Martin Hopf und Mathias Ropohl [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



Dieses Werk ist lizenziert unter der Creative Commons Attribution 4.0 Lizenz CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z. B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

Logos Verlag Berlin GmbH 2024

ISBN 978-3-8325-5859-8

ISSN 1614-8967

DOI 10.30819/5859

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<https://www.logos-verlag.de>

Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie

Dissertation

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

– Dr. rer. nat. –

vorgelegt von

Florian Trauten

geboren in Essen-Kupferdreh

Fakultät für Chemie

der

Universität Duisburg-Essen

2023

Gutachter: Prof. Dr. Maik Walpuski

Prof. Dr. Elke Sumfleth

Vorsitzender: Prof. Dr. Mathias Eppe

Tag der Disputation: 11. März 2024

Für alle, die an mich geglaubt und mich so weit gebracht haben.

*„Wer noch nie einen Fehler gemacht hat,
hat sich noch nie an etwas Neuem versucht.“*

Albert Einstein

Inhaltsverzeichnis

Einleitung	3
Kapitel 1: Studienabbruchforschung und Studieneingangsphase	5
1.1. Faktoren von Studienabbruch	7
1.2 Hürden und Herausforderungen in der Studieneingangsphase	10
1.3 Fachspezifische Forschungsergebnisse	12
Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen	17
2.1 Klassische Lerntheorien, ein Überblick	17
2.2 Der Begriff Feedback, eine definitorische Klärung	21
2.3 Feedback in Lernsituationen	24
2.4 Zum Forschungsstand von Feedback in Lernaufgaben	36
Kapitel 3: Gestaltungsprinzipien Multimedialen Lernens	43
3.1 Cognitive Load Theory	43
3.2 Cognitive Theory of Multimedia Learning	47
3.2 Modell zur Entwicklung digitaler Lernumgebungen	53
Kapitel 4: Forschungsfragen, Ziele und Hypothesen	55
Kapitel 5: Studiendesign und Testinstrumente	59
Kapitel 6: Aufgabenentwicklung	65
6.1 Lernzielorientierte Aufgabenkonstruktion	65
6.2 Das Software-Tool	69
6.3 Der Feedback Algorithmus	72
6.3 Implementierung der JACK® Aufgaben und Gestaltung des Moodle-Kurses	77
Kapitel 7: Stichprobenbeschreibung	79
Kapitel 8: Ergebnisse der durchgeführten Studien	81
8.1 Pilotierung	82
8.1.1 Ergebnisse	82
8.1.2 Implikationen für die Hauptstudie	84
8.2 Hauptstudie	87
8.2.1 Arbeitsverhalten in der Lernumgebung	87

8.2.2 Einfluss des Feedback-Typs auf das Fachwissen (Post).....	91
8.2.3 Lernzuwachs in Abhängigkeit von der Kurswahl.....	95
8.2.4 Einfluss des Feedback-Typs auf das Klausurergebnis	97
8.2.5 Kohortenvergleich mit der Vergleichskohorte aus ALSTER I.....	100
Kapitel 9: Zusammenfassung und Diskussion.....	105
9.1 Zusammenfassung der Ergebnisse	105
9.2 Theoretischer und praktischer Ertrag.....	108
Kapitel 10: Limitationen und Ausblick.....	111
Abbildungsverzeichnis.....	117
Tabellenverzeichnis.....	119
Literaturverzeichnis	121
Anhang	133
1. Skalendokumentation der eingesetzten Tests	133
1.1. Mathematisches Wissen	133
1.2. Fachwissen Chemie.....	135
1.3. Kognitive Fähigkeiten.....	137
2. Bonuspunktregelung	139

Einleitung

Mit der durch den Bologna Prozess initiierten „widening participation“ Agenda (Weedon & Riddell, 2016) hat man in Deutschland u. a. durch Reduzierung der Eingangsvoraussetzung begonnen, (MINT-)Studienfächer einer breiteren Anzahl an Studierenden zu öffnen. Damit ist sehr früh der Grundstein gelegt worden, „der sich verändernden Tätigkeitsnachfrage und der zunehmenden Wissensbasierung der deutschen Produktion“ (Brossardt, 2015, S. 67) entgegensteuern zu können, die neben dem demografischen Wandel mit verantwortlich für den drohenden Fachkräftemangel gemacht werden. Bereits 2015 wurde für den naturwissenschaftlichen Sektor der Arbeitslandschaft 2040 prognostiziert, dass „im Schnitt zwölf Prozent aller Stellen für Hochqualifizierte unbesetzt“ bleiben (Brossardt, 2015, S. 62). Im Feld der Forschung (fachrichtungsunabhängig) wurde sogar ein Arbeitskräftesaldo für Personen mit Hochschulabschluss von rund -18 % prognostiziert.

Geringere Eingangsvoraussetzungen führten allerdings fachübergreifend auch zu einer Zunahme von Studierenden mit geringerem (fachspezifischen) Vorwissen (Organisation for Economic Co-operation and Development, 2008), so auch in MINT-Fächern. Da fachliche Defizite den Hauptgrund für Studienabbruch darstellen und die Entwicklung der Abbruchquote im Fach Chemie mit 23 % im Jahre 1999 (Heublein, Schmelzer & Sommer, 2005) sich zu mittlerweile 47 % im Jahre 2016 drastisch erhöht hat (Heublein, Richter & Schmelzer, 2020), deutet das aus Sicht der oben genannten Aspekte auf einen massiven Handlungsbedarf hin. Vergleichbare Befunde für den MINT-Bereich liegen auch für den angloamerikanischen Raum vor (X. Chen, 2015).

Die Ergebnisse von Averbek (2020) aus der ersten Förderphase des Projektes „Akademisches Lernen und Studienerfolg in der Eingangsphase von naturwissenschaftlich-technischen Studiengängen“ (ALSTER; FOR 2242) belegen zudem, dass das Fachwissen im Fach Chemie auch in späteren Semestern insbesondere von den in der Schule erworbenen Chemiekenntnissen abhängt. Lücken im Vorwissen werden zudem im ersten Semester in der Regel nicht durch die bestehenden Hochschullernangebote kompensiert, was sich für Betroffene stark auf die Abbruchsintention auswirkt (Heublein et al., 2017). Hinzu kommt, dass sich die Studierenden zu Beginn des Studiums auch in sozialer Hinsicht in einer Umbruchsphase befinden. Neben der Überwindung von Leistungsdefiziten müssen die Studierenden sich in neue Sozialgefüge einordnen (Tinto, 1988), sodass das Verhältnis zu den Lehrpersonen/Mitstudierenden gerade in der frühen Studieneingangsphase eher distanziert ist. Feedback

muss nun aktiv eingefordert werden (Heublein et al., 2017; Rost, 2018). Dass Feedback an Hochschulen besonders selten von Studierenden aus der Gruppe der Studienabbrechenden in Anspruch genommen wird (Heublein et al., 2017), lässt in Verbindung mit der Prämisse, dass Feedback eine der wichtigsten Einflussgrößen auf Lernprozesse darstellt (Hattie & Timperley, 2007), den Einfluss dieser Feedback-Problematik auf den Performance-Sektor fast als logische Folge erscheinen. Die Schlüsselrolle, die der Disziplin der Allgemeinen Chemie für das gesamte Chemiestudium zukommt, bietet daher berechnete Gründe, genau dort eine Intervention anzusetzen, die den Studierenden aktiv Feedback bereitstellt. Über z. B. informatives tutorielles Feedback (ITF) in einer Online-Lernumgebung (Narciss, 2006), welches mit typisch auftauchenden Fehlern bei der Aufgabenbearbeitung umzugehen weiß, könnten Leistungsdefizite spezifisch adressiert werden. So kann dann in einer durch soziale Eingewöhnung geprägten Phase eine Unterstützung bereitgestellt werden, um dem Mangel an Feedback entgegenzuwirken.

Doch selbst wenn Feedback erhalten wird, ist das noch kein Gelingensgarant, denn Feedback ist nur dann wirksam, wenn es das Vorwissen der lernenden Person angemessen berücksichtigt. Grundsätzlich ist jedoch von einer heterogenen Verteilung des Vorwissens bei den Erstsemesterstudierenden auszugehen, allein schon auf Basis der Grundlage, ob ein Leistungskurs in der Oberstufe belegt wurde oder nicht (Averbeck, 2020). Eine Interventionsmaßnahme, die das Vorwissen aller Studierenden konzeptionell berücksichtigen soll, ist daher mit einer Reihe an Hürden konfrontiert. Bereits aus der Kombination von *Feedback-Principle* (Johnson & Priest, 2014; Moreno & Mayer, 2007), *Worked Examples Principle* (Renkl, 2014b) und *Expertise Reversal Effect* (O. Chen, Kalyuga & Sweller, 2017) ergeben sich Probleme für die Passung bei Lernenden mit unterschiedlicher Expertise. Aufgrund großer Varianz hinsichtlich der Effektstärken von Feedback in Verbindung mit unterschiedlichen Forschungsdesigns kann die Feedbackforschung zur Zeit auch keine gesicherten Erkenntnisse hinsichtlich des Vorwissenseinflusses auf die Wirksamkeit von Feedback unterschiedlichen Umfangs liefern (Fyfe, Rittle-Johnson & DeCaro, 2012; Narciss, 2006; Smits, Boon, Sluijsmans & van Gog, 2008). Für chemienahe Kontexte konnten aber Indizien für einen solchen Zusammenhang gefunden werden (Albacete & Van-Lehn, 2000; Moreno, 2004; Narciss & Huth, 2004). Über einen systematischen Vergleich zweier Feedbackalgorithmen sollten so weitere Erkenntnisse in Bezug auf die lernförderliche Wirkung von Feedbackkomponenten gewonnen werden. Gleichzeitig sollte auf diese Weise eine effektive Fördermaßnahme der Binnendifferenzierung für Erstsemesterstudierende im Fach Chemie generiert werden.

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

Studienabbruch¹ kann als komplexer, multikausaler Prozess verstanden werden, der auf der misslingenden Passung zwischen Voraussetzungen des Individuums und dem daraus resultierenden Verhalten im Studium auf der einen sowie den institutionellen Anforderungen und Rahmenbedingungen auf der anderen Seite beruht (Neugebauer, Heublein & Daniel, 2019). Das folgende Kapitel soll dazu dienen, die verschiedenen für diese Arbeit relevanten Aspekte des Studienabbruchs und die zentrale Rolle der Studieneingangsphase darzulegen und bekannte Wirkmechanismen zu erläutern.

Vor dem Hintergrund der zunehmenden Wissensbasierung der deutschen Produktion (Brossardt, 2015) ist für „die Versorgung von Gesellschaft und Volkswirtschaft mit akademisch qualifizierten Fachkräften [...] neben der Zusammensetzung der Absolventenkohorten deren Umfang entscheidend. Von ihm ist abhängig, ob auf dem Arbeitsmarkt zukünftig in ausreichendem Maße Akademikerinnen und Akademiker zur Verfügung stehen werden“ (Wissenschaftsrat, 2015, S. 82). Die auf den Bologna Prozess folgende „widening participation“ Agenda (Weedon & Riddell, 2016) führte in Deutschland wie auch in vielen anderen Ländern dazu, dass sich die Studienfächer einer breiteren Masse an Studierenden geöffnet haben (Organisation for Economic Co-operation and Development, 2008). Wie aber bereits angeklungen, regt die Lage des Studienabbruchs insbesondere in den innovierenden MINT-Fächern gleich aus mehreren Perspektiven zum Handeln an. Denn hohe Abbruchquoten (Heublein et al., 2020) bedeuten immer, dass ein Teil der gesamtgesellschaftlichen Investitionen für das Hochschulsystem nicht zur Ausbildung akademisch qualifizierter Fachkräfte für den Arbeitsmarkt führt (Neugebauer et al., 2019), also folglich nicht effizient eingesetzt werden.

Es gibt forschungshistorisch verschiedene Herangehensweisen, wie man an das Thema Studienabbruch herantreten kann. Eine der wohl bekanntesten Theorien ist das von Tinto 1975 publizierte „Student Integration Model“. Demzufolge erhöht sich das Risiko des Studienabbruchs, umso weniger sich Studierende sozial und akademisch an der Hochschule integriert haben oder sich diese beiden Integrationsbereiche in einem Ungleichgewicht befinden. Soziale Integration bezieht sich dabei auf

¹ „Studienabbruch liegt vor, wenn Studierende kein Studium abschließen. Fach- oder Hochschulwechsel bedeuten keinen Studienabbruch, sofern das Studium noch erfolgreich abgeschlossen wird. Die Abbruchquote wird in einem Schätzverfahren ermittelt (vgl. www.dzhw.eu/pdf/pubfh/fh-201203.pdf)“ (Autorengruppe Bildungsberichterstattung und W. Bertelsmann Verlag (2018, S. 164).

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

die soziale Eingebundenheit, die in Schlusche, Schnaubert und Bodemer (2021) untersucht wird und wurde. Akademische Eingebundenheit bezieht sich auf die studienbezogenen Leistungen und Lernfortschritte.

„Thus, although academic dismissal is most closely associated with grade performance, drop-out in the form of voluntary withdrawal is not. Such withdrawal, instead, appears to relate to the lack of congruency between the individual and both the intellectual climate of the institution and the social system composed of his peers“ (Tinto, 1975, S. 116).

Einige Forschende steigen noch tiefer in das komplexe Geflecht der Wirkmechanismen der Soziologie ein (Thomas, 2002) und ziehen Bordieus Habitus² Theorie heran und erklären eine Abbruchgefährdung mit einer unzureichenden Passung zwischen dem herkunftsspezifischen Habitus der Studierenden und dem an der Hochschule geforderten akademischen Habitus. Es wird bemängelt, dass die Herangehensweise seitens der Universitäten sehr häufig dem traditionellen Ansatz verfallen sei, die Fehler grundsätzlich bei den Studierenden zu suchen, anstatt den akademischen Habitus dem der Studierenden anzunähern (Thomas, 2002). Hinzu kommt der Aspekt, dass Studierende sich in der Eingangsphase ihres Studiums grundsätzlich in einer schwierigen Umbruchsphase befinden (Trautwein & Bosse, 2017), die sich auch mit dem Habitus in Verbindung bringen lässt. Ein Habitus wird aktiv durch Praxis (Mit-Tun) aber größtenteils unbewusst und damit geradezu selbstverständlich und kaum thematisiert angeeignet (Portele, 1985). Daher könnten auch „fachspezifische Studien- und Lernstrategien, Arbeitsformen und Diskussionsstile sowie Lebensweisen als Habitus strukturierend und manifestierend gelten und zur Operationalisierung herangezogen werden“ (Huber, 1991, S. 423). Beim Eintritt in das Hochschulsystem sind in der Regel zwölf bzw. dreizehn Jahre im institutionellen Kontext der Schule verbracht worden. Während der schulische Kontext bis zum Ende durch eine gewisse Nähe zu den Lehrpersonen und durch angeleitetes Lernen geprägt ist, dominieren gerade beim Eintritt in die Hochschullandschaft Eigenverantwortlichkeit und eine gewisse Distanz zu den Lehrpersonen und meist auch den Mitstudierenden. Hier liegt zumindest eine Diskrepanz eher aber ein Umbruch zu den bisherigen gewohnten Arbeitsweisen und Umgangsformen vor. Damit ist auch der Sozialisationsaspekt mit dem Thema Studienabbruch eng verwoben.

Als Konsens aus diesen Herangehensweisen verschiedener Tiefe, die bislang immer noch nicht hinlänglich untersucht sind, kann der etwas weiter gefasste Standpunkt

² Bourdieu (1987) definiert den Habitus als die durch Sozialisation erworbenen Denk-, Wahrnehmungs- und Handlungsschemata.

von Heublein (2014) herangezogen werden. Er verortet die Bedingungen für ein (nicht) erfolgreiches Studium in einem Zusammenspiel aus individuellen Qualifikationen und dem Studieverhalten auf der einen und den Rahmenbedingungen der Institution Hochschule und des Soziotops des Individuums auf der anderen Seite, was der obigen Sichtweise sehr ähnelt. Diese Wirkzusammenhänge hat Heublein (2014) in einem Studienabbruchsmodell (Abbildung 1) zusammengefasst. In diesem Modell, wie auch in vielen anderen der oben genannten Theorien, kommt der Studieneingangsphase eine zentrale Bedeutung im Studienabbruchsprozess zu, welche im Folgenden genauer erläutert werden soll.

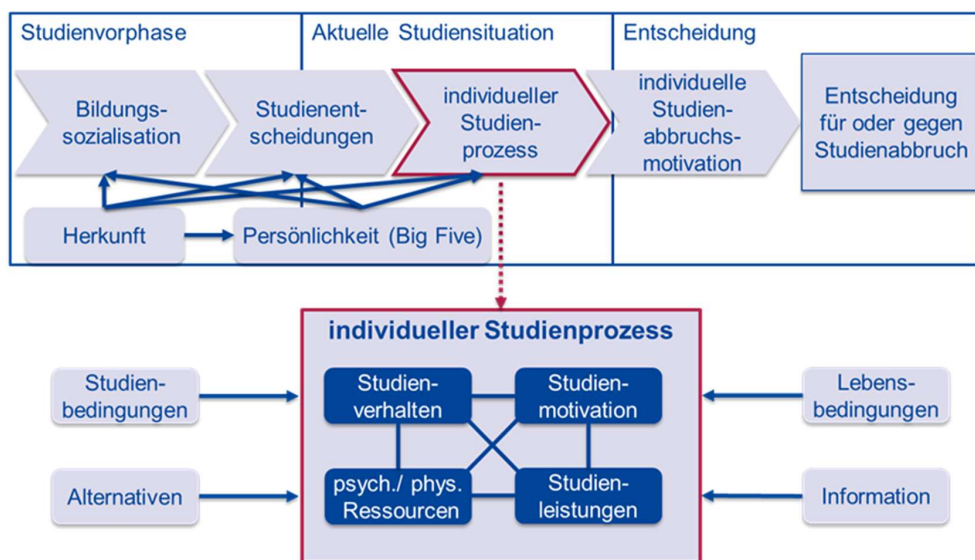


Abbildung 1: Studienabbruchsmodell (Heublein, 2014; Paczulla, Fischer, Sumfleth & Walpuski, 2019)

1.1. Faktoren von Studienabbruch

Die Umstellung der alten Studienstruktur auf das Bachelor-Master-System soll den Ausgangspunkt für die folgenden Überlegungen bilden. Denn die mit den Bologna Prozess angestoßene Harmonisierung der akademischen Ausbildung führte dazu, dass eine ganze Sparte des Bildungsapparates reformiert, reorganisiert und restrukturiert werden musste. Ob der inhaltliche Umfang auch immer an die Studienzeit angepasst wurde, bleibt fragwürdig:

So wurden die „ersten Jahre nach der Einführung der Bachelor- und Masterstudiengänge an deutschen Hochschulen ... von den Bemühungen bestimmt, die Studieninhalte in die neue

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

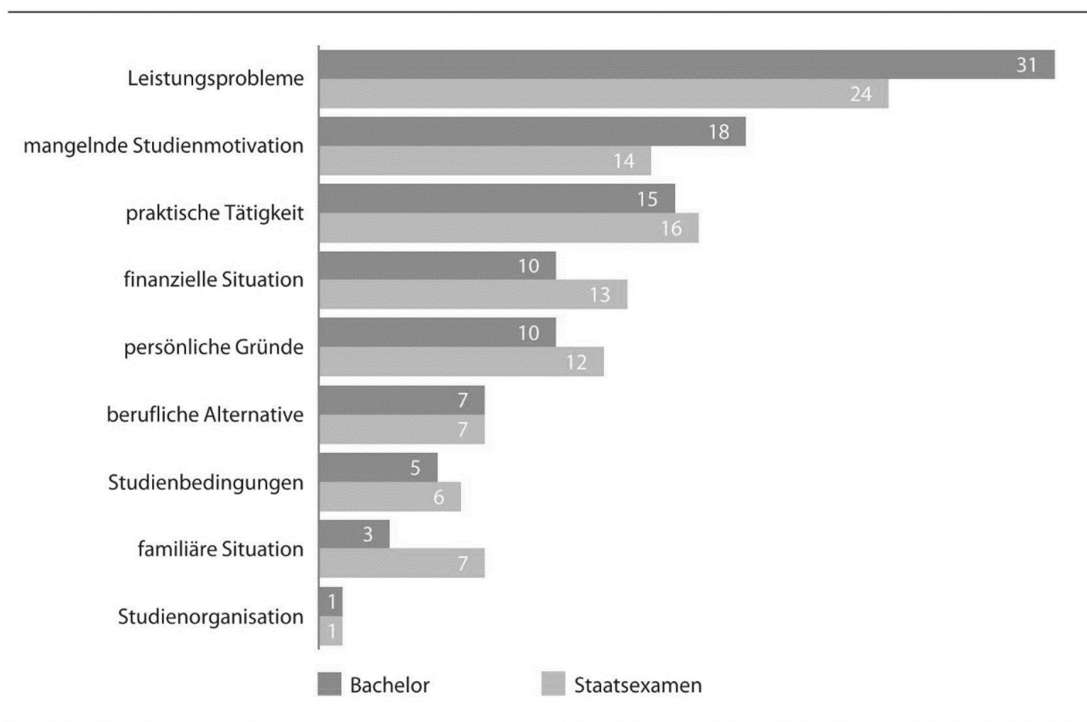
zeitliche Struktur des Studiums zu übertragen und den Anteil der Studierenden, die ihren Abschluss innerhalb der Regelstudienzeit erreichen, zu erhöhen. Die Qualifikationsziele wurden zunächst hingegen nicht immer vollständig an die neuen Stufen angepasst und die Curricula nicht flächendeckend auf eine Kompetenzorientierung umgestellt“ (Wissenschaftsrat, 2015, S. 71).

Hinzu kommt der Aspekt, dass sich im Vergleich zu früheren Diplom- und Magisterstudiengängen die Prüfungsdichte, der für die Abschlussnote wirksamen Prüfungen, in dem umgestellten System massiv erhöht hat. Manche sprechen hier auch schon von einem Prüfungsmarathon (Winter, 2009), denn, während früher zu zwei Zeitpunkten sehr umfangreiche Prüfungsphasen stattfanden (zum Ende des Grund- und des Hauptstudiums), finden sich im neuen System Prüfungen als Abschluss eines jeden Moduls und damit mehrere pro Semester, besonders in den ersten Semestern meist in Form von Klausuren. Das hat den Vorteil, dass sich die Gesamtleistung viel besser auf das Studium verteilt, doch es hat auch den Nachteil, dass vergleichsweise weniger Zeit zur Verfügung steht, um Wissenslücken aufzuarbeiten. Natürlich gab es in den alten Studiengängen ebenfalls Klausuren. Einige befürchteten aber, „die neuen Studiengänge und ihr studienbegleitendes Prüfungssystem verlangten eine wesentlich höhere Prüfungsdichte als die alten Diplom- und Magister-Studiengänge [...]. Tatsächlich wurden [...] vielerorts für einzelne Veranstaltungen innerhalb eines Moduls mindestens eine, zumeist mehrere Prüfungen angesetzt“ (Winter, 2009, S. 52). Dies scheint sich auch auf das Abbruchverhalten der Studierenden auszuwirken. Vor der Einführung des neuen Systems wurde im Schnitt nach dem vierten Semester das Studium abgebrochen. Seit der vollständigen Umsetzung dieser Maßnahme des Bologna Prozesses – hier wird das Jahr 2010 als Zeitmarke genannt – findet ein Abbruch im Schnitt nach dem ersten oder zweiten Semester statt (Heublein, 2014). Auch in den Abbruchintentionen lässt sich dieser Ansatz wiederfinden:

Denn rund „jeder dritte Studienabbrecher ist in erster Linie aufgrund von Leistungsproblemen gescheitert (30 %). Damit ist Überforderung mit Abstand der wichtigste Abbruchgrund. Die Studienabbrecher aus Leistungsgründen schaffen es nicht, die fachlichen Anforderungen (5 %) oder die Menge des dargebotenen Stoffes (2 %) zu bewältigen. Die Überforderung zeigt sich aber vor allem auch darin, dass Prüfungen endgültig nicht bestanden werden und die Studierenden sich gezwungen sehen, aus ihrem Studium auszusteigen. Dies ist bei 13 % aller Studienabbrecher der Fall“ (Heublein et al., 2017, S. 24).

Allerdings muss man mitunter auch aus gesellschaftlicher und volkswirtschaftlicher Sicht bei leistungsbedingten Abbrechenden zwischen zwei theoretischen Grundtypen unterscheiden (Young, Schaffer, James & Gallardo-Williams, 2021). Auf der einen Seite finden sich Studierende, die nicht in der Lage sind, den Stoff (in einer rea-

listischen Zeit) zu erlernen und zu festigen, auf der anderen Seite finden sich Studierende, deren schlechte Prüfungsergebnisse nicht von einem Mangel an mentaler Kapazität, sondern von einem Mangel an Unterstützung bei der Überwindung von Hürden im Lernprozess herrühren. Während im ersteren Falle schlechte Bewertungen, die dann unter Umständen zu einem Abbruch führen, vollständig gerechtfertigt sind, stellt das Ausscheiden von Studierenden des zweiten Typs, also potentiell guten Fachleuten, sowohl auf der Ebene des Individuums in Sachen unnötig aufgewandter Lebenszeit, wie auch auf Ebene der Gesellschaft, hinsichtlich der ertraglos gebliebenen Aufwendung von Hochschul- und damit Gesellschaftsressourcen einen kapitalen Verlust dar (Young et al., 2021). Dies gilt insbesondere für den Fall, dass dieser Verfehlung an Ausschöpfung an vorhandenem Humankapital eine Systematik zugrunde liegt, welche sich auf Basis der sehr umfangreichen Forschung zum Studienabbruch aus dem Kreis um Heublein zumindest nicht unwahrscheinlich darstellt (Heublein et al., 2005; Heublein, 2014; Heublein et al., 2017; Heublein et al., 2020; Neugebauer et al., 2019), wie man auch anhand der folgenden Grafik sehen kann.



DZHW-Studienabbruchstudie 2016

Abbildung 2: Ausschlaggebende Studienabbruchmotive nach Abschlussart, unabhängig vom Studiengang (Heublein et al., 2017, S. 23)

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

Während sich die ausschlaggebenden Abbruchsursachen zwischen Bachelorstudium und Staatsexamen in vielen Belangen die Waage halten, erkennt man bei dieser Selbstauskunft der Studierenden von Staatsexamen hin zum Bachelor zwar erfreulicherweise einen Rückgang hinsichtlich der familiären Situation und der finanziellen Gründe, was prinzipiell ein Indiz dafür sein könnte, dass der sozialen Herkunft der Studierenden durch geänderte Studienbedingungen eine weniger starke Rolle zukommt. Erstaunlich ist hier allerdings, dass sowohl Leistungsprobleme als auch die Studienmotivation, also zwei Aspekte, denen durchaus auch ein Zusammenhang mit der Gestaltung des Studiums seitens der Hochschule unterstellt werden kann, die Spitze der aufgeführten Abbruchmotive bilden.

Heublein et al. mutmaßen hier, dass für Studierende

„die relativ frühe Konfrontation mit fachlichen Defiziten sowie mit ihrer persönlichen Leistungsfähigkeit auch zu einem frühzeitigen Rückgang der Fachidentifikation führt. Die Möglichkeit, in dieser Situation durch einen Abbruch die einstige Entscheidung für einen Bildungsweg zu korrigieren, nutzen Bachelorstudierende vermutlich häufiger, da sie bis dato weniger Zeit in das Studium investiert haben“ (Heublein et al., 2017, S. 23–24).

Grundsätzlich kann man daher an den Universitäten von einem *Performance Problem* in den MINT-Fächern sprechen, da für 31 % der Abbrechenden Leistungsprobleme der entscheidende Grund sind und für 84 % der Abbrechenden dem Leistungsproblem mindestens eine „eher große Rolle“ bei der finalen Abbruchsentscheidung zu kommt (Heublein et al., 2017). Bei einer Abbruchquote von 47 % im Jahre 2016, wie im aktuell letzten Bericht zum Studienabbruch (Heublein et al., 2020) aufgeführt, ist das sowohl mit Blick auf die Zukunft hinsichtlich der Bedeutsamkeit der innovierenden Berufe wie auch mit Blick auf die damit ertraglos aufgewandten Bundes- und Landesressourcen im Hier-und-Jetzt ein alarmierender Befund. Heublein et al. (2017) weisen zudem in Bezug auf das Leistungsproblem darauf hin, dass die abbruchsfördernde Wirkung besonders dann zum Tragen kommt, wenn die Lücken im Vorwissen nicht zu Beginn des ersten Semesters geschlossen werden können. Damit kommt der Studieneingangsphase bereits aus dieser sehr fokussierten Perspektive auf die Leistungsproblematik eine Schlüsselrolle zu.

1.2 Hürden und Herausforderungen in der Studieneingangsphase

Die Schlüsselrolle der Studieneingangsphase begrenzt sich nicht nur auf die Überwindung von Leistungsproblemen, sondern auch auf weitere ineinander verzahnte

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

Bereiche, unter anderem auch die der im akademischen Rahmen stattfindenden Sozialisationsprozesse. Fachlich kompetente Mitstudierende (Peers) zu finden, um beispielsweise Lerngruppen zu bilden, scheint für die Studierenden im ersten Jahr hinsichtlich der sozialen Perspektive die größte Herausforderung zu sein (Trautwein & Bosse, 2017). Gleichzeitig berichten besonders Studienabbrechende von Schwierigkeiten, ein gutes Verhältnis zu den Lehrenden aufzubauen. Zudem konnte gezeigt werden, dass besonders von diesen Studierenden auch signifikant seltener Feedback von den Lehrenden eingefordert wird (Heublein et al., 2017). Dies rekurriert möglicherweise auf die gewohnten Verhaltensmuster, die über zwölf Jahre im institutionellen Kontext der Schule erworben und gefestigt worden sind. Rückmeldungen der Lehrenden und auch der Mitschülerinnen und Mitschüler zu den eigenen Lernprodukten werden an der Universität nicht mehr bereitwillig zur Verfügung gestellt oder in Form von Diskussionen im Unterricht angeleitet, sondern müssen nun aktiv eingefordert werden (Rost, 2018).

Denn „die Kommunikation mit den Lehrenden ist kein einseitiger Kontakt [mehr], sie ist nur dann erfolgreich, wenn auch die Studierenden bewusst den Erhalt von Feedback und Rat der Lehrkräfte anstreben. Eine Lehrsituation, die auf solche Kommunikation, auch neben den Lehrveranstaltungen, aufbaut, ist ein wichtiger Studienerfolgsmoment“ (Heublein et al., 2017, S. 154).

Grundsätzlich ist hier jedoch nochmals darauf zu verweisen, dass sich Studierende in der Studieneingangsphase in einer komplexen Umbruchsphase befinden. „Dass Studierende sich involvieren müssen, um ihrerseits involviert und dadurch sozialisiert zu werden ist hochplausibel – oder trivial“ (Huber, 1991, S. 427). Die aktive Kontaktaufnahme ist aber die letzte der drei von Tinto bereits 1988 beschriebenen Phasen der Integration in das Studium. Die erste Phase wird als überwiegend durch die Lösung vom früheren Umfeld geprägt beschrieben, während die zweite Phase einen Übergangsprozess darstellen soll, in dem neue Normen und Verhaltensmuster übernommen werden. Die dritte Phase zeichnet sich dann durch darauf basierende soziale Einbindungsprozesse aus (Tinto, 1988). „Hier wird die Kontaktaufnahme zu anderen Studierenden und zu Lehrenden zu einem entscheidenden Faktor für den Verbleib an der Hochschule“ (Saraceni & Müller, 2011, S. 237). Das heißt allerdings, dass die Kontaktaufnahme zu den Mitstudierenden wie auch den Hochschullehrkräften zu Beginn des Studiums nicht sichergestellt ist, da zunächst einmal die Einbindungsprozesse anlaufen müssen, bevor diese wirksam werden können. Das hängt natürlich auch mit den individuellen Persönlichkeitsmerkmalen der Studierenden zusammen, die als Big-Five auch im Studienabbruchmodell von Heublein (2014) berücksichtigt werden. Prinzipiell kann aber zu Beginn des Studiums und damit zu Beginn

der allgemein als entscheidend gesehenen Studieneingangsphase nicht davon ausgegangen werden, dass oben genannte Strukturen ad hoc bestehen, sondern vielmehr über die Studieneingangsphase aufgebaut werden müssen. Damit starten die Studierenden mit einem potentiellen Mangel an Feedbackgelegenheiten in das Studium. Feedback bildet jedoch eine der wichtigen Einflussgrößen auf Lern- und Motivationsprozesse (Hattie & Timperley, 2007) – dies wird im zweiten Kapitel ausführlich erklärt – wodurch sich eine direkte Verbindung zum Leistungsproblem ergibt.

Bei genauerer Überlegung ergibt sich aus dem Zusammenspiel von *Feedback-Problem* und dem oben angesprochenen *Performance-Problem* ein regelrechtes Dilemma. Studierende erhalten Feedback nämlich nur, wenn sie es aktiv einfordern. Dabei sehen sich die Studierenden sowohl einer sozialen (Tinto, 1988) als auch einer fachlichen Hürde gegenübergestellt. Einerseits müssen sie sich erst einmal trauen, nachzufragen – hier spielt auch die Angst sich zu blamieren eine Rolle – andererseits muss das Problem dazu erst einmal selbstständig formuliert oder zumindest verbal eingegrenzt werden. Genau das setzt bereits ein gewisses Maß an Vorwissen voraus³. Gelingt es nicht, das Problem verständlich zu formulieren, kann es zudem passieren, dass langfristig auch das Fähigkeitsselbstkonzept⁴ der jeweiligen Lernenden darunter leidet. Denn dann bleiben sowohl Kompetenzerleben hinsichtlich der Aufgabenlösung als auch hinsichtlich des Hilfesuchens mit einer gewissen Regelmäßigkeit aus (Ryan & Deci, 2017). Auch das kann zu dem von Heublein et al. (2017) erwähnten Rückgang der Fachidentifikation führen.

1.3 Fachspezifische Forschungsergebnisse

Bereits im Jahre 2010, also gut zehn Jahre nach den Beschlüssen des Bologna Prozesses, wurden detaillierte, auf den Bachelorstudiengang Chemie in Deutschland fokussierte Ergebnisse publiziert, die nahelegen, dass bereits bei Studienbeginn fachspezifische Kenntnisse gefordert werden, die nur dann Vorliegen, wenn ein Chemie Leistungskurs in der Schule belegt wurde (Heublein, Hutzsch, Schreiber, Sommer & Besuch, 2010). Damals wie heute sind Leistungsprobleme der Hauptgrund für einen

³ Laut Schuhmacher (2008) muss bei mathematischen Aufgaben beispielsweise ein konzeptuelles Verständnis gewährleistet sein, damit Fehler überhaupt selbstständig gefunden werden können.

⁴ Hellmich (2011) versteht unter dem „Fähigkeitsselbstkonzept [...] die *Gesamtheit der kognitiven Repräsentationen eigener Fähigkeiten*. Dies schließt Vorstellungen über Höhe, Struktur und Stabilität der eigenen Fähigkeiten ein. Affektiv-evaluative Bewertungen der eigenen Fähigkeiten werden dagegen nicht dem Fähigkeitsselbstkonzept zugeordnet“.

Studienabbruch. Das *Performance-Problem* ist somit auch spezifisch für das Fach Chemie schon sehr lange bekannt. Ein Blick auf die Fachwissensausprägung in der Allgemeinen Chemie der ALSTER Kohorte 2016/2017 zeigt, dass der eklatante Nachteil der Studierenden ohne Chemie Leistungskurs in der Schule weiterhin besteht. So konnte für den Wissenszuwachs in der Allgemeinen Chemie im ersten Semester in Abhängigkeit von der Kurswahl (vgl. Abbildung 3) zwar gezeigt werden, dass jede Gruppe dazulernt, allerdings zeigt sich, dass die Studierenden, die keinen Chemie(leistungs-)kurs in der Schule besucht hatten, den Vorsprung der Mitstudierenden mit Leistungskurs nicht aufholen.

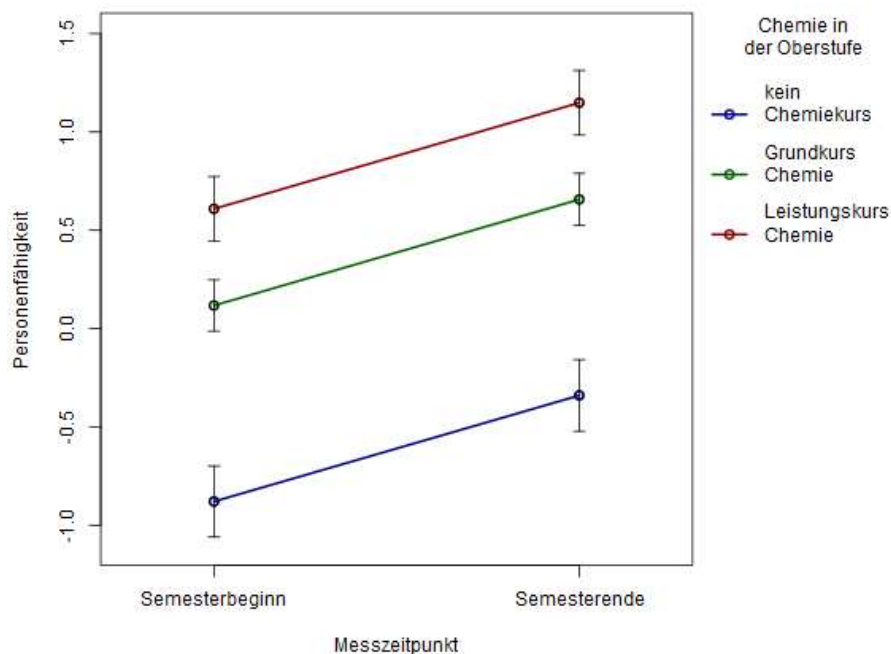


Abbildung 3: Fachwissen in der Allgemeinen Chemie zu Beginn und zum Ende des ersten Semesters in Abhängigkeit zur Kurswahl in der Oberstufe (nach Averbek, 2020)

Im Mittel erreicht sogar die Gruppe der Studierenden, die einst einen Chemie Grundkurs in der Schule besucht hat, gerade einmal das Eingangsniveau der Studierenden mit Leistungskurswahl am Ende des ersten Semesters (vgl. Abbildung 3). Das legt den Schluss nahe, dass es bestehende Hochschullernangebote nicht vollbringen, Defizite über das erste Semester hinreichend aufzuarbeiten oder gar alle Studierenden auf ein vergleichbares Niveau zu bringen (Schwedler, 2017). Interessant ist zudem, dass im Rahmen einer qualitativen Studie an der Universität Bielefeld festge-

Kapitel 1: Studienabbruchforschung und Studieneingangsphase

halten werden konnte, dass sich Studierende z. B. hinsichtlich der Vorlesung zur Allgemeinen Chemie im Falle von Vorwissensdefiziten weniger über die qualitative Überforderung mit dem Stoff der nachzuarbeitenden Inhalte als über den erhöhten Zeitaufwand der Nacharbeitung beklagen (Schwedler, 2017).

„Beschäftigt man sich mit der Erweiterung des Wissens durch Lernen, steht die Frage nach der Struktur des bereits vorhandenen Wissens ebenfalls im Mittelpunkt. Für den Aufbau von Wissensstrukturen müssen die Inhalte fachlich sinnvoll und für das Individuum schlüssig kognitiv verarbeitet werden und dabei effizient in das bestehende Wissensnetz integriert werden“ (Fischer, Glemnitz, Kauertz & Sumfleth, 2007, S. 659).

Von Renkl (1996) wird Vorwissen als die Summe aus Kenntnissen („Wissen, dass“) und Fertigkeiten („Wissen, wie“) einer Person in einem bestimmten Gegenstandsbe-
reich (Domäne) beschrieben. Auf der Internetseite des Landesinstituts für Schule (QUA-LiS NRW) findet sich folgender Ausschnitt:

„[Als Bruttounterrichtszeit ergeben sich] [...] für die Einführungsphase 90 Unterrichtsstunden, für den Grundkurs in der Q1 ebenfalls 90 und in der Q2 60 Stunden und für den Leistungskurs in der Q1 150 und für Q2 90 Unterrichtsstunden“ (QUA-LiS NRW, 2022).

Demnach ergibt sich für einen Chemie Leistungskurs ein Zugewinn an Lernzeit, die der gesamten Lernzeit eines Grundkurses in der Einführungsphase entspricht. Im Rahmen eines Leistungskurses ist folglich viel mehr Zeit gegeben, durch kumulatives Lernen ein Wissensnetz aufzubauen. Fischer et al. (2007) resümieren beispielsweise die Ergebnisse von PISA und TIMSS und indizieren, dass „[m]athematischer und naturwissenschaftlicher Unterricht durch eine geringe vertikale Vernetzung gekennzeichnet“ ist, was sich laut Baumert et al. (2002) in Problemen beim Wissenstransfer widerspiegelt. Neben der Tatsache, dass den Studierenden, die früher einen Leistungskurs besucht haben, grundsätzlich viel mehr domänenspezifische Lernzeit zu unterstellen ist, werden im Rahmen eines Grundkurses manche Teilbereiche der Chemie nicht nur weniger ausführlich, sondern manche auch überhaupt nicht thematisiert. Die geringere Lernzeit bietet im Grundkurs in jedem Fall deutlich weniger Möglichkeiten, Zusammenhänge im selben Umfang wie in einem Leistungskurs zu vermitteln.

Das legt den Schluss nahe, dass es für einige Studierende nicht kognitiv zu anspruchsvoll ist, den Stoff zu verstehen, sondern es ihnen aufgrund verschiedener Faktoren im Rahmen der vorgegebenen Zeit nicht möglich ist, die Defizite im Vorwissen aus eigener Kraft unter den gegebenen Bedingungen aufzuholen. Die Auswirkungen von Performanz-Defiziten in der Disziplin Allgemeine Chemie lässt sich in den Längsschnittdaten der ersten Förderphase der ALSTER-Forscherguppe bereits

im zweiten Semester sehr deutlich erkennen. So konnte hinsichtlich möglicher leistungsförderlicher Interdependenzen der einzelnen Subdisziplinen der grundständigen Chemiestudiengänge gezeigt werden, dass auch die Performance in Disziplinen des zweiten Semesters „maßgeblich von dem Wissen und den Fähigkeiten im Bereich der Allgemeinen Chemie bestimmt wird, die die Studierenden bereits während der Schulzeit erlernt haben“ (Averbeck, 2020, S. 288).

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Durch digitale Aufgaben, die adaptives bzw. fehlerspezifisches Feedback als Hilfestellung aktiv den lernenden Personen bereitstellen, könnte eine geeignete Form der Binnendifferenzierung erzielt werden, um der natürlichen Heterogenität hinsichtlich des Vorwissens in Lerngruppen mit den sich daraus ergebenden individuellen Hürden im Lernprozess angemessen zu begegnen. Um den Einfluss des Feedbacks auf Lehr- und Lernprozesse und die Bedingungen für gutes Feedback hinlänglich aufzeigen zu können, muss in einem vorherigen Schritt zunächst einmal der Begriff Lernen⁵ beleuchtet werden.

2.1 Klassische Lerntheorien, ein Überblick

In einem Artikel des *Journal of Educational Psychology* legt Mayer (1992) drei Sichtweisen dar, die kennzeichnend für das Verständnis von Lernen im Verlauf des vorherigen Jahrhunderts sind. Die drei Konzepte (1) Lernen als Erwerb von Reaktionen, (2) Lernen als Wissenserwerb und (3) Lernen als Wissenskonstruktion stehen laut Klauer und Leutner (2012) repräsentativ für die behavioristische, die kognitivistische und die konstruktivistische Psychologie.

Behaviorismus

Der Behaviorismus bildet wohl einen der ältesten und bekanntesten und zugleich auch den simpelsten Vertreter aller Lerntheorien. Lernen ist weitestgehend auf ein Reiz-Reaktionssystem reduziert. Klassische Konditionierung zurückgehend auf den Pawlow'schen Hund (1905), wie auch die später entstandenen Theorien zum operanten Konditionieren (B. F. Skinner) bzw. Theorien instrumentellen Lernens (E. L. Thorndike) sind allesamt durch eine sehr mechanistische Sichtweise geprägt, die Lernen als kausalen Prozess versteht (Caduff, Pfiffner & Bürgi, 2018, S. 15–17).

⁵ „Kann als zentraler Gegenstand der Pädagogischen Psychologie das Konstrukt „Lernen“ benannt werden, so wurden zwar bislang Strukturen und Prozesse des Lernens aus der Perspektive kognitiver, motivationaler und emotionaler Prozesse betrachtet, der Begriff Lernen selbst aber nicht definiert. Gemäß der neueren entwicklungspsychologischen Orientierung an der Lebensspanne wird Lernen als lebenslanger Prozess betrachtet. Die Fähigkeit zu Lernen gilt als das entscheidende Potential des Menschen, kulturelle Errungenschaften zu reproduzieren und Neues zu schaffen“ (Fritz & Tobinski, 2018, 222 f.).

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Dennoch lassen sich bereits hieraus wertvolle Implikationen für zukünftige Lernumgebungen ableiten. So folgern Caduff et al. (2018, S. 17), dass „Lernumgebungen [an] die Lernenden angepasst [zu] gestalten [...] wesentlich dazu bei[trägt], dass das Erleben von Erfolgen möglich ist. Zudem sollen beobachtbare Strategien und Verhaltensweisen, die zum Erfolg geführt haben, konkret beschreibend zurückgemeldet werden“. In den behavioristischen Lernverständnissen – klassisches Konditionieren, operantes/instrumentelles Konditionieren – meint eine Einflussnahme auf ein Subjekt dann das Senden von verstärkenden Reizen für gewünschtes Verhalten bzw. von hemmenden Reizen für unerwünschtes Verhalten (Kopp & Mandl, 2014). Erst durch das Modell-Lernen nach Bandura (1979) kann auch der Erwerb von komplexeren Verhaltensdispositionen erklärt werden (Kopp & Mandl, 2014). Lernen bzw. Lernprozesse finden hier allerdings keine Erwähnung, so sie denn stattfinden, sind sie in der sog. „Blackbox“ verortet, sodass ihnen keine weitere Beachtung geschenkt wird.

Kognitivistisches und konstruktivistisches Lernverständnis

Die kognitive Wende der 1950er bis 1960er Jahre rückte die Erforschung dessen, was in der sog. „Blackbox“ abläuft, also die Erforschung von Kognitionen – Prozessen der Wahrnehmung, des Denkens und der Erinnerung – in den Vordergrund und verdrängte damit den behavioristischen Ansatz aus dem Fokus der Lehr-Lernforschung, sodass man wenig später zu der Ansicht gelangte, dass der Beitrag des lernenden Subjekts den Lernprozess in besonderer Weise prägt. Lernen wurde von Konditionierung basierend auf einem passiven Reiz-Reaktions-Verhältnis zu einem aktiven Prozess des Subjekts mit vielen Subprozessen (Caduff et al., 2018; Kopp & Mandl, 2014).

Daher gehen auch einige Forschende so weit zu sagen, dass die kognitive Wende der 1960er Jahre eigentlich eine konstruktivistische war (Reusser, 2006). Klauer und Leutner (2012) betonen zwar ebenfalls, dass es im Rahmen der kognitiven Wende an der Zeit war, die Lernenden beim Lernen stärker in den Fokus zu nehmen, allerdings weisen sie im gleichen Zuge darauf hin, dass die Ideen und Ansätze des konstruktivistischen Verständnisses bereits in ähnlicher Form von vielen Psychologen und Pädagogen vertreten wurden.

Ungeachtet dessen, ob diese Sichtweise nun als konstruktivistisch zu bezeichnen ist oder ob diese Sichtweise immanent als Modifikation in der kognitivistischen Psychologie enthalten ist oder umgekehrt, ist Lernen als aktiver Prozess kumulativen, ideal-

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

erweise problemlösend erfolgenden Aufbaus von immer komplexer werdenden Wissens- und Denkstrukturen zu verstehen (Caduff et al., 2018; Klauer & Leutner, 2012; Reusser, 2006). Der Transfer bzw. die Anwendung von Wissen funktioniert dann besonders gut, wenn Lernende das Wissen vorher selbst durch eigenständiges Problemlösen oder problemorientierten Nachvollzug als Beziehungsnetze bzw. mentale Modelle konstruiert und an Beispielen durchgearbeitet haben (Reusser, 2006).

„Das bedeutet, dass Wissen nicht von einer Person auf die andere übertragen werden kann, sondern von jedem Menschen neu konstruiert wird und so dessen subjektive Realität darstellt. Diese ist unter anderem abhängig von seinem Vorwissen und seinen Einstellungen“ (Caduff et al., 2018, S. 20).

Wenn von einem Erkennen oder Wiedererkennen die Rede ist, dann tritt dadurch das lernende Subjekt zwangsläufig in Interaktion mit der Welt. Während die kognitiv-konstruktivistischen Sichtweise von einer Interaktionsbeziehung von dem Subjekt als solo-lernende Person mit der gegenständlichen Welt ausgeht, geht der sozio-kulturelle Konstruktivismus davon aus, dass Lernen auch durch personale und soziale Impulse, also durch Widersprüche, Meinungsdivergenzen und produktive Lerndialoge, also reziprok und damit (dialogisch) ko-konstruiert erfolgt (Reusser, 2006).

„Der soziale Aufbau von Wissens- und Weltbildstrukturen wird durch kognitive und kulturelle Werkzeuge und Symbolsysteme unterstützt. Neben der Alltagssprache gehören die wissenschaftlichen Symbolsysteme (Logik, Zahlensystem, naturwissenschaftliche Repräsentationsformen) ebenso dazu wie die Rituale und symbolischen Interaktionsformen in Erziehung, Politik, Religion, Literatur und Kunst“ (Reusser, 2006, S. 156).

Wissenschaftliche Symbolsysteme nehmen in der Chemie einen besonderen Stellenwert ein (Dickmann, 2019), da diese Disziplin der Naturwissenschaften – vielleicht mehr als jede andere – durch Symbolsysteme verschiedener Arten konstituiert ist. Jedoch müssen auch diese Symbolsysteme zunächst einmal erlernt und in Wissensstrukturen eingebunden werden. Damit gehen Lernprozess und Informationsverarbeitungsprozess miteinander einher, so sie denn überhaupt für den Menschen voneinander trennbar sind.

Betrachtet man neuere Gedächtnismodelle, die den Informationsverarbeitungsprozess klar in den Fokus setzen, werden die oben genannten Aspekte besonders deutlich. Informationsbearbeitung ist immer ein aktiver, individueller Prozess, der durch Wissen aus dem Langzeitgedächtnis – Vorwissen – bedingt und zum Teil auch gesteuert wird. Abbildung 4 stellt ein solches Informationsverarbeitungsmodell dar, dass die im bisherigen Verlauf vorgestellten Sichtweisen verbindet und die einstige

„Blackbox“ des Behaviorismus mit Erkenntnissen über (mögliche) Subkomponenten (Sensorisches, Arbeits- und Langzeitgedächtnis) und Prozessmodellierungen füllt.

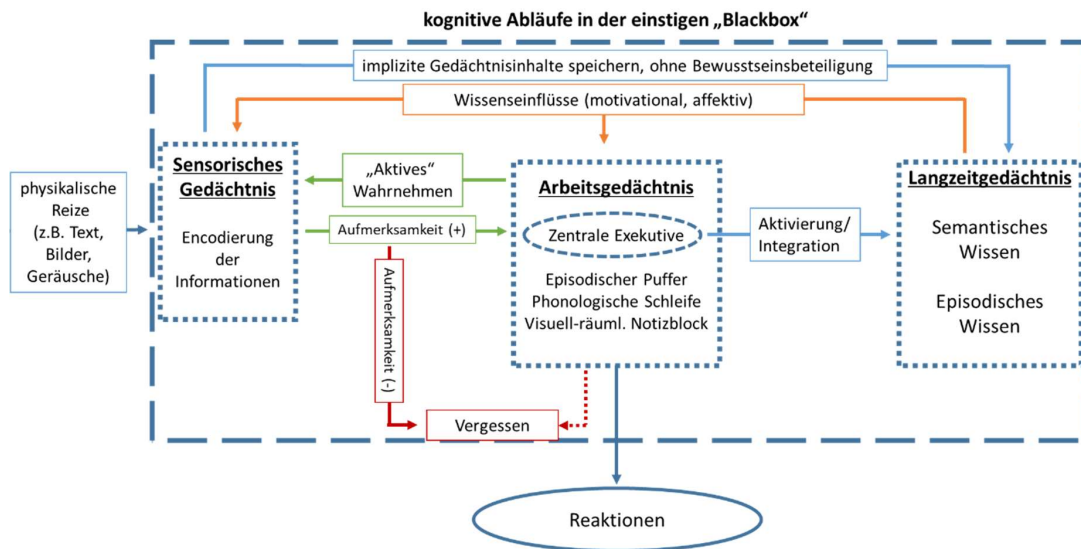


Abbildung 4: Informationsverarbeitungsmodell des Gedächtnisses adaptiert nach Woolfolk (2014, S. 280), Hussy, Fritz und Tobinski (2018, S. 84) und Moreno und Mayer (2007, S. 314)

Im sensorischen Gedächtnis werden die physikalischen Reize, die auf das Individuum einprasseln, bereits auf Basis von Wissensseinflüssen aus dem Langzeitgedächtnis encodiert. Durch aktive Wahrnehmung wird dieser Informationspool gescannt.

„Alle drei Elemente des Systems interagieren miteinander, um die Wahrnehmung zu steuern; sie repräsentieren, organisieren und interpretieren die Informationen; sie nutzen und ändern Propositionen, Konzepte, Vorstellungsbilder, Schemata und Strategien; sie konstruieren Wissen und lösen Probleme“ (Woolfolk, 2014, S. 280).

Die Informationen, die in den Fokus der Aufmerksamkeit gelangen, werden im Arbeitsgedächtnis durch verschiedene registerspezifische Vorgänge aufrechterhalten und führen (falls vorhanden/abrufbar) zu einer Aktivierung spezifischer Wissensstrukturen. Unter der Beachtung von diesen Wissensseinflüssen aus dem Langzeitgedächtnis wird die Information dann verarbeitet. Mit diesen können Sie natürlich konfliktieren oder diese hinsichtlich ihrer „Passung“ bestärken. Im Idealfall wird dann das Resultat dieses Verarbeitungsprozesses im Langzeitgedächtnis in bereits bestehende Wissensstrukturen integriert (Fritz & Tobinski, 2018; Woolfolk, 2014).

Die (pädagogische) Psychologie und die Kognitionswissenschaft haben über die Zeit viele der im Modell aufgezeigten Prozesse und Subeinheiten identifizieren können.

Diese verbleiben allerdings ohne komplexe Messverfahren und kontrollierte Laborsettings verborgen. Wenn es also darum geht, jemanden in seinem Lernprozess zu unterstützen, dann ist Intervenieren besonders dann sinnvoll, wenn in einer Lernsituation (mit spezifischen Reizen) eine „unerwünschte“ Reaktion des lernenden Subjekts, gemeinhin bekannt als „Fehler“, vorausgeht. In der Regel wird in solchen Fällen dann eine Rückmeldung gegeben, um diese unerwünschte Reaktion zur gewünschten zu wandeln. Informationen, die Lernende aus einer solchen Rückmeldung erhalten, werden daher im Rahmen des kognitivistischen bzw. konstruktivistischen Ansatzes im Vergleich zum behavioristischen weniger als Verstärker gesehen sondern als Informationsquelle, die zur Regulierung der Lernaktivitäten herangezogen wird bzw. nach konstruktivistischem Verständnis herangezogen werden kann, da Letzteres einen deutlicheren Fokus auf die Systemexploration des Individuums setzt (Kopp & Mandl, 2014). Eine solche Rückmeldung findet daher idealerweise, aber nicht zwingend Beachtung. Wann also ein Lernprozess stattgefunden hat, ist für Forschende im Feld tatsächlich auf die behavioristische „Blackbox“-Sichtweise limitiert, denn „[d]er eigentliche Lernprozess ist nicht direkt beobachtbar, sichtbar ist nur das Resultat der Lernprozesse in Form von Verhaltensänderungen, vorausgesetzt diese werden gezeigt“ (Fritz & Tobinski, 2018, S. 223).

2.2 Der Begriff Feedback, eine definitorische Klärung

Eine Rückmeldung wird in der Lehr-Lernforschung in der Regel mit dem Begriff „Feedback“ synonym verwendet. Es gibt zwar einige, die eine Unterscheidung dieser beiden Begriffe anstreben (Ditton & Müller, 2014), doch findet sich diese Unterscheidung in gängigen Feedback-Theorien und Metastudien zu diesem Thema nicht wieder (Butler & Winne, 1995; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Narciss, 2006, 2012; Wisniewski, Zierer & Hattie, 2019). Entlang der Forschung zu verschiedenen Lerntheorien wurde der Feedbackbegriff, wie zuvor erwähnt, aus unterschiedlichen Perspektiven beleuchtet. Im Rahmen der behavioristischen Ansätze steht überwiegend die stärkende Wirkung von positivem Feedback auf Stimulus-Response-Assoziationen im Mittelpunkt, die dann zu einer Erhöhung der Auftretenswahrscheinlichkeit des verstärkten Verhaltens führt (Kopp & Mandl, 2014). Bei kognitivistischen als auch in konstruktivistischen sowie instruktionspsychologischen Lerntheorien verlagert sich der Fokus auf den Informationsgehalt von Feedback und darauf, welche kognitiven Voraussetzungen auf Seiten des lernenden Subjektes gegeben sein müssen, um mit spezifischen Feedback-Inhalten am besten die Korrektur

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

von Fehlern zu unterstützen (Narciss, 2018). So erscheint folgende Definition von Feedback als valide, um zunächst Feedback im weiten Sinne zu beschreiben:

„Feedback is all post-response information which informs the learner on his/her actual state of learning or performance in order to regulate the further process of learning. This information can be provided by external sources of information (i.e., teachers, peers, parents, computer-based trainings), and by internal sources of information (i.e., information perceivable by the learner while task processing)“ (Narciss, 2012, S. 1285).

Wichtig ist hier noch einmal zu erwähnen, dass Feedback *sämtliche* Informationen umfasst, die dem lernenden Subjekt zurückgemeldet werden⁶. Feedback kann folglich auf zweierlei Weisen erfolgen. Unter *indirektem* bzw. *impliziten* Feedback subsumieren sich all jene nicht intendierten, alle unbeabsichtigten und unbewussten Rückmeldungen durch verbale, nonverbale und paraverbale Signale (Lehmenkühler, Roscher & Theis, 1976). Feedback, das aus einer intendierten (verbalen) Rückmeldung zu einem wahrgenommenen Verhalten resultiert, wird in der Literatur als explizites Feedback bezeichnet. Wird das *explizite* Feedback durch richtungsweisende Komponenten für künftiges Verhalten wie konstruktive Kritik an den Adressaten angereichert, so kann von *konstruktivem* Feedback als Spezialfall des *expliziten* Feedbacks gesprochen werden.

„Damit ein Lehr-Lern-Prozess erfolgreich ablaufen kann, muss die Lehre bestimmte Effekte erzielen. Diese sechs Lehrfunktionen [siehe Abbildung 5] sind erforderlich, soll die Lehre wirksam sein: Die Steuerung und Kontrolle des Prozesses, die Motivation der Lernenden, ihre Informierung, ferner die Informationsverarbeitung, die Speicherung und Abrufbarkeit der Information und der Transfer, d. h. ihr Einsatz in neuen Zusammenhängen“ (Klauer & Leutner, 2012, S. 48).

Klauer und Leutner betonen, dass gerade Feedback im Kontext des multimedialen Lernens eine besondere Bedeutung zukommt und zukommen wird. Sie verweisen hier auf die besondere Rolle als Lernhilfe, die z. B. speziellen Feedbacktypen zusätzlich zur Motivationsfunktion von klassischem Feedback zukommt (Narciss, 2006), so es denn richtig gestaltet ist. Insbesondere im Zusammenhang mit anspruchsvollen und schwierigen Aufgaben sehen sie einen möglichen Vorteil von informativem tutoriellem Feedback gegenüber klassischem Feedback – korrekatives Feedback und Musterlösungen – (Klauer & Leutner, 2012), welches laut Hattie und Timperley (2007) gerade bei komplexeren und schwierigeren Aufgaben Erfolgseinbußen verzeichnet.

⁶ „One cannot *not* communicate“ (Watzlawick, Bavelas & Jackson, 1967, S. 51).

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

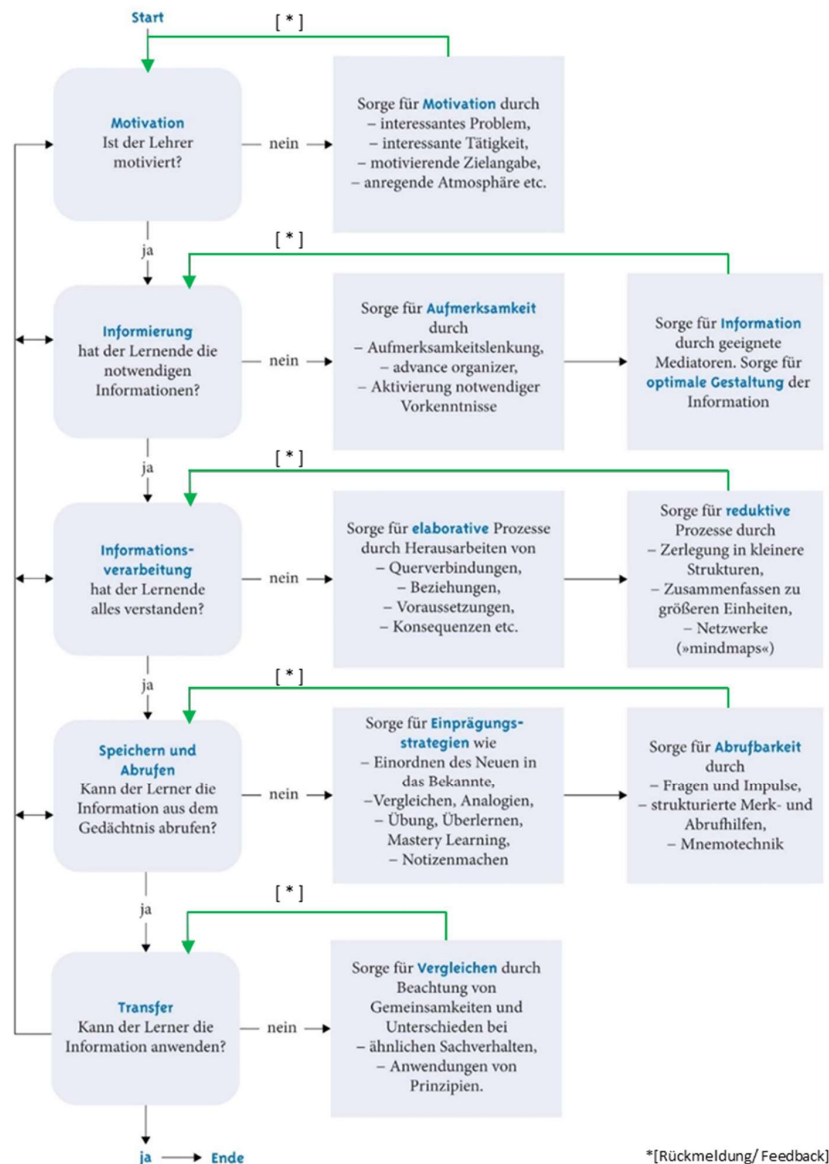


Abbildung 5: Der Lehlgorithmus (nach Klauer & Leutner, 2012, S. 47)

Dass mit Lernhilfe gemeint ist, dass Feedback als ein steuernder Faktor auf jeder Stufe des Lehlgorithmus' (Abbildung 5) also als Respondus der Lehrperson wirken kann, zeigt sich in Tabelle 1, in der Narciss (2006) die diversen Funktionen zusammenträgt, die Feedback über die Jahre von verschiedenen Forschenden zugeschrieben wurden.

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Tabelle 1: Übersicht über postulierte Feedback-Funktionen

Cusella, 1987	Sales, 1993	Wager & Mory, 1993	Butler & Winne, 1995
• Verstärken • Informieren • Hinweisen • Motivieren	• Stimulieren • Informieren • Lenken • Motivieren	• Bestätigen • Informieren • Hinweisen • Motivieren • Korrigieren	• Bestätigen • Informieren • Hinweisen • Korrigieren
• Regulieren • Instruieren	• Regulieren • Instruieren • Bewerten • Beraten	• Instruieren • Bewerten	• Anregen • Ergänzen • Diskriminieren • Restrukturieren

So ist es kaum verwunderlich, dass Hattie und Timperley (2007) *Feedback* in einer umfangreichen Metastudie als einen der größten Einflussfaktoren – positiv wie auch negativ – auf Lern- und Motivationsprozesse klassifizieren konnten.

Wenn man Feedback also als Teil der Lehrfunktionen im Lehralgorithmus bzw. generell in Lehr-Lernprozessen versteht, dann kann mit Feedback im Kontext der Lehr-Lernforschung bzw. Instruktionsforschung nur sinnvoll das *explizite* Feedback gemeint sein. Denn *zielgerichtetes*, also auf einen funktionalen Effekt *abzielendes* Handeln (Klauer & Leutner, 2012) bedeutet nichts anderes als geplantes und damit bewusstes Handeln.

2.3 Feedback in Lernsituationen

Bangert-Drowns, Kulik, Kulik und Morgan (1991) beschreiben Feedback als essentielles Konstrukt für viele der bestehenden Lehr-Lerntheorien. Ein Verständnis der Bedingungen für effektives Feedback sollte daher sowohl die theoretische Entwicklungsphase des Feedbacks als auch die instruktionale Praxis fördern (Bangert-Drowns et al., 1991). Mitunter deshalb steht Feedback auch seit ca. einem Jahrhundert im Fokus von Untersuchungen (Narciss, 2014). Die Ergebnisse weisen, wie bei Hattie und Timperley angeklungen, nicht nur ein breites, sondern ein sich über den positiven und negativen Bereich erstreckendes Spektrum von Effektstärken auf. Eine

Moderatoren-Analyse zeigte, dass die Effektstärke substanziell von dem übermittelten Informationsgehalt abhängt (Wisniewski et al., 2019). Teilweise berichten die vorliegenden Metastudien auch von inkonsistenten und partiell auch sich widersprechenden Ergebnissen beim Vergleich von Einzelstudien (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Wisniewski et al., 2019). Die Heterogenität in der Datenlage zeigt also, dass Feedback nicht als Konstrukt mit eindeutiger Wirkungsweise gesehen werden kann und darf. Um ein Verständnis für die Wirkweise von Feedback zu gewinnen, sollen daher im Folgenden einige bekannte Feedback-Modelle vorgestellt und diskutiert werden, um entscheidende Aspekte für diese Arbeit herauszuarbeiten.

Als Grundlage hierfür soll das Feedbackmodell von Butler und Winne (1995, siehe Abbildung 6) dienen, das versucht diese inkonsistenten Forschungsbefunde vor dem Hintergrund von Modellen des selbstregulierten Lernens zu betrachten (Narciss, 2014).

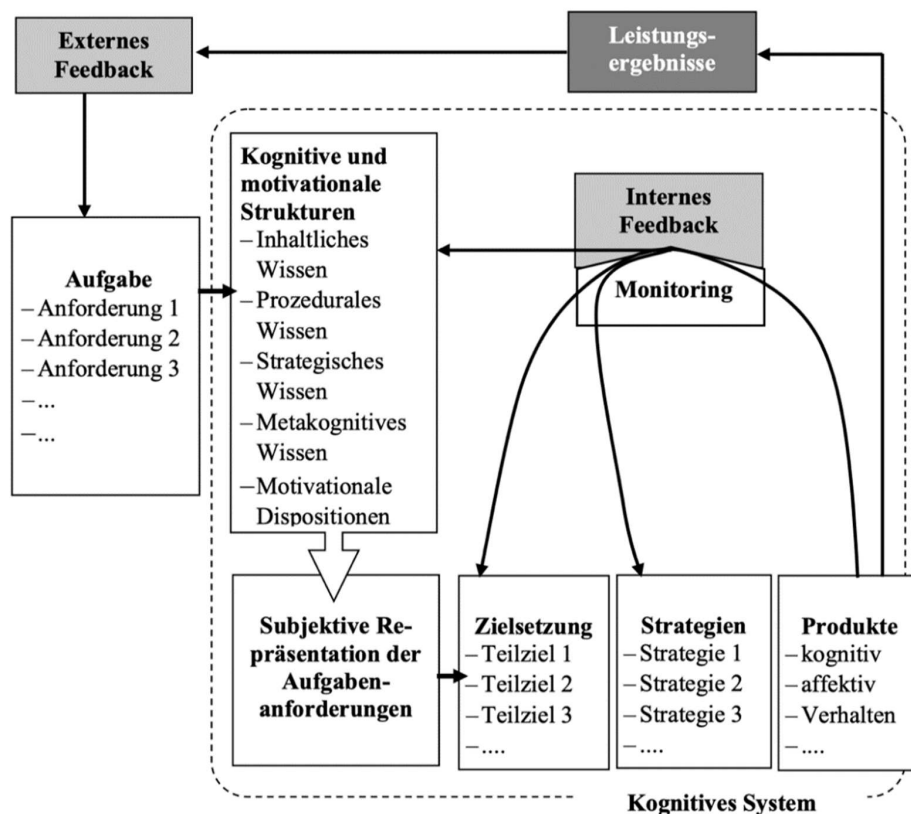


Abbildung 6: Modell des Selbstregulierten Lernens (Narciss, 2006, S. 53) nach Butler und Winne (1995)

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Die Überwachung des Lernprozesses (Monitoring) nimmt hier eine entscheidende Rolle ein und kommt beim selbstregulierten Lernen dem lernenden Subjekt selbst zu. Über Monitoring-Prozesse werden Zielsetzungen, (Lösungs-)Strategien und (Lern-)Produkte miteinander in Beziehung gesetzt und evaluiert, verknüpft bzw. mit bestehenden oder neuen Wissensstrukturen verwoben. Butler und Winne (1995) gehen dabei davon aus, dass vom lernenden Subjekt beim Bearbeiten von Lernaufgaben eine individuelle mentale Aufgabenrepräsentation auf Basis des Vorwissens generiert wird. Auf Basis dieser mentalen Repräsentation findet dann ggf. eine Sequenzierung der Aufgabenteilschritte sowie eine Auswahl von individuellen (Sub-)Zielen und Bearbeitungsstrategien statt. Ausgehend von Modellen selbstregulierten Lernens heben sie daher die Monitoring-Funktion des Feedbacks hervor.

Aus diesem Grund führten Butler und Winne (1995) die für ihr Modell zum selbstregulierten Lernen notwendige Unterscheidung zwischen *externem* und *internem* Feedback ein (siehe Abbildung 6). Die Anwendung der ausgewählten Verhaltensstrategien führt zu Lernprodukten. Das sind zum einen konkrete Handlungsergebnisse, wie z. B. die Aufgabenlösung, zum anderen aber auch mentale Produkte wie z. B. kognitive oder affektive Veränderungen des individuellen Wissens. Das interne Feedback ist dabei als Rückmeldung des Individuums an sich selbst zu verstehen, ob die Ziel- und Strategieauswahl zu einem adäquaten Produkt geführt haben⁷. Wenn dann noch externe Feedbackinformationen angeboten werden, werden sie demnach mit dem beim Monitoring generierten internem Feedback verglichen. Dieses externe Feedback kann dabei sowohl bestätigend oder ergänzend wirken, als auch widersprüchlich zum internen Feedback sein (Butler & Winne, 1995). Die Wirksamkeit von externem Feedback hängt daher von einigen Faktoren ab. Narciss (2014) führt die Wirksamkeit auf der einen Seite auf die Übereinstimmung von externem und internem Feedback zurück, auf der anderen Seite ist es von besonderer Wichtigkeit, wie im Falle von Diskrepanz zwischen externem und internem Feedback mit ebendieser umgegangen wird (siehe Tabelle 2). In beiden Situationen, aber insbesondere bei der letzteren, ist die inhaltliche Qualität des Feedbacks von entscheidender Bedeutung: Wie sehr vertraut die lernende Person auf sich selbst?⁸ Wie sehr

⁷ Ob nun dem Monitoring ein Feedback des Lernenden an sich selbst entspringt oder genau diese Rückmeldung an sich selbst die Monitoring-Prozesse ausmachen, bleibt hier allerdings unklar. So könnten Selbstevaluation und internes Feedback möglicherweise in einem Prozess-Produkt-Verhältnis zueinander stehen.

⁸ Studien zur Einschätzung des eigenen Wissensstandes (*illusion of knowing*, z. B. Glenberg, Wilkinson & Epstein, 1982) zeigen, dass es eine interindividuelle Variabilität solcher Einschätzungen gibt, welche sich unter anderem durch personenspezifische Merkmale (z. B. Vorwissen, motivationale Orientierung, Alter, Fähigkeitsselbstkonzept) erklären lässt (Narciss, 2006).

vertraut das lernende Subjekt auf die externe Feedbackquelle⁹? In der Regel versucht die lernende Person dann diese Diskrepanz zu überwinden, wobei natürlich gewünscht ist, dass je nach Tragweite der Fehler die Fehlvorstellung durch einen Theoriewechsel abgelegt wird und nicht nur oberflächlich modifiziert, die Feedbackinformation ignoriert oder gar umgedeutet wird (siehe Tabelle 2).

Tabelle 2: Sieben mögliche Reaktionen auf Informationen, die den eigenen Erwartungen, Vorstellungen oder Wissensstrukturen widersprechen (Chinn & Brewer, 1993)

Typ der Reaktion	Konsequenzen der Reaktion		
	Wird das Individuum die Informationen annehmen?	Wird das Individuum die Informationen erklären können?	Wird das Individuum seine Fehlvorstellung ablegen?
Ignorieren	Nein	Nein	Nein
Ablehnen	Nein	Ja	Nein
Abwerten/ Ausschließen	Ja oder vielleicht ^a	Nein	Nein
Ausgliedern	Ja	Noch nicht ^b	Nein
Umdeutung	Ja	Ja	Nein
Oberflächliche Veränderung	Ja	Ja	Ja, teilweise ^c
Conceptual Change	Ja	Ja	Ja ^d

^a Das Individuum wird entweder diese Daten als gesichert akzeptieren oder agnostisch über die Qualität der Daten bleiben. ^b Das Individuum erwartet, dass die gegenwärtige Theorie die Informationen schon irgendwann erklären können wird. ^c Es werden nur Randprämissen verändert, die Kernprämissen bleiben unverändert. ^d Kernprämissen werden verändert und führen nach Lakatos (1970) zu einem Theoriewechsel.

Ergebnisbezogene Feedbackarten (*knowledge of result, knowledge of performance*) bieten folglich am wenigsten Unterstützung für die Selbstregulation an, da sie relativ wenige Anhaltspunkte für Handlungs- bzw. Wissensoptimierung geben. *Elaboriertes Feedback* hingegen kann helfen, die für den Lernfortschritt relevanten Informationen herauszustellen (Narciss, 2014). Unter elaboriertem Feedback kann konstruktives Feedback, also korrekatives Feedback angereichert mit strategischen Informationen für die Aufgabenbearbeitung (task validity feedback), Informationen zur Selbstregulation (cognitive validity feedback) oder auch Informationen zur Evaluation der eigenen Selbstregulation (funktional validity feedback) gemeint sein. Um in diese Fülle an verschiedenen vorgestellten Feedbackkomponenten eine Ordnung zu bringen

⁹ Die intraindividuelle Varianz der Selbsteinschätzung beruht unter anderem auch auf der Gestaltung des Lernmaterials (z. B. Menge, Strukturiertheit, Vertrautheit; Narciss, 2006).

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

und auch für weitere Überlegungen die Begriffe mit inhaltlichen Beispielen zu füllen, kann die folgende Tabelle (Tabelle 3) herangezogen werden.

Tabelle 3: Klassifikation unterschiedlicher Feedback-Komponenten nach inhaltlichen Gesichtspunkten (Narciss, 2006, S. 23)

Elaborierte Feedbackkomponenten	Bezeichnung ^a	Beispiele für Feedback-Inhalte
	Knowledge of Performance [KP]	• 15 von 20 Aufgaben richtig • 85 % der Aufgaben gelöst
	Knowledge of result/ response [KR]	• Falsch/ richtig • Nicht richtig/ stimmt
	Knowledge of correct response [KCR]	• Angabe der korrekten Antwort • Markierung der korrekten Antwort
	Knowledge on task constraints [KTC]	• Hinweise auf Art der Aufgabe • Hinweise auf Bearbeitungsregeln • Hinweise auf Teilaufgaben • Hinweise auf Aufgabenanforderungen
	Knowledge about mistakes [KM]	• Anzahl der Fehler • Ort der Fehler/ des Fehlers • Art der Fehler/ des Fehlers • Ursache/n des/r Fehler(s)
	Knowledge on how to proceed [KH] („know how“)	• Fehlerspezifische Korrekturhinweise • Aufgabenspezifische Lösungshinweise • Hinweise auf Lösungsstrategien • Leitfragen • Lösungsbeispiele
	Knowledge on meta-cognition [KMC]	• Hinweise auf meta-kognitive Strategien • Metakognitive Leitfragen

^a Bei der Wahl der Bezeichnungen wurde einerseits darauf geachtet, die in der Feedback-Forschung üblichen Bezeichnungen zu verwenden bzw. weiter zu führen, andererseits sollte die inhaltliche Feedback-Facette in der Bezeichnung deutlich werden.

Während also internes Feedback voraussetzt, dass die lernende Person über ein adäquates Maß an Vorwissen (Fachwissen und Metawissen) verfügt, da sonst die Qualität des internen Feedbacks nicht gewährleistet ist, kann über externes korrektives Feedback der lernenden Person die Diskrepanz zwischen Ist- und Sollzustand

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

aufgezeigt werden. Elaboriertes Feedback (extern) kann dann über strategische Hilfestellung und Erklärungen zur Lösung von Aufgaben führen. In zweiter Instanz kann dieses externe Feedback dann auch Hilfestellung zur Selbstevaluation der ermittelten Lösungen bzw. angewendeten Lösungsstrategien qualitativ hochwertige Informationen für den Abgleich mit dem internen Feedback (siehe Tabelle 3) liefern. Die Produkte dieser kognitiven Prozesse können dann idealerweise in bestehende Wissensstrukturen eingebunden werden.

Letztlich kann das Modell zum Selbstregulierten Lernen in Abbildung 6 als fokussierendes Informationsverarbeitungsmodell für Informationen aus Feedback im Arbeitsgedächtnis herangezogen werden, insbesondere da sich einige Aspekte aus Abbildung 4 (S. 16) in der Argumentation von Butler und Winne (1995) wiederfinden lassen:

„Internal feedback that students generate by comparing evolving states of a task to goals creates conditional knowledge that is the basis for further action... . In our model of self-regulated engagement [siehe Abb. 6], cognitive or behavioral products are created by three sequential cognitive events: (a) perceiving task conditions in terms of extant schemata that weave together knowledge and beliefs, (b) adopting goals, and (c) applying tactics and strategies. We posit that information about each of these cognitive events is, in the sense of a lens model, a cue that relates to cognitive products“ (Butler & Winne, 1995, S. 261).

Für selbstbezogene Metakognitionen bei Arbeits- und Organisationsprozessen in Unternehmen konnten im Rahmen der Feedback-Intervention-Theory¹⁰ von Kluger und DeNisi (1996) tatsächlich einige Vorhersagen hinsichtlich der Feedbackeffektivität bedingt durch die Feedbackgestaltung empirisch bestätigt werden. Grundlage der korrespondierenden Moderatorenanalyse bildeten dabei die Studien, die für ihre Metaanalyse herangezogen wurden. Diese Metaanalyse mit 607 Effektstärken und insgesamt 23663 Fällen reiht sich in die oben beschriebene Befundlage ein, denn sie zeigt ebenfalls eine zwar im Mittel positive Wirkung von Feedback, jedoch verweisen Kluger und DeNisi (1996) bereits in der Zusammenfassung des Artikels darauf, dass über ein Drittel der Feedbackinterventionen negative Effektstärken hinsichtlich der Performanz aufwiesen. Die Moderatorenanalyse konnte folgende theoriebasiert hergeleitete Implikationen für die inhaltliche Gestaltung von Feedback bestätigen, die Narciss (2006, S. 62), wie folgt, zusammenfasst:

¹⁰ Die ersten beiden Hauptannahmen der Feedback-Intervention-Theory nach Kluger und DeNisi (1996) sind, dass (1) Verhalten in der Regel zielorientiert ist und dass (2) Ziele hierarchisch organisiert sind, wobei selbstbezogene Ziele an oberster Stelle, Handlungsziele an unterster Stelle stehen.

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

- „1. Feedback-Effekte auf die Leistung sind geringer bei Feedback-Inhalten, die die Aufmerksamkeit auf selbstbezogene Meta-Kognitionen lenken (z. B. alle Formen summativen Leistungsfeedbacks; z. B. Noten, Lob, sozial vergleichende Wertungen).
2. Feedback-Effekte auf die Leistung sind größer bei Feedback-Inhalten, die die Aufmerksamkeit auf die aufgabenbezogene Motivation (z. B. alle Formen motivierenden Feedbacks) und die aufgabenbezogene Lernebene lenken (z. B. alle Formen informativen tutoriellen Feedbacks).
3. Wenn keine aufgabenbezogenen korrigierenden Informationen im Feedback angeboten werden, sondern nur ergebnisbezogene Informationen, ist der Anstrengungs- und damit der Leistungseffekt dann am größten, wenn die Aufgaben einfach oder bereits gut gekonnt sind.
4. Wenn Feedback und eine Zielsetzungsmaßnahme kombiniert werden, sind die Leistungseffekte größer als beim Einsatz nur einer Maßnahme.
5. Feedback-Inhalte, die zu wichtig erachteten selbstbezogenen Zielen passen, lenken die Aufmerksamkeit von der Aufgaben-Ebene auf die Meta-Ebene und führen daher zu Leistungsminderungen.“

Verglichen mit Feedback zur Optimierung von Arbeits- und Organisationsprozessen in Unternehmensstrukturen ist für Feedback im Kontext von Lehr-Lernsituationen allerdings ein deutlich breiteres Spektrum von Metakognitionen von Interesse (Butler & Winne, 1995).

Obwohl all diese theoretischen Ansätze auf verschiedenste Feedback-Arten (s. o.) eingehen und auch verschiedene Gestaltungsmöglichkeiten von Feedback hervorheben, bleiben sie doch einer genaueren Untersuchung der inhaltlichen Qualität schuldig. Hattie und Timperley (2007) knüpfen an die Meta-Analyse von Kluger und DeNisi (1996) an und entwerfen ebenfalls ein Feedbackmodell (Abbildung 7), welches aber von einem externen Standpunkt aus auf Feedback schaut und weniger den Fokus auf die Wirkweise im lernenden Subjekt, sondern eher auf die Gestaltung des Feedbacks legt.

Hierbei werden drei Fragen an das lernende Subjekt zu zentralen Angelpunkten effektiven Feedbacks: Feed Up – Was sind deine Ziele?; Feed Back – Wo stehst du, welche Fortschritte hast du bislang erreicht?; Feed Forward – Wohin geht es als nächstes, um die Ziele noch besser zu erreichen?

Des Weiteren differenziert dieses Modell vier Wirkebenen (Aufgaben-, Prozess-, Selbstregulations- und Selbstebene), die Anhaltspunkte für die Konzeption von Feedback liefern. Im Einklang mit der Selbstwirksamkeitstheorie von Ryan und Deci (2017) werden, wie auch bei den vorherigen Modellen, für Feedback, das sich auf

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

das eigene Selbst des lernenden Subjektes bezieht (z. B. Begabungslob) ungünstige Effekte postuliert (Narciss, 2018).

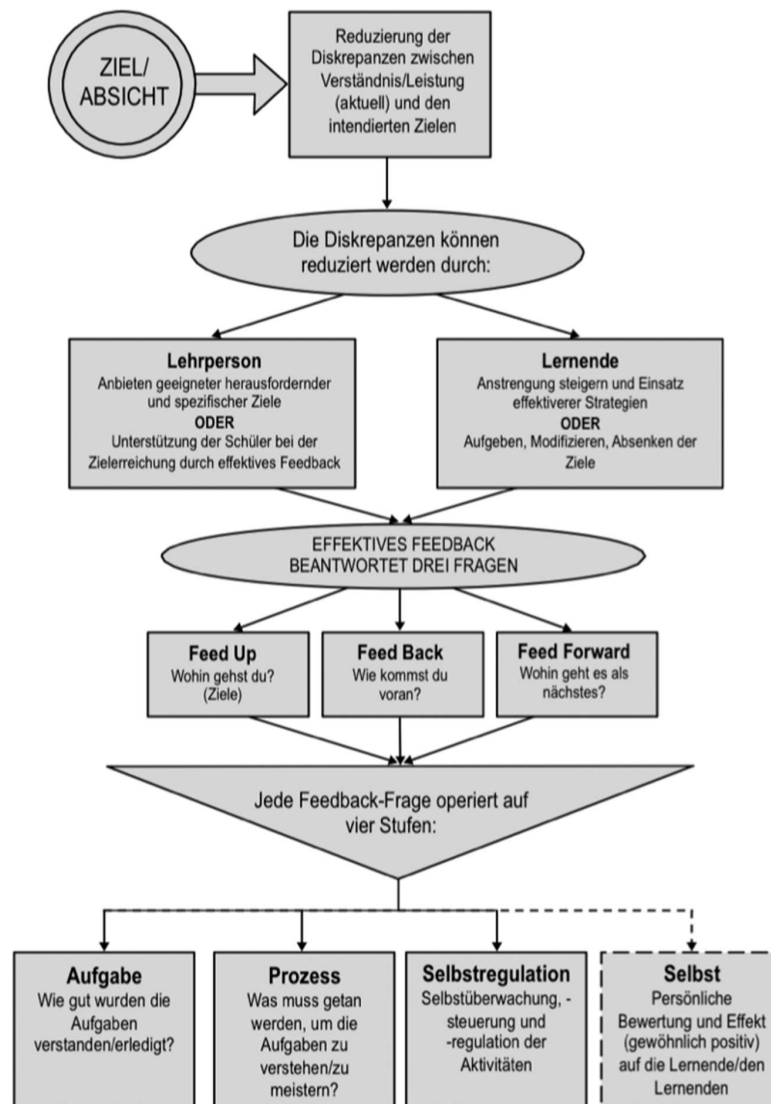


Abbildung 7: Ein Modell über Feedback zur Verbesserung des Lernens (Hattie, 2013)

In Anlehnung an die Modelle von Butler und Winne (1995), Kluger und DeNisi (1996) sowie Hattie und Timperley (2007) und Hattie (2013) entwickelte Narciss (2006, 2008, 2018) für Lehr-Lernsituationen ein eigenes heuristisches Modell (aktualisierte

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Fassung, siehe Abbildung 8) zur Analyse von Feedbackeffekten insbesondere in interaktiven Aufgaben, „das Erkenntnisse aus der Forschung zu Feedback, formativem Assessment, und zum selbstregulierten Lernen mit zentralen Annahmen der Systemtheorie verbindet“ (Narciss, 2018, S. 9).

Die zentrale Funktion von Feedback besteht laut Narciss darin, die lernende Person dahingehend zu unterstützen, Lücken zwischen dem aktuellen und dem angestrebten Wissens- bzw. Kompetenzzustand zu schließen. Um dieser zentralen Funktion von Feedback nachzukommen, verbindet das Modell individuelle und situative Faktoren. Hierzu soll Feedback folglich nicht nur evaluative Funktionen erfüllen, sondern vielmehr wie ein Tutor bzw. eine Tutorin strategisch hilfreiche Informationen zur Überwindung von Hürden im Lernprozess liefern, allerdings ohne direkt die Lösungen anzubieten, um so den Lernenden zur aktiven Wissenskonstruktion anzuregen (Narciss, 2018) und zur Selbstevaluation zu befähigen, wie es auch bei Butler und Winne (1995) angestrebt wird. Narciss (2006) klassifiziert ein so beschaffenes Feedback in ihrem gleichnamigen Buch als *informatives tutorielles Feedback*¹¹ (ITF). Kluger und DeNisi (1996) schreiben so konzipiertem Feedback ebenfalls einen leistungssteigernden Effekt zu.

Das ITFL-Modell (s. Abbildung 8) weist daher deutliche Parallelen zu dem von Butler und Winne (1995) auf, allerdings wird hier ein verstärktes Augenmerk auf die modellierten prozeduralen Abläufe von internen und externen Faktoren sowie die Einflussnahme von externem Feedback auf interne Informationsverarbeitungsprozesse gelegt.

„Die Organisation der zahlreichen Faktoren erfolgt beim ITFL-Modell auf der Basis der Bestimmungslücke zweier interagierender Regelkreise, da auf diese Weise sehr gut deutlich wird, wie situative und individuelle Faktoren zur Regulation eines Lernprozesses mit externem Feedback beitragen“ (Narciss, 2018, S. 11).

¹¹ Informatives tutorielles Feedback macht es sich zum Ziel, regulierend auf Lernprozesse zu wirken und Hilfestellung zu geben, sodass Lernende Wissen(sstrukturen) bzw. Kompetenzen, die sie benötigen erwerben, um die Anforderungen der Lernsituation zu meistern, ohne dabei die Aufgabenlösung preiszugeben. Dazu bedient es sich je nach Zielsetzung (s. Tab. 3) verschiedener elaborierter Feedbackkomponenten (Narciss, 2006, 2014, 2018).

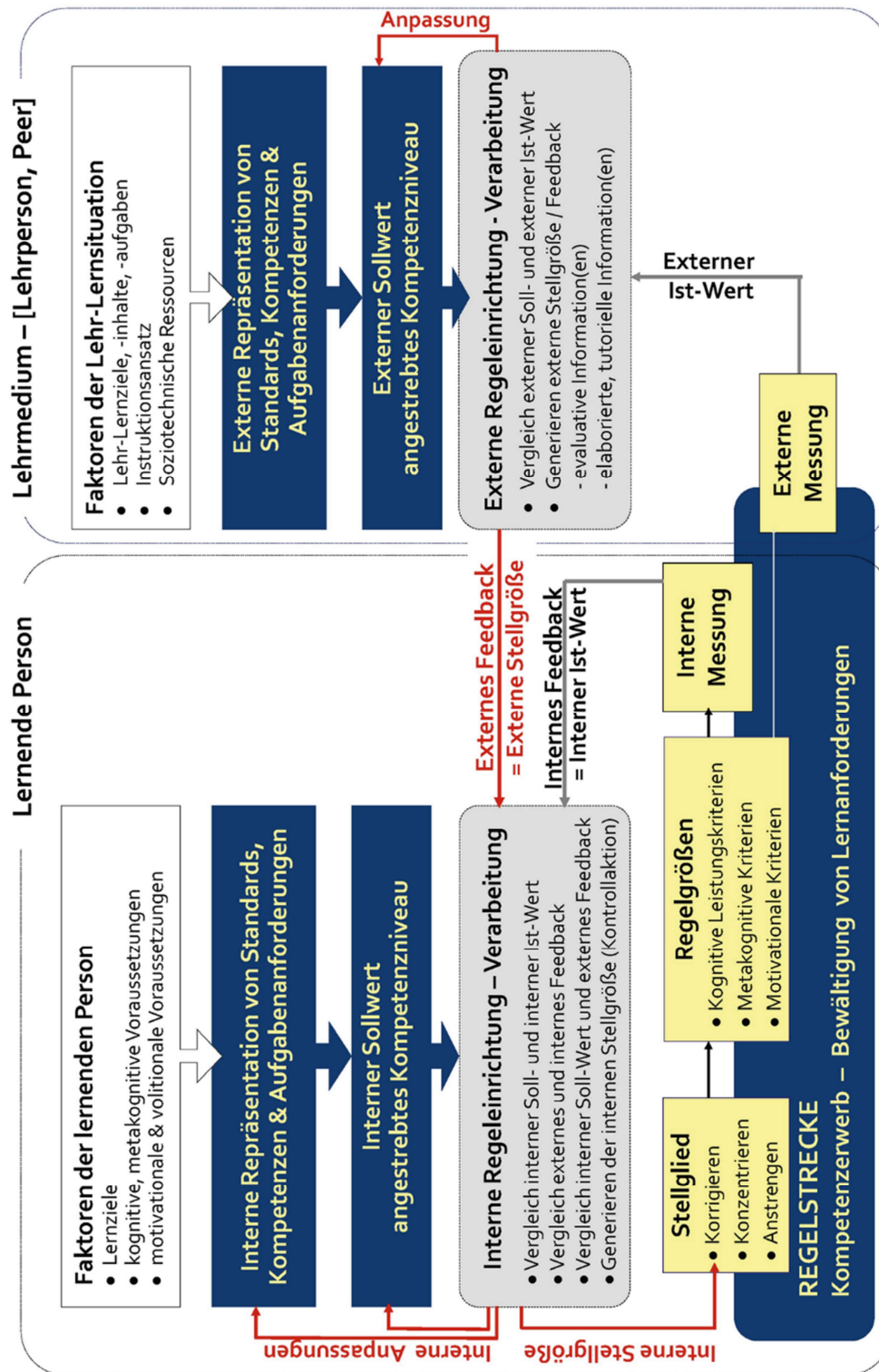


Abbildung 8: Interactive two feedback-loops Model – ITFL-Modell (Narciss, 2018)

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Diese Stelle weist Querbezüge zur Systemtheorie auf. In dem jeweiligen Regelkreis (lernende Person oder Lehrmedium) werden bestimmte Regelgrößen fortlaufend erfasst, sodass sich entsprechend des regelkreisspezifischen Blickwinkels ein *Ist-Wert* ergibt, der innerhalb des Regelkreises dann an die informationsverarbeitende Instanz rückgemeldet wird (*feed back*, s. Abbildung 7).

Liegt eine Diskrepanz zwischen *Ist-* und *Soll-Wert* vor, werden dort, wie auch bei Butler und Winne (1995), Korrekturgrößen oder -maßnahmen spezifiziert (sog. *Stellgrößen*), um eine Veränderung hinsichtlich des Soll-Wertes herbeizuführen. Diese Korrekturen werden dann von einer korrigierenden Instanz (sog. *Stellglied*) im System umgesetzt. Dabei wird ein neuer Ist-Zustand generiert und der Loop beginnt erneut, bis Ist- und Soll-Zustand als kongruent erfasst werden (Narciss, 2018).

„Sowohl für die lernende Person, also den internen Feedback-Loop, als auch den externen Feedback-Loop besteht der zu regulierende Prozess aus dem Erwerb von Wissen und Kompetenzen, die für die Bearbeitung von Lernaufgaben bzw. die Bewältigung der mit Lernaufgaben verknüpften Anforderungen notwendig sind“ (Narciss, 2018, S. 12).

Daraus ergeben sich analog zu Butler und Winne (1995) sowie Kluger und DeNisi (1996) mehrere Anforderungsebenen, die von der Feedback-Funktion angesprochen werden können, denn Feedback kann sowohl auf der kognitiven und motivationalen Ebene als auch auf der metakognitiven Ebene agieren (vgl. Abbildung 9). Da diese Feedback Funktionen aber wiederum eng mit dem Feedback-Inhalt und auch der Feedback-Form (Timing, Modus und Adaptivität) verknüpft sind, die wiederum mit den individuellen Faktoren der lernenden Person wie auch den situativen Faktoren des Lehr-/Lernsettings wechselwirken, vollzieht jedes Individuum die Informationsverarbeitung von Feedback auf eine individuelle Art und Weise. Dieser individuelle Vollzug führt dann zu individuellen (un)mittelbaren Effekten. Um die Wirkmechanismen unter den Determinanten von Feedback verstehen zu können, leitet Narciss (2008) folgendes multidimensionales Modell (vgl. Abbildung 9) von Feedback her.

Dieses Modell liefert zudem eine Grundlage für die trotz Kontrolle des Feedback-Inhalts zu findende Varianz in den Effektstärken von so untersuchtem Feedback (Narciss & Huth, 2004), denn sowohl für die Passung von Feedback und Adressat als auch für die Informationsverarbeitung des Feedbacks sind die individuellen Faktoren der Person ein maßgeblicher Einflussfaktor. Angenommen, alle individualitätsstiftenden Faktoren außer dem Vorwissen würden sich kontrollieren lassen, so würde bereits die Varianz im Vorwissen ausreichen, um bei gleichen Feedback Inhalten zu einer individuellen Verarbeitung des Feedbacks und damit zu einer Streuung im Spektrum der direkten und indirekten Effekte führen.

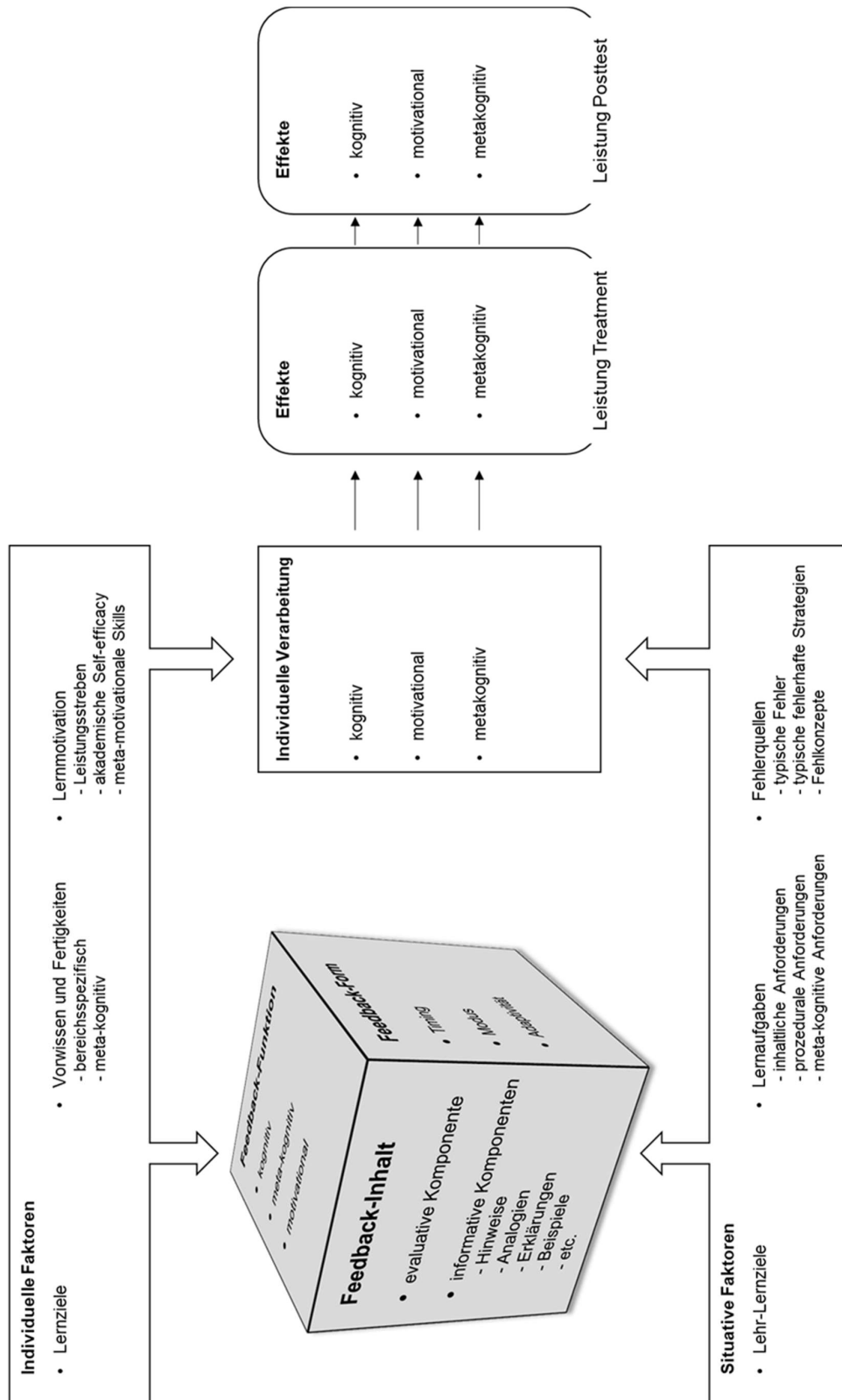


Abbildung 9: Übersicht über Determinanten des Informationswerts von Feedback (nach Narciss, 2008)

Ideales Feedback sollte nach dem Modell von Narciss (2008) daher so gestaltet werden, dass es in der Lage ist, der natürlichen Heterogenität, nicht nur Vorwissen angemessen zu begegnen und bestenfalls jeden Adressaten an seinem Konglomerat an individualitätsstiftenden Faktoren abzuholen – bei genauerer Überlegung wird sich das aber vermutlich als unlösbare Aufgabe darstellen.

2.4 Zum Forschungsstand von Feedback in Lernaufgaben

Bei Lernaufgaben¹² steht allerdings meist der Lernstand und die Veränderung des Vorwissens im Fokus. Denn insbesondere, wenn Feedback eine kognitive Funktion erfüllt, also beim Aufbau von Wissensstrukturen helfen soll, spielt das themenbezogene Vorwissen eine sehr wichtige Rolle. Bestehende Wissensstrukturen müssen aktiviert werden und neues Wissen muss integriert werden (Fritz & Tobinski, 2018; Woolfolk, 2014). *Was soll gekonnt werden? Was kannst du (schon)? Was solltest du als nächstes lernen?* – Das sind die korrespondierenden Fragen zu den von Hattie (2013) formulierten Aspekten *Feed Up*, *Feed Back* und *Feed Forward*. Dass aus der Beantwortung der Fragen bei differentem Vorwissen ein aufgabenspezifisches Feedback mit differentem Umfang erfolgen muss, erscheint hier als logische Folgerung. Für chemienahe Kontexte und die Mathematik konnten auch bereits Indizien für einen solchen Zusammenhang gefunden werden (Albacete & VanLehn, 2000; Moreno, 2004; Narciss & Huth, 2004). Andere Studien liefern hingegen keine gesicherten Erkenntnisse hinsichtlich des Vorwissenseinflusses auf die Wirksamkeit von Feedback unterschiedlichen Umfangs (Fyfe et al., 2012; Narciss, 2006; Smits et al., 2008). Die Einzelstudien zeigen im Vergleich jedoch ähnliche Auffälligkeiten wie in den bereits zitierten Metastudien (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Wisniewski et al., 2019) und lassen daher keine eindeutigen Schlüsse zu.

Als weitere Einschränkung kommt hinzu, dass externes Feedback zusätzlich auf die externe Messung von Zustandsgrößen angewiesen ist, (Hattie & Timperley, 2007; Hattie, 2013; Narciss, 2006, 2014, 2018) und da diese verarbeitet werden müssen, ist eine exakte (Echtzeit-) Passung zwischen Feedback und Individuum unter realen

¹² Lernaufgaben und Testaufgaben müssen in diesem Zusammenhang strikt unterschieden werden, denn Testaufgaben sind darauf ausgerichtet einen Lernstand zu erheben und weniger dazu geeignet, durch sie themenspezifische Handlungskompetenzen zu erlernen. Natürlich fördern auch Testaufgaben bis zu einem gewissen Punkt das Lernen, aber Lernaufgaben sind in diesem Zusammenhang als instruktionales Material zu verstehen, welches dem Lernenden Subjekt dabei hilft, ein Lernziel zu erreichen (Proske, Körndle & Narciss, 2012, S. 1606).

Bedingungen nicht möglich, da eine direkte Betrachtung der Kognitionen (Fritz & Tobinski, 2018) und Berücksichtigung aller Faktoren des Individuums nicht möglich ist.

Dass die ausgegebenen Feedback-Nachrichten kompatibel zum Vorwissen der lernenden Person sein, also verstanden werden können, müssen, um lernwirksam zu werden (Hattie & Timperley, 2007), ist gleichermaßen einleuchtend als auch durch den Aptitude-Treatment-Interaction Ansatz begründet (McLeod, Cronbach & Snow, 1978). Besonders wenn der Wissensstand nicht ausreicht, um Fehler in Lösungen der Übungsaufgaben eigenständig zu finden (Schuhmacher, 2008), können über ITF, z. B. in einer Online-Lernumgebung, effektiv Hilfestellungen gegeben werden (Narciss & Huth, 2004). Vor dem Hintergrund des Expertise-Reversal-Effekts (O. Chen et al., 2017) kann bei vorwissensstarken Studierenden umfangreiches Feedback allerdings auch lernhinderlich wirken (Smits et al., 2008).

Nimmt man also das beobachtbare Verhalten der Lernenden, also die Performanz in einer Aufgabe, als externen Messwert (vgl. Abbildung 8), so kann der externe Ist-Wert (*Was kannst du?*) vom Feedback herangezogen werden, um Defizite aufzuzeigen (*Was soll gekonnt werden?*), und wie das Individuum am besten vorzugehen hat, um sich dem externen Soll-Wert anzunähern, wenn das interne Feedback nicht ausreicht (*Was solltest du als nächstes lernen?*). Für diese Passung von Feedback zum Individuum spielt es daher eine große Rolle, wie oft der Ist-Zustand erhoben wird. Das setzt mindestens eine von Feedback unbeeinflusste Performanz in einer Situation bzw. pro Lernaufgabe, z. B. erster Antwortversuch ohne vorher verfügbares Feedback – Feedback Presearch Availability darf nicht gegeben sein (Bangert-Drowns et al., 1991) – zu einer spezifischen Aufgabenstellung, voraus. Daraufhin ist von entscheidender Bedeutung, wann (Timing), wie oft (Frequenz) und in welchem Umfang (Inhalt) das Feedback eingesetzt wird und welche Funktion (kognitiv, motivational, metakognitiv) es zu erfüllen hat, um zu helfen.

Genau hier liegt die Stärke interpersonaler Lehr-Lernsituationen wie z. B. Tutorien. Geschulte Lehrende sind bestenfalls situativ in der Lage eine Performance Leistung zu analysieren und die Lernenden mit gutem Feedback zu versorgen, was sich unter anderem auch mittelbar in den sehr positiven Ergebnissen, z. B. hinsichtlich akademischer Performance und Haltung, widerspiegelt, die sich bezüglich (schulischer) tutorieller Lernangebote finden lassen (Bloom, 1984; Cohen, Kulik & Kulik, 1982). Verglichen mit universitärer Lehre, beispielsweise Kursräumen, in denen die Lernenden ihre eigenen Ideen durch Kommunikation externalisieren, um so Feedback über ihren Wissensstand und ggf. auch ihre Fehlvorstellung erhalten zu können (Krause, Stark

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

& Mandl, 2009) bzw. im Erfolgsfall auch Kompetenzerleben zu erfahren (Ryan & Deci, 2017), ist es fraglich, ob die tendenziell eigenständig erfolgende Arbeit mit E-Learning-Angeboten eine ähnlich effiziente Arbeit wie ein menschlicher Tutor bzw. Tutorin bezüglich des Feedbacks zu leisten vermag.

Spezifisch für automatisiertes Feedback in Online-Lernumgebungen konnten Cavalcanti et al. (2021) für den gesamten Bereich der Naturwissenschaften eine eklatante Forschungslücke ausmachen: Sie betrachten in ihrem Literatur-Review 3,178 Online-Artikel, die anhand von Schlagwortgruppen über ein automatisiertes Suchverfahren erhalten wurden (u.a. von SpringerLink, Scopus und ScienceDirect). Im anschließende Review-Verfahren bleiben davon 63 Artikel übrig, die sich entsprechend der angelegten Kriterien hinreichend mit dem Thema automatisierten Feedbacks in Online-Lernumgebungen befassen. Mathematik erwies sich dabei mit lediglich drei Artikeln als der einzige naturwissenschaftsnaher Vertreter. Über alle Artikel konnte allerdings unabhängig von der Fachrichtung bei 32 Artikeln ein empirischer Nachweis für positive Effekte von automatisiertem Feedback auf die Performance von Nutzerinnen und Nutzern gefunden werden, wodurch sich tendenziell ein Trend ergibt, den es in chemischen Kontexten zu replizieren gilt. Auf den Einfluss von Vorwissen auf die Wirksamkeit von Feedback wurde in diesem Review nicht eingegangen.

„The lack of social interaction in e-learning might therefore have been a reason for the suboptimal results. Furthermore, example-based learning requires active processing of example information. Probably, many students, especially those with little prior knowledge, did not effectively compare their own solutions with the worked examples; it is likely that they could (or would) not use the (standardized) feedback. Especially, weaker or less motivated students might need specific feedback that refers to individual errors and knowledge gaps. Without social interaction and specific feedback students easily develop misconceptions and illusions of understanding“ (Krause et al., 2009, S. 159).

Die Rechenleistung von Computern hat sich seit der Erfindung des Mikrochips und insbesondere in der letzten Dekade um Dimensionen gesteigert, sodass digitale tutorielle Systeme und ITF immer einfacher und besser zu realisieren sind. Auch die Fehlerevaluation lässt sich nun mit Hilfe der im Software-Tool realisierten tutoriellen Systeme feingliederiger gestalten, sodass insbesondere Studierende mit geringerem Vorwissen, die von Krause et al. (2009) hier angesprochen werden, kognitiv entlastet werden können. Denn bei Fehlern, die einer gewissen Systematik folgen, kann nun auch maßgeschneidertes Feedback ausgegeben werden. Dazu müssen diese Fehler im Software-Tool allerdings definiert und mit spezifischen Feedback-Nachrichten

versehen werden. Eine interaktive Komponente lässt sich ebenfalls realisieren, indem man nach Erhalt des Feedbacks erneut einen Lösungsversuch gewährt. Durch die wiederholte tutorielle Hilfestellung in Interaktion mit der Performance der lernenden Person während der Aufgabenbearbeitung, handelt es sich dann nicht mehr nur um eine einfache Lernaufgabe, sondern um eine interaktive Lernaufgabe¹³.

„Interactive learning tasks reciprocally tutor learners through the series of cognitive operations and actions necessary for solving a task which is specifically designed for active engagement in knowledge construction“ (Proske, Körndle & Narciss, 2012, S. 1606–1607).

Die Limitationen für ein solches Vorhaben liegen insbesondere für komplexe Vorgänge der Chemie mit einem hohen Grad an Visualisierungen und fachspezifischen Schreibweisen (z. B. Reaktionsgleichungen, domänenspezifische Symbolik) immer noch teilweise auf der Seite der Softwarelösung, wie Trauten, Eitemüller und Walpuski (2020) für ausgewählte Beispiele zeigen konnten. Das Spannungsfeld in einer Lernsituation erstreckt sich im Kontext E-Learning also nicht nur zwischen Individuum und Feedback mit all den einflussnehmenden Bedingungen (Vorwissen und Fähigkeiten, (Lern-)Zielsetzungen, kognitive Fähigkeiten, Motivation, Persönlichkeitsmerkmalen, Feedback-Funktion, -Inhalt, -Form), wie die Herleitung von Abbildung 9 aufzeigt (Narciss, 2008), sondern schließt die Software als wechselseitigen Korrespondenten zwischen den beiden Spannungsfeldern als bedingenden Faktor mit ein (siehe Abbildung 10).

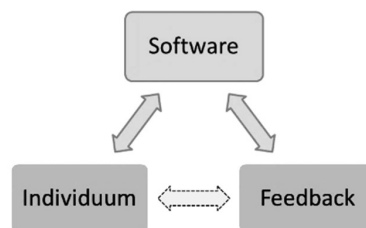


Abbildung 10: Dreigliedriges Spannungsfeld in Feedback präsentierenden Softwarelösungen

¹³ Interaktive Lernaufgaben zeichnen sich dadurch aus, dass die lernende Person bereits während der Aufgabenbearbeitung mit dieser interagieren kann. Dies kann durch (a) Wiederholungs- oder Korrekturmöglichkeiten, (b) tutorielle Hilfestellungen beim Bearbeiten der Aufgabe oder (c) reziproke Reaktionen der Aufgabe auf Aktionen der Lernenden geschehen. Hierdurch kann zu Hürden im Lösungsprozess und Fehlvorstellungen in zugrundeliegenden Konzepten Hilfe bereitgestellt werden (Proske et al., 2012, S. 1606).

Kapitel 2: Über die Bedeutung von Feedback in Lehr-/Lernprozessen

Die Leistungsfähigkeit der Software und das Maß, in der diese ausgeschöpft wird, bedingt die Frequenz und den Umfang der externen Messung von *Regelgrößen* (vgl. Abbildung 8) und damit auch die Präzision, mit der *externe Stellgrößen* formuliert werden können (Narciss, 2018). Narciss und Huth (2004) greifen diese Problematik für die Feedback Konzeption ihrer Lernaufgaben auf – allerdings im Bereich der Grundschullehre in Bezug auf Subtraktionsaufgaben und nicht für komplexe Lernaufgaben im Kontext der Allgemeinen Chemie an Hochschulen. Für ein systematisches Vorgehen bei der Konzeption von Feedback für E-Learning Kontexte soll daher das von Narciss und Huth (2004) entwickelte Schema (vgl. Abbildung 11) dienen¹⁴, welches aber fortlaufend mit den technischen Möglichkeiten der Softwarelösung abgeglichen werden muss.

Angelehnt an dieses Schema und auf Basis der Literatur, insbesondere mit einem Blick auf Abbildung 7, muss bei der Konzeption von Aufgaben für eine Lerngruppe auf die folgenden Fragen eingegangen werden (Hattie, 2013; Jacobs, 1998; Krause et al., 2009; Narciss & Huth, 2004):

1. Welche Voraussetzungen (vor allem Vorwissen) bringen die Lernenden mit?
2. Welche Lernziele müssen erreicht werden?
3. Welche Art, Komplexität und Schwierigkeit an Aufgaben resultieren daraus?
4. Welche typischen Fehler werden bei solchen Aufgaben gemacht?
5. Wie kann man bei solchen Aufgaben effektiv ITF geben?

¹⁴ Narciss und Huth (2004) formulieren als Produkt dieses Auswahl- und Spezifizierungsprozesses sog. Bug-Related-Tutoring Feedback. Es besteht aus einer Auswahl von ITF-Komponenten, die zusätzlich zu einem Hinweis auf einen Fehler (KR) den Fehlerort eingrenzt (KM) und strategische Hilfestellung zum Lösungsprozess bereitstellt. In einem Vergleich ohne klare Gruppendifferenzierung konnte dieses Feedback als vorteilig gegenüber korrektivem Feedback klassifiziert werden.

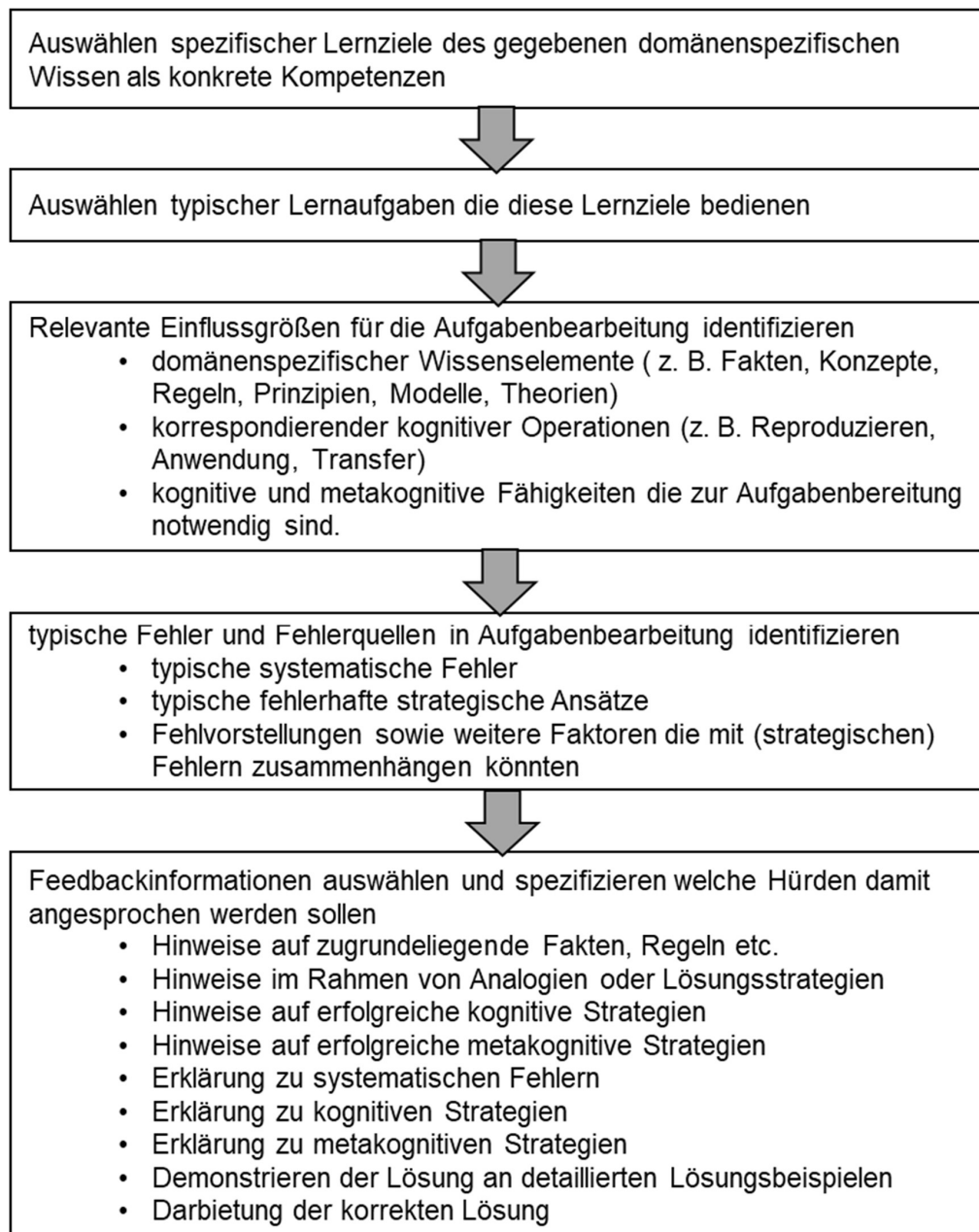


Abbildung 11: Auswahl und Spezifizierung von Feedbackelementen (nach Narciss und Huth, 2004)

Kapitel 3: Gestaltungsprinzipien Multimedialen Lernens

In Kapitel 2.1 stand eher die psychologische Sichtweise des Lernvorgangs im Vordergrund. Die von Sweller, Ayres und Kalyuga (2011) im Rahmen der Cognitive Load Theory (CLT) bzw. die von Mayer (2009) in Cognitive Theory of Multimedia Learning (CTML) gewonnenen Erkenntnisse gelten darüber hinaus als wichtige Grundlagen auf dem Gebiet der Lehr-/ Lernforschung, insbesondere für die Konzeption von Lernmaterial im multimedialen Sektor. Die Vielzahl gestalterischer Möglichkeiten, die sich für digitale Lernangebote ergeben, verleiten Konstrukteure von Lernmaterial immer wieder zu einem übermäßigen Einsatz dieser. Gerade dann, wenn die Konstruktion von (multimedialen) Lernangeboten mit einem großen Aufwand verbunden ist, ist es umso wichtiger, diese so zu konzipieren, dass das Material die Nutzer/-innen beim Lernen unterstützt und diese durch das Material, aber auch die Aufgabenschwierigkeit, nicht überfordert werden. Die beiden oben genannten Theorien liefern hierfür wertvolle Implikationen.

3.1 Cognitive Load Theory

„Cognitive load theory has been designed to provide guidelines intended to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance“ (Sweller, van Merriënboer & Paas, 1998, S. 251).

Wie man weiß, ist die Menge an Informationen, die im Arbeitsgedächtnis verarbeitet werden kann, begrenzt (Hussy, 2018; Klauer & Leutner, 2012; Sweller et al., 1998; Sweller et al., 2011). Diese begrenzte Ressource, vergleichbar mit dem Arbeitsspeicher eines Computers, bildet das Fundament der Cognitive Load Theorie (CLT). Durch Informationsverarbeitungsprozesse wird diese Ressource in einem bestimmten Maße in Anspruch genommen. Folgendes Modell (Abbildung 13) veranschaulicht auf eine vereinfachte Weise die limitierte Informationsverarbeitungskapazität des Arbeitsgedächtnisses anhand von Cognitive Load Komponenten.

Der *Intrinsic Load* wird determiniert durch die Komplexität des zu lernenden Stoffs und ist damit direkt abhängig vom Vorwissen und den Fähigkeiten der lernenden Person. Er ist daher für eine bestimmte lernende Person in einer spezifischen Lernsituation nicht veränderbar. Die CLT geht ebenfalls davon aus, dass Wissen in sog. Schemata organisiert wird. Die Kapazität, die zur Konstruktion bzw. Erweiterung dieser Wissensschemata aufgewandt wird, wird als *Germane Load* bezeichnet.

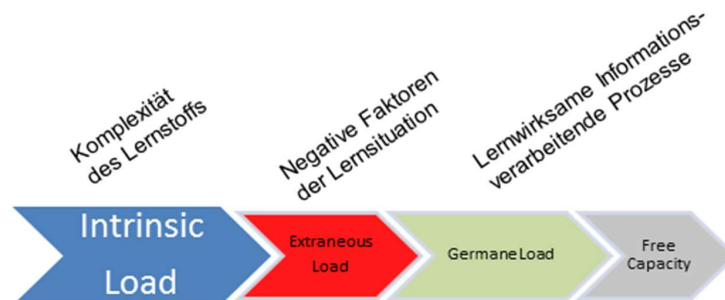


Abbildung 12: Kapazität des Arbeitsgedächtnisses (nach Sweller et al., 1998)

Das Herzstück der CLT bildet jedoch der Extraneous Load oder viel mehr die Kunst diesen zu reduzieren. Neben negativen externen Störfaktoren der Lernsituation (z. B. durch eine unangenehme Lernatmosphäre, Stress u.v.m.) ist ein hoher Extraneous Load häufig auf ein instruktionales Problem zurückzuführen. Oftmals wird der zu lernende Stoff nämlich unnötig verkompliziert.

„Consider another common instructional design. Frequently, instruction includes multiple sources of information such as a combination of mutually referring diagrams and text. In order to understand the diagram or the text, it may be necessary to mentally integrate them“ (Sweller et al., 1998, S. 263).

Wenn die zu lernenden Informationen sehr komplex oder (zudem) verstreut über Text und Bild „verpackt“ sind, dann muss zunächst eine enorme Menge an Kapazität dafür aufgewandt werden, die Informationen in ein mentales Gesamtmodell des Sachverhaltes zu integrieren. Sweller et al. (1998) und Sweller (2010) sprechen hier von der sog. *Element Interactivity*. Beckmann (2010) hält hierzu Folgendes fest: Wenn die Element Interactivity reduziert werden kann, ohne zu verändern, was gelernt werden muss, dann handelt es sich um Extraneous Load. Wenn die Element Interactivity allerdings nur dadurch geändert werden kann, was gelernt wird, dann handelt es sich um *Intrinsic Load*. Ein Lernthema kann also gemessen an einem spezifischen Wissensstand eine hohe interne Komplexität aufweisen, dann ist es für die lernende Person schwer aufgrund der Menge, die gelernt werden muss, oder das Lernmaterial weist einen unnötig hohen instruktionalen Komplexitätsgrad auf, die das Erlernen des Lerngegenstandes (künstlich) erschwert. Als allgemeine Richtlinie lässt sich festhalten:

„Learners' attention must be withdrawn from processes not relevant to learning and directed toward processes that are relevant to learning and, in particular, toward the construction and mindful abstraction of schemas“ (Sweller et al., 1998, S. 264).

Es können jedoch in vorherigen Lerneinheiten Heuristiken, Handlungs- und Wissensschemata – um bei der Computeranalogie zu bleiben: vergleichbar mit Computerprogrammen – erlernt worden sein, die unter bestimmten Bedingungen die Informationsaufnahme und Verarbeitung erleichtern können. Ein Experte weiß z. B. um die Bedeutung einzelner sinnstiftender Elemente in einer Keilstrichformel und kann viel schneller und mehr Informationen aus einer solchen Formel ziehen als jemand, dem diese Darstellungsform noch nicht geläufig ist. Werden allerdings Repräsentationsformen genutzt, die den Lernenden noch nicht mit all ihren Eigenheiten und Konventionen geläufig sind, müssen zusätzlich zum Lerngegenstand noch das Lesen der Repräsentationsform – mit der der Lerninhalt gelernt werden soll – erlernt werden. Dieses sogenannte repräsentationale Dilemma (Rau, 2017) wurde zwar vorwiegend für den Bereich der Organischen Chemie aufgezeigt, doch sollte es auch in der Allgemeinen Chemie berücksichtigt werden. Um diese Problematik besser zu verstehen, soll das integrative Modell der Text und Bildbearbeitung (Abbildung 13) nach Schnotz (2002) herangezogen werden.

Gerade die Informationsverarbeitung via Text-Bild-Integration ist in der Chemie häufig notwendig. Chemie ist für viele ohne mathematische Formeln, Reaktionsgleichungen, Diagramme usw. überhaupt nicht denkbar. „Chemistry is a visual science“ (Wu & Shah, 2004, S. 465). Für Diagramme mag es einleuchtend erscheinen, aber auch bei mathematischen Formeln und Reaktionsgleichungen handelt es sich um Visualisierungen und nicht um Text (Dickmann, 2019). Eine Analyse von Lehrbüchern, die Studierende in der Eingangsphase des Chemiestudiums häufig zum Lernen nutzen, konnte zeigen, dass wenigstens 87 % aller Lehrbuchseiten mindestens eine Visualisierung enthalten. Mit wenigstens 60 % Verteilungsanteil stellen symbolische Repräsentationsformen dabei den Hauptanteil. Im Mittel sind 34 % davon Reaktionsgleichungen (Dickmann, 2019).

Betrachtet man nun die einzelnen Ebenen dieses Modells von Schnotz (2002) zur Text und Bildbearbeitung (siehe Abbildung 13) genauer, so stellt man fest, dass bei der Textanalyse die Analyse der Semantik im Zentrum des Handelns steht und auf Seite der Visualisierung alles auf das Kreieren eines mentalen Modells hinausläuft. Doch sowohl auf der Repräsentations-, als auch auf der Bearbeitungsebene liegt bereits eine wechselseitige Abhängigkeit vor. Deshalb findet man in Lehrbüchern der Chemie häufig Textstellen mit hoher Element Interactivity, wo im Text ein Sachverhalt (z. B. die energetischen Zustände während einer Reaktion) anhand eines Diagramms oder auch mehrerer (z. B. Born-Haber-Kreisprozess) oder mit symbolischen

Kapitel 3: Gestaltungsprinzipien Multimedialen Lernens

Repräsentationsformen (z. B. dem Satz von Hess als mathematische Gleichung) erklärt werden (siehe z. B. Mortimer und Müller, 2003, S. 98). Ist die Darstellungsform und das Lesen dieser erlernt worden, so erzeugt dies weniger Extraneous Load (s. o.) und es kann sich auf den Inhalt des Textes konzentriert werden. Liegt ein repräsentationales Dilemma vor, sind Verständnisprobleme wegen kognitiver Überlastung (overload) nicht auszuschließen.

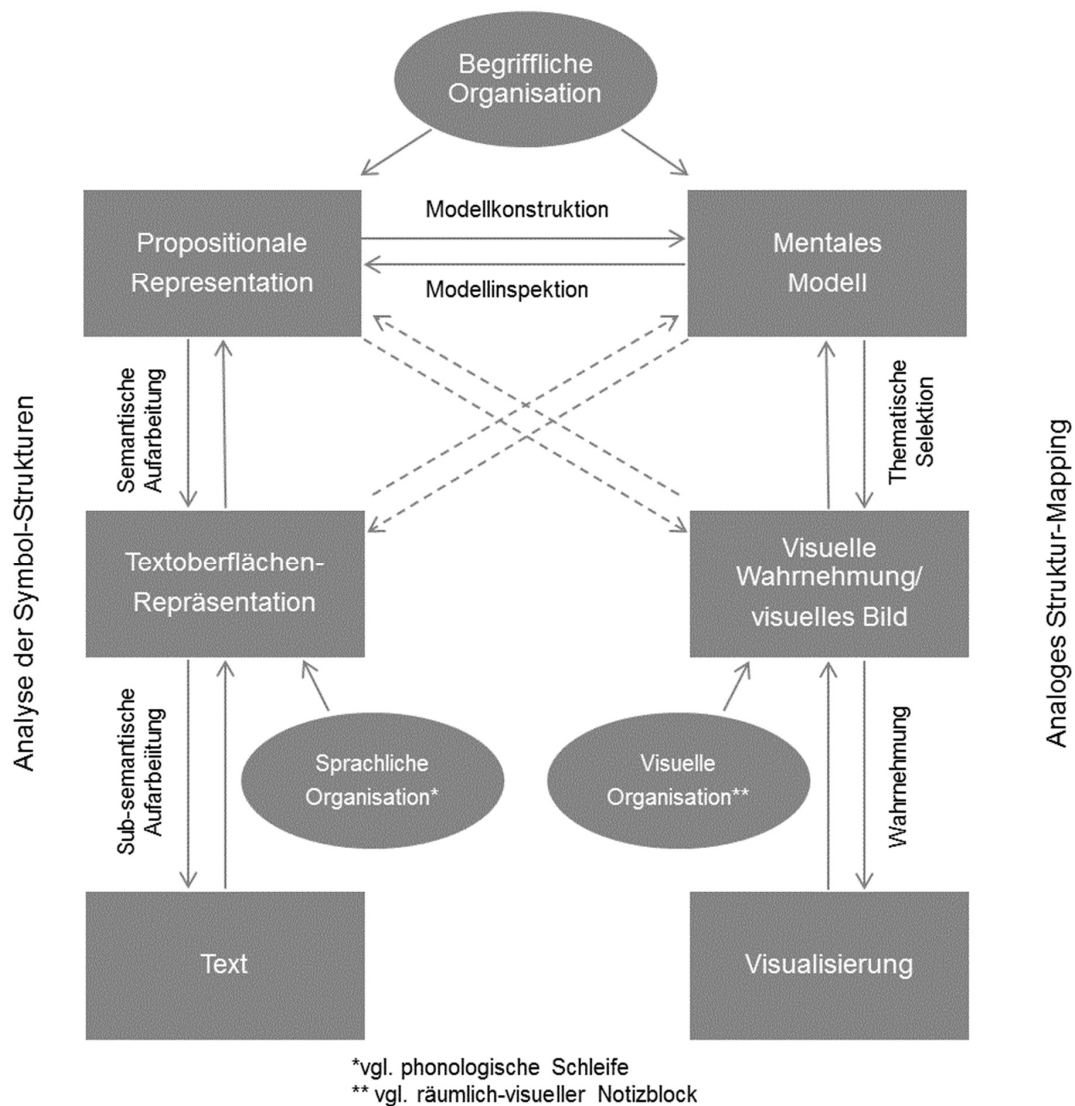


Abbildung 13: Integratives Modell der Text und Bildbearbeitung (nach Schnotz, 2002)

Zusammenfassend zeichnen sich die meisten akademischen Lernsituationen durch multimediale Lernanteile aus, da Studierenden Informationen sehr häufig als Kombination aus Text und Bild z. B. durch Ergänzungen zum Text in Form von z. B. Grafiken bereit gestellt bekommen (Mayer, 2014). Die Chemie nimmt hierbei eine besondere Rolle ein, da es sich aus oben genannten Gründen nicht vermeiden lässt, Informationen auf mehreren Kanälen zu präsentieren (Wu & Shah, 2004). Da diese per se hohe Element Interactivity also dem Lerngegenstand, ja der Disziplin, geschuldet ist, muss nach einem Weg gesucht werden, ein Lernen mit Text und Visualisierungen in Kombination möglichst optimal zu gestalten. Hier liefert die CTML, die sich genau mit diesem Problem beschäftigt, wertvolle Implikationen.

3.2 Cognitive Theory of Multimedia Learning

„Multimedia learning is learning from words and pictures. The rationale for studying multimedia learning is that people can learn more deeply from words and pictures than from words alone. A goal of research on multimedia learning is to understand how to design multimedia learning environments that promote meaningful learning“ (Mayer, 2014, S. 1).

Oft werden in diesem Zusammenhang jedoch die Begriffe Multimedia Learning und Multimedia Instruction synonym verwendet und gelegentlich auch einfach unter dem Begriff Multimedia subsummiert. Folgende Tabelle (Tabelle 4) soll daher dazu dienen, die drei Begriffe auszudifferenzieren:

Tabelle 4: Grundlegende Definitionen der CTML (nach Mayer, 2014, S. 2)

Begriff	Definition
Multimedia	Präsentation von Wörtern (geschriebener oder gesprochener Text) und Bildern (Illustrationen, Fotos, Animationen oder Videos).
Multimedia learning	Aufbau mentaler Repräsentationen auf einer Kombination von Wörtern und Bildern.
Multimedia instruction	Präsentation von Wörtern und Bildern mit der Absicht das Lernen zu fördern.

Multimediale Instruktion kann also multimediales Lernen fördern und beschreibt damit eine potentiell wirkmächtige Gestaltung von Lernmaterial. Allerdings ist die Effizienz einer Lernumgebung, ähnlich wie bei der systematischen Präsentation von Feedback (s. Kapitel 2.3 -2.4), an eine Reihe von Bedingungen geknüpft. Bei der

Kapitel 3: Gestaltungsprinzipien Multimedialen Lernens

Konzeption einer multimedialen Lernumgebung gibt es zwei Ansatzpunkte, von denen man ausgehen kann. Während der technologiezentrierte Ansatz eher der Frage nachgeht, wie man neuste technische Errungenschaften in die Gestaltung von multimedialer Instruktion einfließen lassen kann, geht der auf Lernende zentrierte Ansatz der Frage nach, wie man eine bestehende multimediale Technologie nutzbar machen kann, um das Lernen zu erleichtern. Letzterer Ansatz ist dabei aus didaktischer Sichtweise deutlich zu bevorzugen (Mayer, 2014). Mayer folgt dabei ebenfalls dem kognitivistisch-konstruktivistischen Ansatz als Grundlage für multimediales Lernen (s. o.).

„[A]ccording to the knowledge construction view, the learner [...] is an active sense maker who experiences a multimedia presentation and tries to integrate the presented material into a coherent mental representation [...]. [T]he teacher is a cognitive guide who provides needed guidance to support the learner's cognitive processing. [T]he goal of multimedia presentations is not only to present information, but also to provide guidance for how to process the presented information – that is, for determining what to pay attention to, how to mentally organize it, and how to relate it to prior knowledge” (Mayer, 2014, S. 19).

In diesem Zusammenhang werden von Mayer drei Typen von Learning Outcomes beschrieben (vgl. Tabelle 5), die sich aber durchaus auf das Lernen in Bezug auf weit vernetztes Konzeptwissen gerade im Bereich der Mathematik und der Naturwissenschaften verallgemeinern lässt.

Tabelle 5: Drei Arten multimedialer Learning Outcomes (nach Mayer, 2014, S. 21)

Learning Outcome	Wissensstand	Test-Leistung	
		Behaltensleistung	Transfer
Kein Lernen	Kein Wissen	Schlecht	Schlecht
Routiniertes Lernen	Inselwissen	Gut	Schlecht
Sinnstiftendes Lernen	Integriertes Wissen	Gut	Gut

Die Ergebnisse von PISA und TIMSS, die laut Fischer et al. (2007) auf eine geringe vertikale Vernetzung hindeuten, was sich laut Baumert et al. (2002) negativ auf den Wissenstransfer auswirkt, stärkt somit den von Mayer in Tabelle 5 beschriebenen Typen auf folgenschwere Art und Weise den Rücken. Umso wichtiger ist es daher,

bei der Konzeption von Lernmaterial, das auf die Schließung von Lücken im Vorwissen abzielt, darauf zu achten, dass Text und Bildelemente sinnvoll miteinander verknüpft werden, sodass die lernende Person (1) die Bezüge mit dem bereits vorhandenen Vorwissen aber auch (2) zwischen den Informationsquellen herstellen kann. Während (1) als Aufgabe dem Feedback-Inhalt zuzuordnen ist, wird (2) durch die Darstellung sichergestellt. Mayer liefert hierzu eine Reihe an Leitprinzipien, von denen einige in der folgenden Tabelle (Tabelle 6) dargestellt sind.

Wenn die Darstellung von multimedialen Inhalten eine Mehrzahl der in Tabelle 6 genannten grundlegenden Richtlinien berücksichtigt, denn auch hier sind wieder durch das Spannungsfeld zwischen Feedback-Darstellung und Softwarelösungen (s. Abbildung 10) Einschränkungen möglich, können sogenannte *Advanced Principles of Multimedia Learning* betrachtet werden. Diese Advanced Principles beziehen sich allesamt nicht mehr nur auf das optische Erscheinungsbild, also wie die Informationen präsentiert werden, sondern geben auch konkrete Hinweise hinsichtlich der inhaltlichen Gestaltung. Drei dieser Prinzipien sind besonders wichtig für diese Arbeit, weil sie die in Kapitel 2.4 aufgezeigte Forschungslücke tangieren bzw. übergehen und nun den Fokus wieder auf den Inhalt von Feedback lenken. Diese drei Prinzipien sollen im Folgenden kurz erörtert werden, denn insbesondere für das Prinzip (1) gibt es Diskussionsbedarf.

(1) *Feedback Principle*: Novizen lernen besser durch erklärendes Feedback, das den Grund angibt, warum die gegebene Antwort richtig oder falsch ist, als durch bloßes korrekatives Feedback (Johnson & Priest, 2014; Moreno & Mayer, 2007).

Die in Kapitel 2 herangezogenen Metastudien zur Wirksamkeit von Feedback (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Wisniewski et al., 2019) zeigen deutlich, dass die Faktenlage komplexer zu sein scheint, als das Feedback Prinzip hier suggeriert. Kernelement dieses Prinzips ist der Vergleich der Wirksamkeit von korrektivem (CF; Moreno & Mayer, 2007) und erklärendem Feedback (EF; Moreno & Mayer, 2007). Betrachtet man aber die Feedbackbedingungen einiger diesem Postulat zuarbeitenden Arbeiten unter Zuhilfenahme von Tabelle 3, so erkennt man, dass hier einer weniger differenzierten Auffassung von Feedback gefolgt wird.

Kapitel 3: Gestaltungsprinzipien Multimedialen Lernens

Tabelle 6: Ausgewählte Grundprinzipien multimedialen Designs nach (Mayer, 2009, 2014)

	Prinzip	Erklärung
Reduzieren unnötiger Aufgabenprozesse	1. Coherence Principle	Menschen lernen besser, wenn belanglose Wörter, Bilder oder Geräusche vermieden werden.
	2. Signaling Principle	..., wenn wichtige Begriffe durch zusätzliche Hervorhebungen gekennzeichnet sind.
	3. Redundancy Principle	..., wenn die gleiche Information nicht in mehr als einem Format präsentiert wird.
	4. Spatial Contiguity Principle	..., wenn auf dem Bildschirm eine räumliche Nähe zusammengehöriger Sinnelemente besteht (z. B. Text und Grafik).
	5. Temporal Contiguity Principle	..., wenn zusammengehörige Sinnelemente zeitgleich präsentiert werden.
Aufgabenbearbeitung	6. Segmenting Principle	..., wenn eine multimediale Arbeitseinheit in adressatengerechten Chunks präsentiert wird, statt vollständig und fortlaufend.
	7. Pre-training Principle	..., wenn sie Namen und Randbedingungen des zugrundeliegenden Konzeptes der multimedialen Aufgabeneinheit kennen.
	8. Modality Principle	Menschen können besser mit einer Kombination aus Grafiken und gesprochenem Wort, als von Grafiken und Text auf Bildschirmen lernen.
Unterstützungsmechanismen	9. Multimedia Principle	Menschen können besser mit einer Kombination von Text und Bild als von Text allein lernen.
	10. Personalisation Principle	Menschen lernen besser von multimedialen Arbeitseinheiten, wenn Sätze eher in einer narrativen als in einer formalen Schreibweise präsentiert werden.
	11. Voice Principle	Menschen lernen besser von multimedialen Arbeitseinheiten, wenn gesprochene Textpassagen von einer freundlichen Stimme gesprochen werden als von einer Computerstimme.
	12. Image Principle	Menschen lernen nicht notwendigerweise besser von einer multimedialen Arbeitseinheit, wenn ein Bild des Sprechers präsentiert wird.

„[T]he EF version of the program [...] (a) [...] asked the student to select the plant part that was appropriate for each environment from a presented menu, (b) [...] gave feedback on students' correctness of choice, (c) he explained why the chosen plant part was or was not correct [...] (d) [...] showed students the right choice and moved on to the next step. The CF version had identical functions to the EF version with the exception that it did not include step (c)“ (Moreno, 2004, S. 104).

Problematisch ist hier die nicht trennscharfe Verwendung von Begrifflichkeiten. Dadurch, dass beide Treatments das korrekte Ergebnis präsentieren, handelt es sich in dieser Studie bei „CF“ um KCR bzw. bei dem erklärenden Feedback (EF) um KCR in Verbindung mit KH (siehe Tabelle 3). In folgender Studie (Mayer & Johnson, 2010) wird korrekatives Feedback (CF wird dann als KR bezeichnet) als Treatment-Bedingung mit der richtigen Antwort und einer Erklärung verglichen (EF = KCR & KH).

„After each response, the learner heard a tone — a “ding” for a correct response and a “buzzer” for an incorrect response [...]. In the feedback version [...], after each response, an arrow appeared over the correct answer and a textbox displayed the explanation for the correct answer“ (Mayer & Johnson, 2010, 251 f.).

Zu hinterfragen ist hier, ob der Vorteil des erklärenden Feedbacks nicht auch auf KCR zurückzuführen ist. KCR alleine ist natürlich noch kein erklärendes Feedback. Es präsentiert lediglich die richtige Antwort und indiziert somit erwünschtes Verhalten in einer bestimmten Situation. Bei mehr als zwei Antwortoptionen kann so immerhin *routiniertes Lernen* (siehe Tabelle 5) angestoßen werden. Lediglich sinnstiftendes Lernen, also die Einbettung in Wissensstrukturen durch die Beantwortung der Frage, warum diese Antwort (nicht) richtig ist, bleibt aus. Auf Basis der präsentierten Ergebnisse und vor dem Hintergrund der oben genannten Metastudien zur Zielgerichtetheit menschlichen Handelns (Kluger & DeNisi, 1996) scheint folgende Erklärung konsistenter: Durch erklärendes Feedback (KH) in Kombination mit der Präsentation des korrekten Ergebnisses (KCR) können Novizen besser lernen als durch korrekatives Feedback, da *KCR* vor dem Hintergrund der Zielgerichtetheit menschlichen Handelns eine viel dezidiere Informationsquelle darstellt als *KR*. Zusätzlich kann KH helfen, effizienter Wissen zu konstruieren, da durch Erklärung der Aufgabenschritte in Form einer Musterlösung bestehende Wissensstrukturen aktiviert werden und so die Präsentation des richtigen Ergebnisses mit bestehenden Wissensstrukturen verknüpft werden kann. Hinzu kommt die außerordentliche Themenspezifität und Vorwissensabhängigkeit (Narciss & Huth, 2004; Narciss, 2006) und weitere Faktoren (siehe Abb. 9), die für die Effektivität von Feedback-Komponenten bedeutsam sind. Fraglich ist in dem Zusammenhang auch, ob und in welchem Umfang das Worked Example im Feedback Principle aufgeht. Elaborierte Feedback-Komponenten bedienen sich dieser Lösungsbeispiele sehr häufig.

(2) *Worked Example Principle*: Lernende erlangen ein tieferes Verständnis durch Lernen mit Lösungsbeispielen, wenn sie in der Phase des Fähigkeitserwerbs präsentiert werden (Renkl, 2014b).

Dabei wird das Lösungsbeispiel üblicherweise vor der Bearbeitung von Problemlöseaufgaben präsentiert. Dabei sollen die Lernenden ein Grundverständnis domänenspezifischer Prinzipien und Heuristiken erwerben, um eine Basis für spätere Problemlöseaufgaben zu erlangen (Renkl, 2014a). Das bloße Durcharbeiten von Lösungsbeispielen ist dabei oft mit Motivationseinbußen verbunden. Wenn Lernende aber eine Problemlöseaufgabe mit einem vergleichbaren Problem erwarten, konnten

Sweller und Cooper (1985) zeigen, dass die Motivation, das Lösungsbeispiel durchzuarbeiten, anstieg. Generell zeigen sich für das Lernen mit Lösungsbeispielen, die nicht als Feedback präsentiert werden, bessere Learning-Outcomes (Carroll, 1994; Paas & van Merriënboer, 1994) und auch bessere Transferleistungen (Cooper & Sweller, 1987). Diese drei Forschungsarbeiten bewegen sich zudem alle im Bereich der Mathematik. Die fundamentale Rolle von Mathematik und mathematischen Modellierungsfähigkeiten für das Lernen von Chemie (Goldhausen, 2015) macht diese Befunde daher mittelbar auch äußerst relevant für feedbackgestützte Aufgaben, die unter anderem auch Schulstoff der Oberstufe aufarbeiten sollen. Denn die Allgemeine Chemie bedient sich neben der Grundrechenarten auch der Oberstufenmathematik wie e-Funktionen und Logarithmen, insbesondere in Transferaufgaben, für die mathematisches Modellieren fundamental wichtig ist. Plausibel für diese elaborierte Hilfestellung ist zudem, dass sie in der Phase der Wissensaneignung präsentiert werden sollten, denn zu einem fortgeschrittenen Zeitpunkt des Lernens ist der *Expertise-Reversal-Effect* (O. Chen et al., 2017) zu erwarten, der im dritten Prinzip aufgeht.

(3) *Expertise Reversal Principle*: Mit steigender Expertise in einem Bereich sinkt die Notwendigkeit instruktorischer Hilfestellung (Kalyuga, 2014).

Dem Bedarf für erklärendes Feedback und anschauliche Lösungsbeispiele auf Seite der Novizen steht die zunehmende Bedeutung des Expertise Reversal Effect auf Seite der fortgeschrittenen Lernenden gegenüber. Während für Novizen Hinweise zu relevanten Fakten und Heuristiken für eine Aufgabenlösung notwendig sind und dementsprechend zu großen Effektstärken führen können, wie beispielsweise in der Studie von Moreno (2004, S. 103) mit $d = 1.16$, reicht unter Umständen bei einem Experten z. B. die Präsentation einer Summenformel von Essigsäure im Kontext einer pH-Aufgabe aus, um ein ganzes Netz an Wissen zu aktivieren. All diese Informationen werden dann durch die Summenformel von Essigsäure (re)präsentiert. Die gleichen Informationen noch einmal in einem Feedback (z. B. anhand von einem Lösungsbeispiel) zu präsentieren wird dann redundant (vgl. Tab. 6, *Redundancy Principle*). Durch diese unnötige Menge an zu lesendem Text wird dann kognitiver Aufwand erzeugt (Kalyuga, 2014), der unter Umständen auch durch die Informationen an sich, schädlich wirken kann (O. Chen et al., 2017). Das Zusammenspiel von (1), (2) und (3) führt zudem wieder zu der Schlussfolgerung, die in Kapitel 2.4 schon gezogen wurde. Für standardisiertes Feedback (also ein und dasselbe Feedback zu einem Thema) heißt das entlang des Lernstands von Novize bis Experte, dass die

Effektivität des Feedbacks sinkt oder das Feedback sogar hinderlich sein kann. Folglich muss der Feedback-Inhalt und damit auch der Feedback-Umfang adaptiv zum Vorwissen gestaltet werden.

3.3 Modell zur Entwicklung digitaler Lernumgebungen

Da laut Narciss und Huth (2004, vgl. S. 312) die Gestaltung informativen Feedbacks in multimedialen Lernumgebungen häufiger auf Intuition der Autoren zurückzuführen ist als auf belegte Design-Prinzipien, wurde lange nach einem geeigneten Modell gesucht, das die bislang als relevant herausgearbeiteten Aspekte in sich vereint. Aufgrund der thematischen Nähe zur Mathematik und des hohen mathematischen Anteils in den chemischen Inhalten, die in den Aufgaben Behandlung finden werden, wurde das Modell zur Entwicklung digitaler Lernumgebungen in der Mathematikdidaktik, welches exemplarisch am Beispiel des Bruchzahlbegriffs in Reinhold (2019) dargestellt ist, für diesen Einsatzbereich adaptiert. Dazu wurde die themenspezifisch formulierte Inhaltskomponente des Modells verallgemeinert. Das so erhaltene inhaltlich abstrahierte Modell soll dabei helfen die bereits als relevant ausgearbeiteten Aspekte miteinander zu verbinden.

Wie man anhand des Modells nach Reinhold (2019) sehen kann, greifen die drei Teilbereiche Inhalt, Design und Implementierung ineinander. Die hier im Design verorteten Aspekte finden sich in Kapitel 3.1 und 3.2. Diese rahmen und bedingen über die gestalterische Komponente insbesondere die Präsentation von Informationen und damit die Lernwirksamkeit des Inhalts. Kapitel 2.3 und 2.4 haben aufgezeigt, dass für Feedback ein besonderes Spannungsfeld zwischen Implementierung von Feedback, also den Rahmenbedingungen der Softwarelösung (z. B. naturalistische Eingabemöglichkeiten für präzise Fehlererkennung, Adaptivität der Aufgabe uvm.) und (Feedback-)Inhalt existiert (vgl. Abbildung 10, S. 39). Über die drei Fragen aus dem Modell von Hattie (2013, vgl. Abbildung 7) *Feed Up* (Welche Lernziele wurden schon erreicht?), *Feed Back* (Welche Lernziele sind noch unerreicht? Welche Fehlvorstellungen/Defizite existieren?) und *Feed Forward* (Welches Lernziel kann als nächstes erreicht werden?) wird die zentrale Bedeutung der Inhaltsaspekte für Feedback noch einmal deutlich.

Der einzige Aspekt, der in dieser Arbeit keine Berücksichtigung finden wird, ist der Bereich der Embodied Cognition. Lernumgebungen insbesondere im Bereich des Game-Based Learning und im Rahmen von Virtual bzw. Augmented Reality (Akçayır

& Akçayır, 2017) nutzen immer häufiger Endgeräte mit Touchscreen oder noch weitreichendere technische Möglichkeiten, um mehrere Teile des Körpers oder gar den ganzen Körper in die Lernerfahrung mit einzubeziehen (Reinhold, 2019).

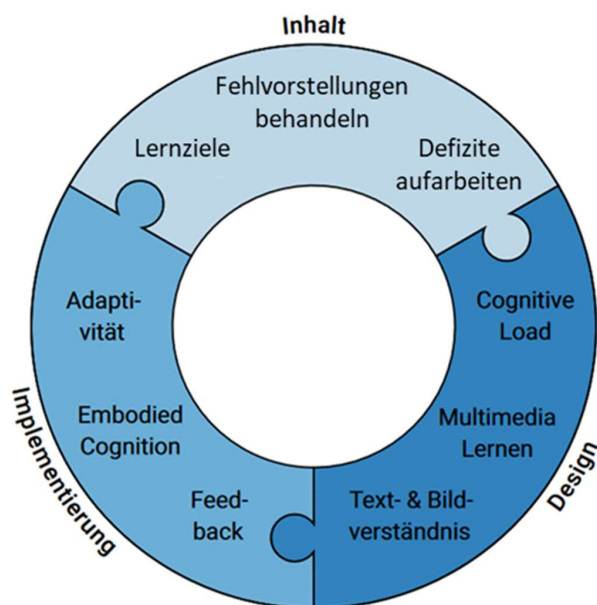


Abbildung 14: Modell zur Entwicklung digitaler Lernumgebungen mit verallgemeinerter Inhaltskomponente (nach Reinhold, 2019)

Die Embodied Cognition Theorie stellt dabei den psychologischen Rahmen für den Fall, dass konkreten körperlichen Erfahrungen eine zentrale Rolle im Lernprozess zukommt (Wilson, 2002). In diesem, in Bezug auf Lernumgebungen allgemeinen Modell (Abbildung 14) wird sie daher aufgeführt, wird aber für diese Arbeit nicht ausreichend relevant sein, da ausschließlich herkömmliche Eingabemethoden (Maus, Tastatur) Verwendung finden werden.

Das nun in allen relevanten Aspekten mit der Theorie verknüpfte Modell soll für die Aufgabenentwicklung als Orientierungshilfe dienen.

Denn „[d]iese drei Dimensionen [Inhalt, Design und Implementierung] sind explizit nicht losgelöst voneinander zu verstehen. Vielmehr sollen bei konkreten Entscheidungen bezüglich der Entwicklung digitaler Lernumgebung Aspekte unterschiedlicher Dimensionen gleichermaßen betrachtet werden“ (Reinhold, 2019, S. 167).

Kapitel 4: Forschungsfragen, Ziele und Hypothesen

Aus dem Zusammenspiel der beiden Problembereiche (Performance-Problem und Feedback-Problem) in Kapitel 1 erwächst unmittelbar ein Handlungsbedarf für die Hochschullehre in MINT-Fächern und damit auch im Fach Chemie, um die in Zukunft notwendigen Fachkräfte in den innovierenden Berufen bereitstellen zu können (Brossardt, 2015). Dass das erste Semester in vielfacher Hinsicht eine bedeutsame wie herausfordernde Phase im Studium darstellt und damit entscheidend für den Verbleib im Studium sein kann (Averbeck, 2020; Heublein et al., 2017), konnte ebenfalls im ersten Kapitel gezeigt werden. Die Abbruchquoten zeigen zudem, dass viele Studierende den Herausforderungen dieser Phase, die geprägt ist von Integration in neue Sozialstrukturen (Saraceni & Müller, 2011; Tinto, 1988; Trautwein & Bosse, 2017) und der Notwendigkeit, Lücken im Vorwissen schnell auszumachen und aufzuholen (Heublein et al., 2017), nicht gewachsen sind.

Daher ist das Ziel dieser Arbeit, eine individuelle Förderungsmaßnahme zu entwickeln, die das Vorwissen der Studierenden konzeptionell stärker berücksichtigt und so insbesondere Studierenden mit geringem Vorwissen die Möglichkeit geben soll, ihr Potential ausschöpfen zu können, gute Chemiker zu werden. Mithilfe von automatisiertem Feedback in Lernaufgaben, integriert in ein digitales Tutorium zur Allgemeinen Chemie Vorlesung, soll in dieser Phase begrenzter Feedbackgelegenheiten (Heublein et al., 2017; Tinto, 1988) eine personenunabhängige Feedbackquelle bereitgestellt werden, um Vorwissensdefizite besser aufzeigen und behandeln zu können.

Wie in Kapitel 2 und insbesondere 2.4 gezeigt werden konnte, ist die lernwirksame Gestaltung von Feedback keineswegs trivial. Zwar sind sich Forschende einig, dass Feedback kompatibel zum Vorwissen sein muss (Hattie & Timperley, 2007), doch hinsichtlich der vorwissensbedingten Notwendigkeit von spezifischen Feedbackkomponenten (siehe

Tabelle 3, S. 28) und deren Umfang gibt es bislang keine eindeutige Befundlage. Die Entwicklung eines solchen hinsichtlich des Vorwissens differenzierenden Hilfsangebots geht daher mit der Evaluation verschiedener Feedback-Komponenten einher. Daher liefert diese Arbeit einen wertvollen Beitrag zum Schließen der konstatierten Forschungslücke. So wurde im Rahmen dieser Arbeit der Vorwissenseinfluss auf die lernförderliche Wirkung zweier ausgewählter Feedback-Maßnahmen untersucht, die sich vom Umfang und hinsichtlich der Feedback-Komponenten her stark unterscheiden.

Kapitel 4: Forschungsfragen, Ziele und Hypothesen

Denn insbesondere, wenn Feedback eine kognitive Funktion erfüllt, also beim Aufbau von Wissensstrukturen helfen soll, spielt das themenbezogene Vorwissen eine sehr wichtige Rolle¹⁵. In einer Studie mit experimentellem Design soll daher der Vorwissenseinfluss auf die unterschiedliche Lernwirksamkeit zweier Feedback-Maßnahmen untersucht werden. Mit Blick auf die Zielsetzung der Arbeit ergibt sich die folgende Forschungsfrage.

FF1: Wie wirken sich die beiden Feedback-Maßnahmen (Error Specific Tutoring-Feedback vs. KORrektives-Feedback) in Abhängigkeit vom fachspezifischen Vorwissen auf den Studienerfolg von Erstsemesterstudierenden in der Allgemeinen Chemie aus?

Als Maß für den Studienerfolg wird in dieser Arbeit der Fachwissenserwerb herangezogen, operationalisiert durch einen unabhängigen Fachwissenstest, sowie den Klausurerfolg am Ende des Semesters. Zur Forschungsfrage ergeben sich theoriegeleitet folgende Hypothesen: Error-Specific-Tutoring (EST) Feedback, das fehlerspezifische Hilfestellung anbietet zeichnet sich nach Narciss, (2006), wie auch beim Bug-Related-Tutoring-Feedback beschrieben (vgl. Kapitel 2.4), durch verschiedene ITF Komponenten aus. Dies erscheint besonders vielversprechend für vorwissenschwache Studierende zu sein, die beim Auftreten von Fehlern elaboriertes Feedback benötigen (Krause et al., 2009; Mayer & Johnson, 2010; Moreno, 2004) insbesondere, da das Finden von Rechen- und Denkfehlern im Lösungsprozess ein gewisses Maß an Vorwissen voraussetzt (Schuhmacher, 2008). Somit erscheint ein EST-Feedback-Algorithmus als vielversprechender Ansatz, das Konzeptverständnis in der Allgemeinen Chemie zu fördern, vor allem vor dem Hintergrund der lernförderlichen Wirkung mehrstufiger fehlerspezifischer ITF-Algorithmen in Studien zum Kompetenzerwerb im Bereich Physik und Mathematik (z. B. Albacete & VanLehn, 2000; Narciss & Huth, 2004). Daraus ergibt sich folgende erste Hypothese:

H1: Studierende mit geringem fachspezifischem Vorwissen profitieren in ihrem Studienerfolg stärker durch ein Training mit einem EST-Feedback als Studierende mit hohem Vorwissen.

¹⁵ Allein wenn man Lernprozesse anhand eines sehr basalen Gedächtnismodells betrachtet – vornehmlich das Aktivieren von Wissensstrukturen und die Integration von Wissen darin (siehe Abbildung 4, S. 16) –, ist der Einfluss von Vorwissen nicht von der Hand zu weisen.

Bei zunehmender Expertise sinkt allerdings der Nutzen von elaboriertem Feedback. Es kann sogar hinderlich werden (Expertise Reversal Effect, siehe Kapitel 3.2; Kalyuga, 2014), sodass für Experten auch Nachteile durch zu umfangreiches Feedback entstehen können (z. B. Smits et al., 2008).

H2: Studierende mit hohem fachspezifischem Vorwissen profitieren in ihrem Fachwissenserwerb bzw. Studienerfolg stärker durch ein Training mit einem KOR-Feedback als Studierende mit geringem Vorwissen.

Kapitel 5: Studiendesign und Testinstrumente

Entscheidend ist laut Heublein et al. (2017), dass die Aufarbeitung von Vorwissensdefiziten besonders früh im Studium erfolgt. Für den Vergleich der zwei Feedback-Maßnahmen bilden Chemiestudierende im ersten Semester die geeignete Grundgesamtheit. Als Lehrveranstaltung wurde das Modul „Allgemeine Chemie“ gewählt, da es an vielen Universitäten Deutschlands im ersten Semester gelehrt wird. Da ein experimentelles Design mit einer wenigstens deutschlandweit randomisierten Zufallsstichprobe mindestens aus logistischen Gründen zum Scheitern verurteilt wäre, wurde ein quasi-experimentelles Design mit Erstsemesterstudierenden der Universität Duisburg-Essen als Teilstichprobe gewählt. Das hat den Vorteil, dass über einen wechselseitigen Austausch ein sehr detaillierter Einblick in die Vorlesungsinhalte sichergestellt werden konnte, um die Passung der späteren Lernaufgaben mit den Lerninhalten zu erhöhen. So wurde das Online-Tutorium auch seitens des Dozenten als fester Bestandteil des Moduls präsentiert, was sicherlich einen Einfluss auf die Arbeitsbereitschaft der Studierenden gehabt haben wird. Allerdings ergeben sich aus dieser Teilstichprobe auch Einschränkungen hinsichtlich der Verallgemeinerung der Ergebnisse auf die Grundgesamtheit, da ein Effekt des Dozierenden nicht kontrolliert werden konnte. Für Replikationsstudien an anderen Universitäten könnte das Maß an Commitment und die spezifische Art, die Lehre zu gestalten, ein Bias der Studie bedeuten. Der Erhebungszeitraum fällt damit auf das erste Fach- und Wintersemester im Studium. Die Einteilung der Studierenden in die Übungsgruppen muss im universitären Lehrbetrieb innerhalb der ersten Semesterwoche erfolgen. In der Kürze der Zeit ist eine Auswertung von Fachwissensdaten aber kaum möglich, sodass andere Indikatoren für das Vorwissen gefunden werden müssen. Daher werden die Studierenden über die Kurswahl in der Oberstufe (Chemie Leistungskurs: ja/nein), die als guter Prädiktor für das Vorwissen gilt (Averbeck, 2020), gleichmäßig zwischen den Interventionsbedingungen aufgeteilt und randomisiert zugewiesen, sodass sich ein 2x2 Design aus Kurswahl und Feedbackbedingung ergibt (IG1 = kein LK+EST, IG2 = kein LK+KOR, IG3 = LK+EST, IG4 = LK+KOR) (vgl. Tabelle 7).

Tabelle 7: Interventionsgruppen im 2x2 Design

Kurswahl Chemie	Feedback-Typ	
	EST	KOR
Kein Leistungskurs	EST _(kein LK)	KOR _(kein LK)
Leistungskurs	EST _(LK)	KOR _(LK)

Zusätzlich ergab sich für weitere Analysen die Möglichkeit durch die in der ersten Förderphase von ALSTER durchgeführten Erhebungen auf die Pre- und Post-Daten

einer Kontrollkohorte (WiSe 16/17) zurückzugreifen. In der ersten Förderphase (ALSTER; FOR 2242) wurde mittels einer Längsschnittstudie mit drei Messzeitpunkten (Anfang des ersten, Ende des ersten und Ende des zweiten Fachsemesters) über die ersten beiden Studiensemester Einflussfaktoren auf den Studienerfolg untersucht. In dieser Studie wurden folglich viele kognitive, aber auch affektiv-motivationale Variablen erhoben, es wurde allerdings keine Interventionsstudie durchgeführt (Averbeck, 2020). Auch die Lehre hat sich seitdem nicht verändert, was den günstigen Umstand mit sich bringt, dass diese Kohorte mithilfe einer herkömmlichen Übung bei aber sonst gleichen Modulinhalten gelernt hat und zugleich auf viele erhobene Variablen zurückgegriffen werden kann, die diese Kohorte sehr detailliert beschreiben. Die Erfassung des Fachwissens der Allgemeinen Chemie bildet im Rahmen dieses Pre-Post-Vergleichsgruppendesigns das wichtigste Testinstrument. Hier wurde der Fachwissenstest der Allgemeinen Chemie (Averbeck, 2020) herangezogen, der sich im Rahmen der ersten Förderphase als Weiterentwicklung des Fachwissenstestes von Freyer (2013) als valides Messinstrument erwiesen hat. Unter Berücksichtigung der schwierigkeitserzeugenden Merkmale aus dem Projekt ESNaS (Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I, Walpuski et al., 2010) ist der Test während der ersten Förderphase „hinsichtlich der Komplexität bewertet und im Anschluss [durch] weitere [Items] zu den unterrepräsentierten Stufen“ ergänzt worden (Averbeck, 2020, S. 81). Er liefert eine akzeptable bis gute EAP/PV-Reliabilität von .765 zum ersten und von .802 zum zweiten Messzeitpunkt. Die Itemreliabilität ist mit .75 für den ersten Messzeitpunkt bzw. mit .78 zum zweiten Messzeitpunkt zufriedenstellend und ermöglicht daher eine relativ präzise Schätzung der Personenfähigkeiten hinsichtlich des Konstruktes „Fachwissen der Allgemeinen Chemie“, auf das der Test abzielt. Der Item-Fit aller 35 Items (Range: 0.76–1.275) liegt zudem innerhalb der von Bond, Yan und Heene empfohlenen Spanne für gute Infit- und Outfit-Parameter für Item Mittelwertsquadrate (Range: 0.7–1.3; 2020, S. 244). Die verschiedenen Themenbereiche des Fachwissenstests¹⁶ beruhen auf einer vorangegangenen Analyse des Modulhandbuchs sowie empfohlener Begleitlektüre und Material der Vorlesung zur Allgemeinen Chemie an der Universität Duisburg-Essen und wurden auch an anderen Standorten, z. B. Ruhr-Universität Bochum, validiert. Der Test wird im Paper-Pencil Format durchgeführt und umfasst 35 Multiple-Choice Aufgaben im Single Select Format. Da laut Narciss (2006) der Effekt

¹⁶ Stoffe und Elemente, Stöchiometrie, Atomaufbau und Periodensystem, Trends im Periodensystem, chemische Bindungen: kovalente, ionische, metallische Bindungen, Moleküle und Orbitale, chemische Energetik, chemische Kinetik, chemisches Gleichgewicht, Säuren und Basen, Oxidation und Reduktion, Komplexbildung

von Feedback und damit der des Treatments maßgeblich von individuellen Personenmerkmalen anhängig ist, wurden neben allgemeinen demografischen Daten entlang des Semesters einmalig zusätzlich weitere leistungsbezogene und motivationale Personenmerkmale mithilfe von weiteren bereits validierten Testinstrumenten erhoben (siehe Tabelle 8).

Tabelle 8: Testinstrumente zur Erfassung von Kontrollvariablen

Skala	Quelle	N _{Items}	EAP/PV	Cronbachs α	Format
Kognitive Fähigkeiten	Heller und Perleth (2000)	25	.77		Paper-Pencil
Mathematische Fähigkeiten	Müller et al. (2018)	23	.86		Paper-Pencil
Studienmotivation (Erwartung, Wert, Kosten)	Kosovich, Hul- leman, Barron und Getty (2015)	3, 3, 3		.84, .82, .71	Onlineum- frage
Allgemeine Lern- und Leistungsmoti- vation im Studium	Spinath, Stiens- meier-Pelster, Schöne und Dick- häuser (2002)	8		.78	Onlineum- frage
Fachbezogenes Akademisches Selbstkonzept	Schöne, Dickhäu- ser, Spinath und Stiensmeier-Pelster (2002)	5		.85	Onlineum- frage
Lernstrategien im Studium	Wild und Schiefele (1994)	41		.88	Onlineum- frage

Die verwendeten Paper-Pencil basierten Testinstrumente wurden in zwei Versionen mit gegenläufiger Item-Reihenfolge ausgeteilt. Der detaillierte Ablauf der Pilotierung wie auch der Hauptstudie wird in der folgenden Tabelle (Tabelle 9) zusammenfassend dargestellt. Da die Studierenden bereits in der ersten Semesterwoche innerhalb einer sehr kurzen Zeitspanne in die Übungsgruppen eingeteilt werden mussten, war im Vorfeld bereits klar, dass die harten Fachwissensdaten (Pre) nicht für eine Einteilung in die Gruppen herangezogen werden können.

Kapitel 5: Studiendesign und Testinstrumente

Tabelle 9: Studiendesign für Pilotierung und Hauptstudie 1, Semesterplan

Semesterverlauf	IG 1 + IG 3 (EST-Feedback)	IG 2 + IG 4 (KOR-Feedback)
Erste Semesterwoche	Demografie-Fragebogen Allgemeine Chemie (Pre) Kognitiver Fähigkeitentest Test für mathematische Fähigkeiten Tutorial für den Umgang mit JACK Studienmotivation (Erwartung, Wert, Kosten) Allgemeine Lern- und Leistungsmotivation im Studium	
Semesterbegleitende Intervention	Arbeiten mit den Lernaufgaben empfundene Aufgabenschwierigkeit/ Denkanstrengung/Nützlichkeit des Feedbacks/Nutzerfreundlichkeit der Aufgaben, time-on-task	
Letzte Semesterwoche	Allgemeine Chemie (Post)	
Vorlesungsfreie Zeit	Klausur/ Nachschreibklausur	

Daher wurde, wie bereits erwähnt, die Chemie Kurswahl in der Oberstufe, die in der ersten Vorlesungssitzung im Rahmen des Demografie-Fragebogens erhoben wurde, herangezogen, um in der Kürze der verfügbaren Zeit die Studierenden auf zwei vom Vorwissen her vergleichbare Interventionsgruppen zu verteilen. Da aufgrund des Schwundes an Probanden damit zu rechnen war, dass die Gesamtstichprobengröße nicht ausreichend für die anschließende Auswertung sein könnte, wurde für das Folgejahr die Kohorte durch einen zweiten Durchlauf der Hauptstudie erweitert. Für die Erweiterungsstudie im WiSe 2020/2021 (HS 2) wurden allerdings nicht alle Testinstrumente benutzt. Da es sich hierbei um eine vollständig Online durchgeführte Erhebung während der COVID-19 Pandemie handelte, wurde auf einige Fragebögen verzichtet, um die Studierenden nicht durch zu viele Erhebungen zu demotivieren. In dieser Kohorte wurden lediglich der Demografie-Fragebogen sowie kognitive Fähigkeiten (figurale Skala) zum Pre-Messzeitpunkt sowie das Fachwissen in der Allgemeinen Chemie am Pre- und Post-Termin erhoben. Über die Messzeitpunkte ergibt sich folgende Verteilung der Messinstrumente (siehe Tabelle 10: Eingesetzte Instrumente je Kohorte). Die Lernaufgaben selbst, deren Entwicklung sich das folgende Kapitel widmen wird, wurden den Studierenden über einen je Interventionsgruppe spezifischen Kursraum mithilfe des Lern Management Systems (LMS) Moodle präsentiert. Das hat sowohl den Vorteil, dass die Lernenden sich sehr schnell in den Kursräumen zurechtfinden, da diese Software-Oberfläche von der Mehrzahl aller belegbaren Kurse der Universität genutzt wird, als auch, dass die so erzeugten Daten

im hauseigenen Serverstandort der Universität Duisburg-Essen verbleiben, was datenschutzrechtlich eine sehr günstige Ausgangslage darstellt. Mithilfe der Gestaltungsmöglichkeiten, die Moodle bereit hält, lässt sich zudem mit wenig Aufwand eine Kursoberfläche mit übersichtlicher Struktur von Dokumenten und Tools erzielen.

Tabelle 10: Eingesetzte Instrumente je Kohorte

Skala	Pilotierung	HS1	HS2
Fachwissen Allg. Chemie (Pre u. Post)	✓	✓	✓
Kognitive Fähigkeiten	✓	✓	✓
Mathematische Fähigkeiten	✓	✓	x
Studienmotivation (Erwartung, Wert, Kosten)	✓	✓	x
Allgemeine Lern- und Leistungsmotivation im Studium	✓	✓	x
Fachbezogenes Akademisches Selbstkonzept	✓	✓	x
Lernstrategien im Studium	✓	✓	x

Um vor der Bearbeitung der Lernaufgaben den Umgang mit den softwarespezifischen Eingabetools zu trainieren, wurde ein kurzes Tutorial, das sechs beliebig oft wiederholbare Tutorial-Aufgaben umfasst, vorgeschaltet, bei dem die Lernenden Rückfragen stellen konnten, so sich denn Schwierigkeiten ergaben. Nach der Bearbeitung einzelner Themenfelder wurde mit je einer Frage innerhalb einer vierstufigen Likert-Skala eine aufgabentypbezogene Einschätzung der Studierenden erhoben. Diese Zwischenerhebungen sind bewusst sehr kurz gehalten worden. Des Weiteren sollten die Systemdaten Rückschlüsse auf die Zeit der Aufgabenbearbeitung ermöglichen.

Fragen der Zwischenerhebung:

1. Wie schwierig waren die Aufgaben, die Sie gerade bearbeitet haben, zu lösen?
☐ sehr einfach ☐ einfach ☐ schwierig ☐ sehr schwierig
2. Bei der Bearbeitung der Aufgaben war meine Denkanstrengung insgesamt ...
☐ sehr gering ☐ eher gering ☐ eher hoch ☐ sehr hoch
3. Wie hilfreich war das Feedback bei der Lösung der Aufgaben?
☐ sehr hilfreich ☐ eher hilfreich ☐ eher nicht hilfreich ☐ nicht hilfreich
4. Wie intuitiv haben Sie die Eingabe bei den Aufgaben empfunden?
☐ sehr intuitiv ☐ eher intuitiv ☐ eher nicht intuitiv ☐ nicht intuitiv

Als motivationale Komponente wurde nach Rücksprache mit der Lehrperson ein Bonuspunktesystem etabliert, das den Studierenden bei hinreichender Bearbeitung von Themenblöcken inkl. Zwischenerhebungen im Falle einer bestandenen Klausur

Kapitel 5: Studiendesign und Testinstrumente

durch den Erhalt einiger Bonuspunkte eine Verbesserung der Modulnote ermöglichen konnte (Bonuspunktregelung, siehe Anhang 2).

Kapitel 6: Aufgabenentwicklung

Die Auswahl spezifischer Lernziele des gegebenen domänenspezifischen Wissens als anzustrebende Kompetenz sowie typischer Aufgaben, die diese kompetenzbildenden Lernziele bedienen, stellen die ersten beiden Schritte des von Narciss und Huth (2004) beschriebenen Schemas zur Auswahl und Spezifizierung von Feedbackelementen dar (siehe Abbildung 11, S. 41). Dem liegt die Analyse von Übungsaufgaben, Klausuraufgaben und weiteren Modulinhalten nach Lernzielen zugrunde.

6.1 Lernzielorientierte Aufgabenkonstruktion

Diese Lernziele liefern bereits über die Operatoren die nötigen Anforderungen an das Software-Tool. Parallel dazu müssen typische Fehlerquellen ermittelt und bei der Konzeption von Feedback berücksichtigt werden, wodurch sich weitere Anforderungen an das Software-Tool, durch die dafür notwendige automatisierte Evaluation späterer Antworten ergeben. Dieser Prozess ist in Abbildung 15 verdeutlicht.

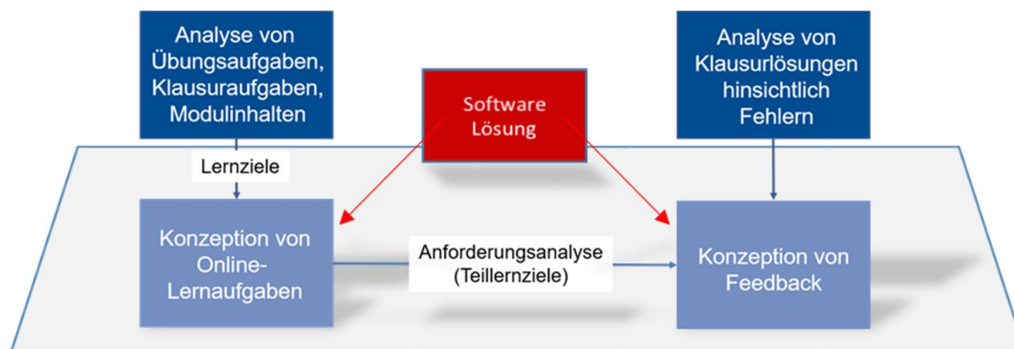


Abbildung 15: Überführung der Papier-Bleistift-Aufgaben in digitale Lernaufgaben unter Berücksichtigung der Einschränkungen durch die Software Lösung

Zur Verdeutlichung der Bandbreite an Operatoren für das Fach Chemie als Naturwissenschaft und dem daraus erwachsenden Anspruch an das Tool seitens der Aufgabenstellung, soll die Formulierungshilfe für Lernziele im universitären Rahmen, die von Cursio und Jahn (2015) für die naturwissenschaftliche Fakultät der Universität Nürnberg-Erlangen erstellt wurde, benutzt werden (siehe Tabelle 11). Darin finden sich zudem auch die von IQB und KMK (2020) veröffentlichten einheitlichen Operatoren für Aufgaben für die Fächer Biologie, Chemie und Physik wieder. Hieraus und aus der besonderen Rolle, die Visualisierungen (Dickmann, 2019; Rau, 2017; Wu & Shah, 2004) und Mathematik (Goldhausen, 2015; Kimpel, 2017) in der (Allgemeinen)

Kapitel 6: Aufgabenentwicklung

Chemie einnehmen, erwächst ein direkter Anspruch an die Lernaufgaben, sowie in zweiter Instanz an ein Software-Tool, das diese Aufgaben abbilden soll. Bereits um in der Chemie basale Zusammenhänge fachlich korrekt darlegen zu können, muss die domänenspezifische Symbolschreibweise der Summenformeln bzw. Reaktionsgleichungen beherrscht werden.

Tabelle 11: Formulierungshilfe für Lernziele (Handlungskomponente) nach Cursio und Jahn (2015)

Stufe	Beschreibung der Taxonomiestufe	Verben
1. Wissen	Lernende können Informationen wiedergeben. Sie kennen konkrete Termini, Definitionen, Fakten, Daten, Regeln, Verfahren, Merkmale Kriterien etc.	angeben, nennen, wiedergeben, beschreiben, aufzählen, (er)kennen
2. Verstehen	Lernende können Sachverhalte erklären, Beispiele anführen, Aufgabenstellungen interpretieren oder ein Problem in eigenen Worten wiedergeben	erläutern, zusammenfassen, interpretieren, darstellen, übersetzen, auslegen, berichten, abstrahieren, skizzieren, illustrieren, herausstellen.
3. Analysieren	Lernende können ein neues Problem durch Transfer des Wissens lösen.	anwenden, ordnen, berechnen, übertragen, zuordnen, erklären, aufstellen, einordnen, berichten, voraussagen, unterscheiden, beschreiben, vergleichen, verallgemeinern, organisieren, einteilen, quantifizieren
4. Evaluieren (Beurteilen)	Lernende können evidenzbasierte, qualitative und quantitative Urteile zu Sachverhalten anhand von Kriterien anstellen.	bewerten, entscheiden, beurteilen, ermitteln, hinterfragen, überprüfen, vergleichen, wählen, unterscheiden, Kriterien aufstellen, einschätzen, folgern, gewichten, ermessen, einstufen.
5. Erschaffen	Lernende können aus mehreren Elementen eine neue Struktur aufbauen oder eine neue Bedeutung erschaffen, können neue Lösungswege vorschlagen, neue Schemata entwerfen oder begründete Hypothesen entwerfen.	entwickeln, erzeugen, erstellen, planen, ausarbeiten, anfertigen, herstellen, konzipieren, gestalten.

Diese Symbolschreibweise hat eigene Regeln mit sich wechselseitig bedingenden Beziehungen, die zusätzlich zu mathematischen Gesetzmäßigkeiten, eine Kombination aus grafischen und semantischen Anforderungen an das Software-Tool stellt, da

die Antwort der Lernenden sowohl in domänenspezifisch üblicher Weise abgegeben werden können muss, als auch in dieser Form vom System evaluiert werden muss (siehe „Dreigliedriges Spannungsfeld in Feedback präsentierenden Softwarelösungen“, Abbildung 10, S. 39). Die Anforderungen an die Software hinsichtlich Darstellung der Aufgabeninhalte, insbesondere von chemiespezifischen Schreibweisen in Lernaufgaben zu einem bestimmten Lernziel, ist folglich abhängig von den Operatoren (Cursio & Jahn, 2015).

Bei der Überführung von Paper-Pencil Aufgaben in ein online-taugliches Format begegnet man zudem immer wieder der Kritik, dass durch diesen Transfer die angesprochenen Lernziele der Aufgaben sich verändern würden. Dies konnte für chemische Kontexte allerdings unter Beachtung einiger Aspekte widerlegt werden:

„[T]here is no medium effect when specific traditional paper examinations in Chemistry and Computing are transferred into electronic format[...]. Typical differences between computer and paper assessments are the reading of questions from a computer screen, and the entering of answers via a keyboard or mouse instead of by pen on paper. In addition, the use of a computer to present and answer questions can mean that the questions may require changes to the wording and presentation, perhaps because the form of the required answer cannot be produced for, or marked by, the computer, often resulting from limitations of the assessment package“ (Ashton, Beevers, Korabinski & Youngson, 2005, S. 771).

Um die Anforderungen an die Software zu verdeutlichen, wurden die Aufgaben zuvor in Form von Skripten hinsichtlich Formatierung und Layout unter Beachtung von Tabelle 6 (S. 50) vorformuliert (z. B. *Signaling*, *Spacial Contiguity* und *Segmenting Principle*, siehe Abbildung 16).

Kapitel 6: Aufgabenentwicklung

Teilreaktionen formulieren

a) Mit Hilfe der Oxidationszahlen wird ermittelt, wie viele **Elektronen** aufgenommen bzw. abgegeben wurden.

Oxidation: $\text{SO}_2 \rightleftharpoons \text{SO}_4^{2-} + 2e^-$

Reduktion: $\text{MnO}_4^- + 5e^- \rightleftharpoons \text{Mn}^{2+}$

Da die Oxidationszahl des Schwefel-Atoms von +IV auf +VI steigt, müssen 2 Elektronen abgegeben worden sein.
Da die Oxidationszahl des Mangan-Atoms von +VII auf +II sinkt, müssen 5 Elektronen aufgenommen worden sein.

b) Dann wird überprüft, ob die Ionenladungen auf beiden Seiten der Gleichung übereinstimmen.
Ggf. wird mit einer entsprechenden Anzahl an **H^+ – Ionen** in saurer Lösung bzw. **OH^- – Ionen** in basischer Lösung ausgeglichen.

Oxidation: $\text{SO}_2 \rightleftharpoons \text{SO}_4^{2-} + 2e^- + 4\text{H}^+$

Reduktion: $\text{MnO}_4^- + 5e^- + 8\text{H}^+ \rightleftharpoons \text{Mn}^{2+}$

Da die Reaktion in saurer Umgebung abläuft, wird in diesem Fall mit **H^+ – Ionen** ausgeglichen.
Bei der Oxidation liegt auf Eduktseite eine Ionenladung von -1 vor, auf der Produktseite eine Ionenladung von -3. Um die Ionenladungen auszugleichen, müssen auf der Produktseite **2H^+ – Ionen** ergänzt werden.
Bei der Reduktion liegt auf der Eduktseite eine Ladung von -6 vor, auf der Produktseite eine Ladung von +2. Um die Ladungen auszugleichen, müssen auf der Eduktseite **8H^+ – Ionen** ergänzt werden.

c) Anschließend wird die Atombilanz überprüft und ggf. durch eine entsprechende Anzahl an **H_2O – Molekülen** ausgeglichen.

Oxidation: $\text{SO}_2 + 2\text{H}_2\text{O} \rightleftharpoons \text{SO}_4^{2-} + 2e^- + 4\text{H}^+$

Reduktion: $\text{MnO}_4^- + 5e^- + 8\text{H}^+ \rightleftharpoons \text{Mn}^{2+} + 4\text{H}_2\text{O}$

Schritt 3: Aufstellen der Redoxgleichung

Erweiterungsfaktor

Damit die Anzahl der abgegebenen Elektronen und der aufgenommenen Elektronen übereinstimmt, werden die Teilgleichungen mit dem entsprechend kleinstmöglichen **Faktor** multipliziert.

Abbildung 16: Realisierung multimedialer Prinzipien am Beispiel eines elaborierten Feedbackausschnitts von Aufgaben zur Redoxreaktion

6.2 Das Software-Tool

Gesucht wurde daher eine Softwarelösung, die sowohl die Möglichkeit für eine chemiespezifische Schreibweise, als auch eine automatisierte synchrone Evaluation von Antworten, die unter anderem in dieser Schreibweise erfolgen muss, ermöglichte. In Zusammenarbeit mit dem Moodle-Kompetenzzentrum der Universität sowie im interdisziplinären Austausch mit der Informatik rückte das vom Paluno (The Ruhr Institute for Software Technology) entwickelte Java-basierte E-Assessment System „JACK® – Lernen. Testen. Prüfen.“ in den Fokus.¹⁷ Bis 2011 wurde dieses Tool zwar ausschließlich im Kontext der Informatik verwendet, Ende 2011 wurde es jedoch um einige Aufgabentypen erweitert, die nun die automatisierte Evaluation von Multiple-Choice Fragen und vor allem auch Fill-In Aufgaben ermöglichten. Damit öffnete sich dieses Tool auch anderen Anwendungsgebieten jenseits der Informatik. Besonders die Anbindung eines externen Computer Algebra Systems, welche die Evaluation (komplexerer) mathematischer Eingaben ermöglichte, machte JACK® dabei attraktiv für die Anwendung in der Hochschullehre der Mathematik (Striewe, 2016).

Damit erschien das E-Assessment System JACK® aus zweierlei Gründen attraktiv für dieses Projekt. Einerseits bietet das Übungs- und Prüfungssystem gute Voraussetzungen für die regelbasierte Evaluation von mathematischen Aufgaben, die eine bedeutende Rolle in der Chemie spielen (Goldhausen, 2015; Kimpel, 2017). Auch für automatisierte Rückmeldung zu spezifischen Fehlern bestehen für mathematische Aufgaben bereits weitreichende Möglichkeiten, da Fehler, die einer Systematik folgen, in Form von abprüfbaren Regeln hinterlegt werden können. Andererseits legt das Paluno besonders Wert auf rege „interdisziplinäre Zusammenarbeit mit Fachdidaktikern aus vielen verschiedenen Domänen“ (Goedicke, 2022), was es als Teil der Fakultät für Informatik der Universität Duisburg-Essen für einen sehr direkten Austausch attraktiv machte.

Zu Beginn der Entwicklungsphase der Aufgaben begrenzte sich die Auswahl auf die folgenden in JACK® realisierbaren Operatoren (siehe Tabelle 12). Gerade zu Beginn der Entwicklungsphase wurden noch weitere Aufgabenformate konzipiert (z. B. Drag-and-Drop-Aufgaben, Aufgaben die Keilstrichformeln miteinschließen), die dann

¹⁷ Sowohl formative also auch summative Assessments waren hiermit zu realisieren. Jede Einreichung einer Aufgabe wird dabei vollständig unabhängig von einer anderen getestet. Bis 2015 wurde JACK® im Informatikkontext an vier Universitäten in Deutschland und England eingesetzt und erhielt dabei sehr viel positiven Zuspruch seitens der Studierendenschaft.

letztlich an der Software scheiterten. Daher wurde sich im weiteren Verlauf der Aufgabenentwicklung regelmäßig mit den Entwicklern von JACK[®] ausgetauscht, um auch einige problematische Aufgaben zu realisieren.

Tabelle 12. Verteilung verwertbarer Operatoren (Trauten, 2018)

möglich	nicht möglich
	deuten
	diskutieren
	analysieren
bestimmen	erläutern
nennen	vergleichen
klassifizieren	zeichnen
berechnen	skizzieren
auswerten	Hypothesen aufstellen
prüfen	interpretieren
	vergleichen
	beurteilen
	bewerten
	erklären

Während Freitext und Zeichenaufgaben aufgrund von hochkomplexen Evaluationsstrukturen für fehlerspezifisches Feedback nicht erschlossen werden konnten, wurde Anfang 2019 der Editor für chemische Summenformeln und Reaktionsgleichungen fertiggestellt (Pobel & Striwe, 2019), sodass für die Hauptstudie Aufgaben, die das Formulieren von Reaktionsgleichungen in chemiespezifischer Schreibweise enthielten, integriert werden konnten. Des Weiteren wurde seitens Paluno ein Eingabetool für Atomorbitaldiagramme fertiggestellt, das zumindest anhand eines lernzielspezifischen Aufgabentyps das freie Zeichnen unnötig machte. So ergab sich für die domänenexklusiven Schreibweisen der Chemie mit ihrer spezifischen Semantik die nötige Fehlersensitivität für passgenaues Feedback.

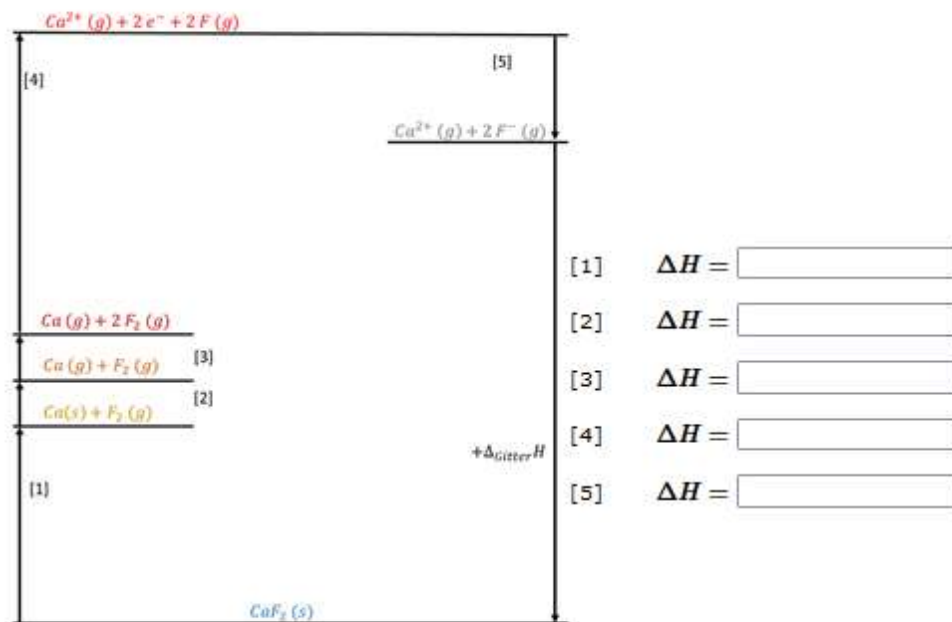
Für die Operatoren, die sich realisieren ließen, ergaben sich lediglich geringfügige Einschränkungen hinsichtlich der Antwort-Texte. Diese konnten, wie beschrieben, nicht als Freitextantwort abgefragt werden. Stattdessen wurden entsprechende Antwortsätze formuliert und an entsprechenden Stellen mit Fill-In Feldern versehen (siehe Abbildung 17).

Berechnung der Gitterenthalpie: Der Born-Haber-Kreisprozess

Nutzen Sie für diese Aufgabe bitte das **"MATERIAL Thermodynamische Kennwerte ausgewählter Elemente und Verbindungen"** im Moodle Kurs.

Betrachtet wird das Salz Calciumfluorid (CaF_2) im festen Zustand.

1. Ordnen Sie die jeweiligen Enthalpiebeträge (in kJ) den entsprechenden Reaktionen 1-5 (s. Abb. unten) zu. Sollten Sie das Vielfache von einem Wert ausrechnen müssen, so geben Sie bitte den ausgerechneten Wert ein.



2. Stellen Sie die Gleichung zur Berechnung der Gitterenergie auf. Orientieren Sie sich dabei an folgendem Beispiel:

$$-156 \text{ kJ} + 18 \text{ kJ} + 120 \text{ kJ} \dots$$

$$\Delta_{\text{Gitter}}H = \text{}$$

- 2.1 Berechnen Sie die Gitterenthalpie für Calciumfluorid (CaF_2).

$$\Delta_{\text{Gitter}}H = \text{}$$

Abbildung 17: Beispiel strukturierender Antwortsätze aus dem Aufgabenset zum Born-Haber-Kreisprozess

Diese Antwortsätze gaben somit teilweise auch Hinweise für nötige Zwischenschritte. Diese Strukturierung des Lösungsvorgangs ist aber unproblematisch, da es sich hier um Lernaufgaben handelt und Lernaufgaben anders als Testaufgaben (Proske et al., 2012) den Lernvorgang direkt unterstützen dürfen und sollen. Zudem wird über die Anzahl abgefragter Zwischenschritte – vgl. Abbildung 17 – ein detaillierteres Feedback ermöglicht, das der Rückmeldung eines menschlichen Tutors nahe kommen kann. Das Feedback, das nach Einreichen der Antwort den Studierenden angezeigt wurde, konnte mithilfe des Tools zudem immer so angezeigt werden, dass die Studierenden das Feedback und ihre fehlerhafte Eingabe simultan betrachten konnten, sodass ein Abgleich mit dem Feedback und insbesondere mit den fehlerspezifischen Hilfestellungen im Feedback ermöglicht wurde.

6.3 Der Feedback Algorithmus

Die nächste Frage, die es zu beantworten galt, war die nach der Feedback-Form, also nach dem Modus, dem Timing und der Adaptivität. Wie bereits in Kapitel 2.2 beschrieben, wird eine vorherige Leistungsmessung benötigt, um fehlerspezifisches Feedback geben zu können. Dazu bedarf es einer Aufgabe, die einem multiple-try Schema mit mindestens zwei Antwortversuchen folgt, denn zusätzlich zu einer fehlerspezifischen Rückmeldung soll auch tutorielles Feedback bereitgestellt werden. Das heißt, wie in Kapitel 2.3 beschrieben, dass passend zu bestimmten Fehlerkonstellationen automatisiert eine Hilfestellung im Feedback enthalten sein soll, die den Lernenden hilft, die Aufgabe selbst zu vollenden, aber ohne die Lösung der Aufgabe zu präsentieren. Ähnliche Überlegungen wurden von Narciss und Huth (2004) als Hilfestellung für Bruchrechenaufgaben angestellt und ein sog. Bug-Related-Tutoring (BRT) Feedback Algorithmus entwickelt. Dieser dreistufige Feedback-Algorithmus scheint eine vielversprechende Rahmenkonstruktion für etwaige digitale Lernaufgaben darzustellen. Ursprünglich ist dieser Algorithmus allerdings konzipiert worden, um die lernförderliche Wirkung von verschiedenen Feedback-Typen miteinander zu vergleichen und Erkenntnisse hinsichtlich der Effektivität zu gewinnen. Im ersten Feedback wurde dazu lediglich *Knowledge of Result* (KR) präsentiert. Erst im zweiten Feedback wurde dann zusätzlich zum KR noch ITF Feedback präsentiert (KTC, KM, KH; siehe Tabelle 3). Das letzte Feedback enthielt dann neben Hinweisen zu den Fehlern (KM), die im dritten Lösungsversuch immer noch gemacht worden sind, zusätzlich noch *Knowledge of Correct Response* (KCR), die richtige Lösung wurde dabei feingliedrig mit Teilschritten präsentiert. ITF wurde, wenn man der Definition

von Narciss (2006) folgt, also ausschließlich im Anschluss an den zweiten Lösungsversuch präsentiert. Nach dem dritten Versuch wurde zwar elaboriertes fehlerspezifisches Feedback, jedoch kein ITF präsentiert, da die Präsentation der korrekten Lösung ITF ausschließt. Der in dieser Arbeit verwendete EST-Feedback Algorithmus orientiert sich hinsichtlich der verwendeten Feedback-Typen stark an diesem, wurde aber inhaltlich für chemische Inhalte modifiziert. Um einen sauberen Vergleich anstellen zu können, wurde analog zu dem auf Narciss und Huth (2004) Arbeit basierenden multiple-try Algorithmus ein korrekativer Feedback Algorithmus konzipiert. Die Gegenüberstellung der beiden Algorithmen findet sich in Abbildung 18 (S. 74). Sowohl die Moderatorenanalyse von Kluger und DeNisi (1996) als auch andere Studien (z. B. Schuhmacher, 2008; O. Chen et al., 2017) liefern allerdings eine theoretische Grundlage dafür, im Sinne adaptiven Feedbacks die erste Feedback-Nachricht mit KR beizubehalten, damit Lernende mit hinreichendem Vorwissen im Falle kleinerer Patzer (z. B. Flüchtigkeitsfehler, Zahlendreher o. Ä.) selbst ihre(n) Fehler berichtigen können, anstatt möglicherweise von einer umfangreichen Feedback-Nachricht aufgehalten oder gar verwirrt zu werden. Die Chemie im ersten Semester hält teilweise Aufgaben für die Studierende bereit, die sich (1) in Bezug auf den Grad der Komplexität stark von simplen Subtraktionsaufgaben für Grundschüler, für die dieser Algorithmus entwickelt wurde, unterscheiden und (2) auch noch in eine digitale Form überführt werden sollen, die sinnvolles automatisiertes Feedback präsentieren kann.

Inwiefern also ein Analogon des oben beschriebenen BRT-Algorithmus für ein online Tutorium zur Erstsemestervorlesung der Allgemeinen Chemie genutzt werden kann, muss daher noch geprüft werden. Der Algorithmus selbst sieht allerdings keine Begrenzung der Aufgabenmenge auf dem Weg zur Testaufgabe vor. Hier wurde ein Maximum von sechs Lernaufgaben pro Lernziel festgelegt, um eine adäquate Menge an Lernaufgaben zum Üben bereitzustellen, aber auch um die Studierenden nicht durch eine zu große Flut an Aufgaben langfristig zu demotivieren. Mit der erfolgreichen Bearbeitung der Testaufgabe konnten dann anteilig Bonuspunkte für die Klausur erworben werden, was zusätzlich motivieren sollte (s. o.).

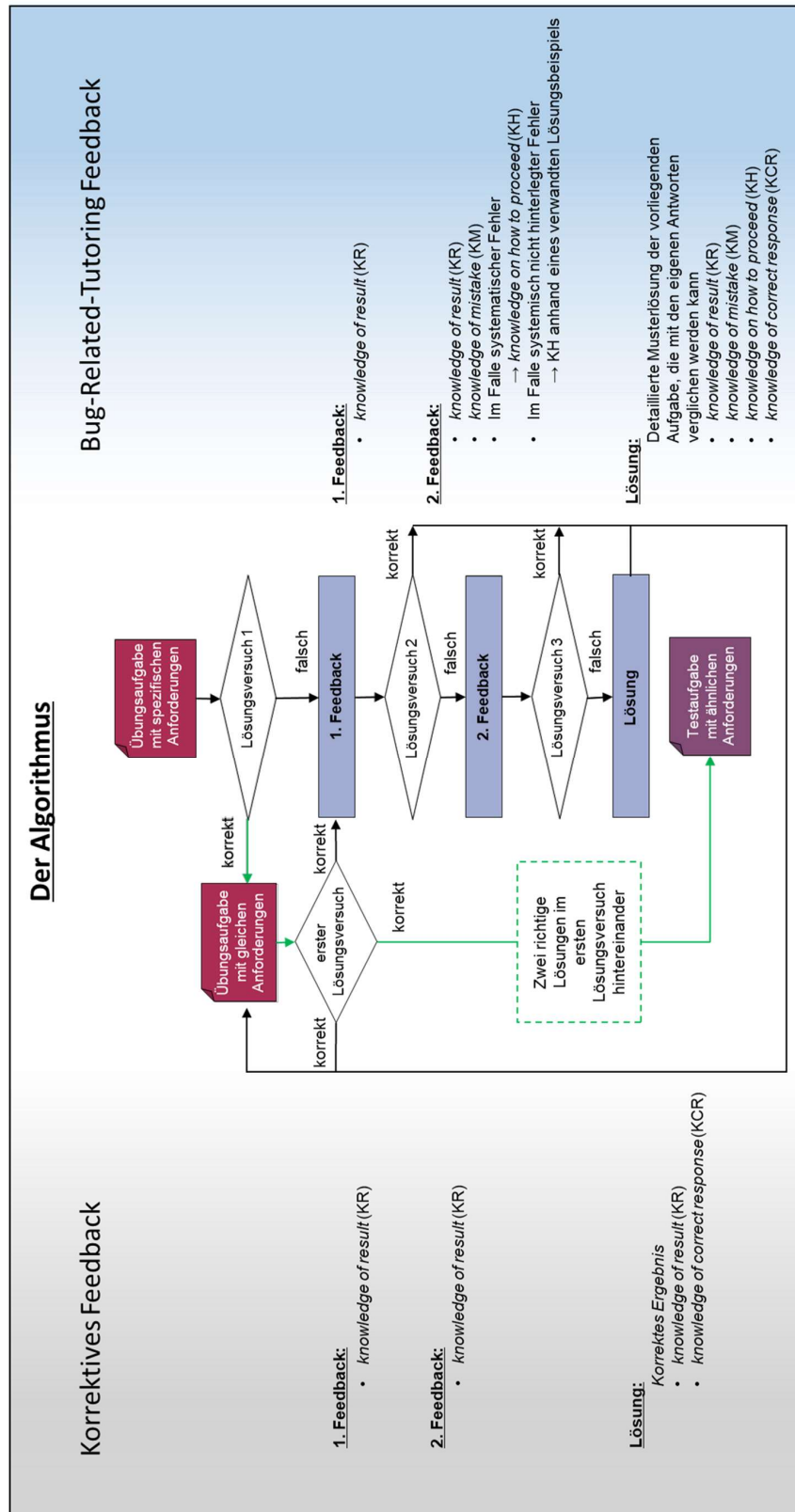


Abbildung 18: Gegenüberstellung des Korrekativen Feedback Algorithmus mit dem EST-Feedback-Algorithmus nach Narciss und Huth (2004)

Scheinbar bildet die Software-Lösung nicht nur für die Feedback-Konzeption, sondern auch für die Übertragung in das Online-Format den limitierenden Faktor, da die Aufgaben ihren Anforderungen herkömmlichen Lernaufgaben möglichst stark ähneln sollten, um die angesprochenen Lernziele möglichst nicht zu verändern.

Während der Feedback-Algorithmus in JACK® in die Programmierung der Aufgaben mit eingebaut werden musste, konnte die Verknüpfung zwischen den Aufgaben verhältnismäßig einfach realisiert werden. Da das Tool JACK® über eine Schnittstelle mit dem LMS Moodle kommunizieren kann, ließen sich JACK® Aufgaben als sog. „externes Tool“ direkt in der Moodle Kursoberfläche einbetten. Die Leistungsdaten der Studierenden in den Einzelaufgaben lassen sich so mit Weiterleitungsbedingungen – sogenannten Lernpfaden – im LMS Moodle verknüpfen. Um im Sinne von Abbildung 18 den jeweiligen Algorithmus zu vollziehen, gibt es zwei Varianten. Eine Möglichkeit ist, dass zwei Aufgaben hintereinander im ersten Antwortversuch, also ohne (EST) Feedback erhalten zu haben, korrekt gelöst werden. Alternativ erfolgt nach sechs Aufgaben, die mit beliebiger Performanz gelöst werden können, die direkte Weiterleitung zur Testaufgabe. Abbildung 19 zeigt exemplarisch die Fortführung des in Abbildung 20 zu sehenden Aufgabensets zum Arbeiten mit dem Periodensystem.

Jede Aufgabe eines Sets ist mit einer solchen Lernpfad-Regel versehen, die Summe dieser Lernpfad-Regeln realisiert die von der Performanz in den Einzelaufgaben abhängige Vermittlung zwischen den JACK®-Aufgaben und damit den aufgabenübergreifenden Anteil des in Abbildung 18 beschriebenen Algorithmus.

Kapitel 6: Aufgabenentwicklung

▼ Voraussetzungen

Voraussetzungen

Teilnehmer/in muss von den folgenden Bedingungen mindestens eine erfüllen

Teilnehmer/in muss von den folgenden Bedingungen alle erfüllen

Bewertung Ü1| 1.1 Arbeiten mit de

☐ muss $\geq$ sein % ✕

☒ muss $<$ sein 50 %

und

Bewertung Ü1| 1.2 Arbeiten mit de

☒ muss $\geq$ sein 50 % ✕

☐ muss $<$ sein %

Voraussetzung hinzufügen

oder

Teilnehmer/in muss folgende Bedingung erfüllen

Bewertung Ü1| 1.2 Arbeiten mit de

☐ muss $\geq$ sein % ✕

☒ muss $<$ sein 50 %

Voraussetzung hinzufügen

Voraussetzung hinzufügen

Abbildung 19: Lernpfad-Regeln exemplarisch für die dritte Aufgabe des Aufgabensets „Arbeiten mit dem PSE“

6.4 Implementierung der JACK® Aufgaben und Gestaltung des Moodle-Kurses

Das finale Produkt des in Abbildung 15 dargestellten Prozesses umfasst insgesamt 120 feedbackgestützte Lernaufgaben zu 20 Lernzielen, die sich aus den Modulinhalten ergeben, sowie 20 Testaufgaben für den Bonuspunkterwerb. Um die Qualität der Aufgaben hinsichtlich fachlicher und didaktischer Aspekte sicherzustellen, wurden die Aufgabenstämme und das Feedback der EST-Aufgaben durch insgesamt zwölf Chemiedidaktiker/-innen mit langjähriger akademischer Erfahrung überprüft und bewertet. Jede Aufgabe wurde dabei von zwei Expertinnen/Experten bewertet. Diese Einschätzungen hinsichtlich sprachlicher und inhaltlicher Qualitätsaspekte wurden für die Optimierung der Aufgaben herangezogen.

Die technischen Möglichkeiten des Software-Tools JACK®- bieten vielseitige Möglichkeiten für die Realisierung von Aufgaben, insbesondere für den mathematischen Anteil der (Allgemeinen) Chemie, bei dem Studierende immer wieder Probleme haben (Goldhausen, 2015; Kimpel, 2017), aber dennoch ergaben sich durch die Bandbreite der Modulinhalte Grenzen (Freitexteingabe und Zeichenaufgaben), die durch das Tool auch bislang nicht behandelt werden können. Zu diesen Inhalten wurden dann im Moodle-Kurs spezifische Übungsblätter als Offline-Übung angeboten, die entsprechende Übungsaufgaben der Übungen zur Allgemeinen Chemie Vorlesung aus den Vorjahren bediente. Diese Übungsblätter wurden wöchentlich freigeschaltet und mussten innerhalb einer Woche bearbeitet werden. Die Inhalte des Moodle-Kurses wurden mit dem Modulverantwortlichen zusammen ausgewählt. Die sog. Offline-Übungen wurden bereits für die Pilotierung angefertigt und sind über alle Erhebungen, von kleineren Änderungen abgesehen, in ihrer inhaltlich ursprünglichen Form genutzt worden.

Die Oberfläche des LMS Moodle bietet ebenfalls – wie bereits angeklungen – eine gute Ausgangslage für die strukturierte Präsentation der Tutoriumsinhalte. Jede Art von möglichen Aktivitäten hatte inhaltsspezifische Tokens, die die bessere Orientierung im Moodle-Kurs gewährleisten sollten (z. B. Puzzleteile für JACK®- Aufgaben, Fragebögen für Zwischenerhebungen, PDF-Symbole für Druckbares Material und Übungsblätter, siehe Abbildung 20). Wie man hier sehen kann, wurden bearbeitete digitale-Kursinhalte (JACK®-Aufgaben, Abgaben der Offline Übungen) durch Check-Boxen (am rechten Rand) automatisch als bearbeitet gekennzeichnet. Für alle herunterladbaren Inhalte konnte manuell von den Lernenden ein Haken gesetzt werden. Der Kursraum baute sich wöchentlich, inhaltlich parallel zur Vorlesung, sukzessiv auf. Von jedem JACK®-Aufgabenset, das das für die entsprechende Vorlesungswoche vorgesehen war, wurde die erste Aufgabe freigeschaltet. Wie viele Aufgaben

Kapitel 6: Aufgabenentwicklung

dann zusätzlich freigeschaltet wurden, oblag der Performanz der Lernenden in den JACK®-Aufgaben.



Abbildung 20: Ausschnitt aus dem Moodle-Kurs des Online-Tutoriums WiSe 19/20

Bereits für die Pilotierung erhielt jede Interventionsgruppe (KOR/EST) nur Zugang zu einem eigenen IG-spezifischen Moodle-Kurs. Diese Kurse waren dabei optisch nicht voneinander unterscheidbar und vollkommen identisch aufgebaut. Andere Kursteilnehmende konnten auch nicht, wie sonst üblich, im Kursraum eingesehen werden. Diese Funktion war für die beiden Kursräume vom Moodle Kompetenzzentrum deaktiviert worden. Für die Hauptstudie wurden die bestehenden Kursräume jeweils immer in ein neues Semester übertragen und durch neue JACK®-Aufgaben an inhaltlich zu Vorlesung passender Stelle ergänzt. Das Grundgerüst der Moodle-Kurse blieb folglich immer bestehen.

Kapitel 7: Stichprobenbeschreibung

Die innerhalb dieser Studie erhaltene (unbereinigte) Gesamtkohorte¹⁸ von $N = 338$ ($M_{\text{Alter}} = 20.72$ Jahre, 57.6 % männlich, 42.4 % weiblich) aus Erstsemesterstudierenden der Fächerwahl Bachelor Chemie und Bachelor Water Science hat einen Anteil von 17.6 % Probanden, die nicht in Deutschland geboren worden sind. Darüber hinaus haben 40.2 % einen Migrationshintergrund, d. h., dass mindestens ein Elternteil nicht in Deutschland geboren worden ist. 29.1 % der Probanden gaben an, dass kein Elternteil in Deutschland geboren wurde. Insgesamt gaben 69.8 % der Probanden die deutsche Sprache als Muttersprache an.

Der Gesamtdatensatz speist sich aus der Pilotierungskohorte (WiSe 2018/2019), der Kohorte der ersten Hauptstudie (WiSe 2019/2020) und der Kohorte des zweiten Erhebungsjahres der Hauptstudie (WiSe 2020/2021). Das letzte Erhebungsjahr wurde abweichend von der Planung durchgeführt, um den Datensatz zu erweitern, um den Schwund an Probanden auszugleichen, der sich vorwiegend durch verpasste Teilnahme am Allgemeinen Chemie Test zum Post-Zeitpunkt ergeben hatte, wie in Tabelle 13 zu erkennen ist.

Tabelle 13: Stichprobenbeschreibung¹⁹

	Demografie			+	Allg. Chemie (Pre)			+	Allg. Chemie (Post)			+	Aktive Übungsteilnahme		
	N	%	%(♀) ¹		N	%	%(♀) ¹		N	%	%(♀) ¹		N	%	%(♀) ¹
Ges	291	100 ²	42.6		234	80.4 ²	42.7		149	51.2 ²	36.9		122	41.9 ²	40.2
Pilot	84	100 ³	39.3		77	91.7 ³	39.0		54	64.3 ³	35.2		39	46.4 ³	43.6
HS1	92	100 ³	41.3		83	90.2 ³	42.2		58	63.0 ³	25.7		53	60.4 ³	24.5
HS2	115	100 ³	44.3		74	64.3 ³	47.3		37	32.2 ³	56.8		30	26.1 ³	63.3

Anmerkung: ¹ Anteil weiblicher Studierender bezogen auf den jeweiligen MZP; ² Bezogen auf die Gesamtstichprobe (Studierende, die den Demografie-Fragebogen bearbeitet haben); ³ bezogen auf die Teilstichprobe je WiSe (Studierende, die den Demografie-Fragebogen bearbeitet haben)

¹⁸ Einige Studierende sind erst in der zweiten Vorlesungswoche erschienen und haben somit den Demografie-Fragebogen verpasst und trotz Angebot nicht nachgeholt. Andere Studierende haben zwar am Demografie-Fragebogen teilgenommen, aber zu einem anderen MZP mindestens einen Test nicht oder zu wenige JACK® Aufgaben bearbeitet.

¹⁹ Diese Tabelle baut auf dem bearbeiteten Demografie-Fragebogen auf und beschreibt die verbleibenden Probandenzahlen, die sich ergeben, wenn alle Fragebögen bis hin zur aktuellen Spalte ausgefüllt worden sind.

Kapitel 7: Stichprobenbeschreibung

Es wurden zudem auch Nachholtermine mit flexiblen Zeiten angeboten, diese konnten aber nur in einem begrenzten Zeitraum (Semesteranfang und Semesterende je drei Wochen nach dem MZP) angeboten werden, um die Vergleichbarkeit mit den anderen Probanden zu gewährleisten, da in den ersten Semesterwochen bereits viel Schulstoff wiederholt und die Klausurphase dicht hinter der Vorlesungszeit liegt und so durch klausurbedingtes Lernen auch eine Verfälschung der Ergebnisse zu erwarten war. Der Hauptschwund (48.8 %) ergibt sich also durch einen versäumten Post-Test. Da zudem nur sehr wenige Studierende alle verfügbaren Aufgabensets im Sinne des Algorithmus vollständig bearbeitet haben²⁰, wurde bereits im Rahmen der Pilotierungsstudie, in der sich dieses Verhalten bereits abzeichnete, die aktive Arbeit mit den Übungsaufgaben an die Regel geknüpft, dass mindestens zwei Aufgaben aus mindestens fünf der Aufgabensets bearbeitet worden sein mussten, um als „aktiv lernende Person“ zu gelten. Fünf Aufgabensets stellten dabei die Hälfte der in der Pilotierung verfügbaren Aufgabensets dar. Um am Ende der Studie einen Effekt auf den Studienerfolg der Studierenden untersuchen zu können, wurde diese weichere Regel für aktive Übungsteilnahme auch für die Folgejahre beibehalten. So ergibt sich folgende Verteilung der in JACK®- „aktiven“ Gesamtkohorte für die Evaluation des Fachwissenszuwachs (siehe Tabelle 14).

Tabelle 14: Verteilung der „aktiv“ Studierenden pro Interventionsgruppe

DEMO	Interventionsgruppe		Σ		„aktiv“	Interventionsgruppe		Σ
	KOR	EST				KOR	EST	
Kein Leistungskurs	90	96	186		Kein Leistungskurs	28	33	61
Leistungskurs	50	55	105		Leistungskurs	30	31	61
Σ	140	151	291		Σ	58	64	122

²⁰ Die Bearbeitung der Übungsaufgaben ist mit Bonuspunkten für die Klausur und des Weiteren mit einer Probandenvergütung von 50 € pro Person belohnt worden. Dazu hätten die Studierenden mindestens zwei Aufgaben eines Sets in Folge richtig lösen oder alle sechs Aufgaben eines Aufgabenpakets sowie die anschließende Testaufgabe als Lernzielabschluss bearbeiten müssen.

Kapitel 8: Ergebnisse der durchgeführten Studien

Wie Trauten, Eitemüller und Walpuski (2023) bereits beschrieben haben, wurden die Leistungsdaten mithilfe der Item Response Theorie (Rasch Modelle) ausgewertet. Daraus ergibt sich der Vorteil, dass sich auch sukzessiv entstandene Stichproben gut auf einer gemeinsamen Skala abbilden lassen. Dies wird durch die Fixierung der Itemschwierigkeiten auf Basis der Messdaten einer Teilstichprobe zu einem Messzeitpunkt erreicht. Wird die Itemschwierigkeit (σ) auf einen Messzeitpunkt fixiert, liegen die in diesem Modell geschätzten Personenfähigkeiten (θ) auf derselben Skala und lassen sich so hinsichtlich ihrer Aussagekraft gleichwertig verwenden und vergleichen. Walter (2005) klassifiziert den WLE-Schätzer als geeignetsten Schätzer für Personenfähigkeiten im Rahmen von IRT-Modellen. Allerdings resultieren hierbei keine z-standardisierten Personenfähigkeiten, sondern mit den WLEs (Weighted Maximum Likelihood Estimates; Warm, 1989), also Personenfähigkeitsparametern auf einer Logit-Skala.

Diese Methode wird *fixed common item parameter method* (Q. Chen, 2013) genannt. Die Fixierung der Itemschwierigkeiten auf einen Messzeitpunkt kann jedoch einen unerwünschten Item-Parameter-Drift hervorrufen. Besonders zwei Aspekte deuten hier augenscheinlich in gegensätzliche Richtungen. Einerseits, so Q. Chen (2013), lassen sich präzisere Ergebnisse erhalten, wenn auf einen Messzeitpunkt einer Teilstichprobe mit gutem Fähigkeitsniveau und Varianz fixiert wird. Für Tests, die den Lernzuwachs während einer Zeitspanne abbilden sollen, sind dies häufig Post-Messzeitpunkte. Hierbei geht es auch darum, Boden- und Deckeneffekte zu vermeiden. Andererseits sollten Tests auch nicht zu häufig durchgeführt werden, da so ein Lerneffekt des Tests einsetzen und schlimmstenfalls die Itemschwierigkeiten auf einen sehr späten Zeitpunkt durch sog. *item over-exposure* deutlich verfälscht werden können (Q. Chen, 2013). Im Rahmen der ersten Förderphase des interdisziplinären Projektes ALSTER wurden die Itemschwierigkeiten auf den Pre-Messzeitpunkt der in dieser Arbeit als Kontrollkohorte bezeichneten Teilstichprobe fixiert (WiSe 2016/2017 MZP1; Averbek, 2020). Da der Test ursprünglich für die Erfassung des Chemiewissens am Ende der Oberstufe konzipiert wurde (Freyer, 2013), war von einem bereits der Schwierigkeit des Testinstruments angemessenen Fähigkeitsniveau auszugehen. Zudem war die Stichprobe bislang durch diesen Test unbelastet. Die Größe der Stichprobe ermöglichte hierbei eine valide Schätzung der Itemschwierigkeiten, wie es sich auch in den Testparametern in Kapitel 5 zeigt. Die anderen Tests wurden aus Konsistenzgründen ebenfalls auf diesen Messzeitpunkt dieser Teilstichprobe fixiert. Der überwiegende Anteil der Items beim Einsatz in den Folgejahren, so auch bei

dieser Kohorte, führte mit den fixierten Itemschwierigkeiten aus ALSTER I zu guten Infit- und Outfit-Parametern innerhalb der von Bond et al. empfohlenen Spanne für Item-Mittelwertsquadrate (Range: 0.7-1.3; 2020, S. 244). Einige wenige Items wurden dabei frei geschätzt (MZIP1: Mathematik, 4 Items; Allg. Chemie, 9 Items; MZIP2: Allg. Chemie, 8 Items), sodass sie in die Skalierung der fixierten Items einsortiert werden konnten²¹. Damit ist ein Vergleich zwischen den Jahrgängen hinsichtlich folgender Leistungsdaten möglich: Fachwissen in der Allgemeinen Chemie und mathematische Fähigkeiten. Die Testparameter der kognitiven Fähigkeiten liegen nur für diese Kohorte IRT-skaliert vor und wurden daher frei geschätzt.

8.1 Pilotierung

Im Rahmen der Pilotierung der JACK®-Aufgaben, welche im selben Design wie in der Hauptstudie durchgeführt wurde und sich lediglich durch Menge an verfügbaren JACK®-Aufgaben von der Hauptstudie unterscheidet, wurden in der ersten Vorlesung über den Demografie-Fragebogen $N = 84$ Probanden erfasst. Regressionsanalysen für die Sub-Gruppen ergaben keine signifikanten Unterschiede für mathematische Fähigkeiten ($F(3, 77) = 1.355, p = 0.263.; R^2 = 0.013$), jedoch einen erwartungskonformen signifikanten Vorteil im Fachwissen Allgemeine Chemie zum Pre-Messzeitpunkt ($\beta = 0.465, p \leq 0.001$) für KOR- und EST-Probanden der IG3 und IG4, die in der Schule einen Leistungskurs Chemie belegt hatten ($F(3, 73) = 7.558, p \leq 0.001.; R^2 = 0.206$). Ein signifikanter Unterschied hinsichtlich der am Pre-Messzeitpunkt abgefragten figuralen Skala des Tests für kognitive Fähigkeiten konnte zwischen den Gruppen nicht gefunden werden ($F(3, 73) = 1.725, p < 0.169.; R^2 = 0.206$).

8.1.1 Ergebnisse

Insgesamt liegen $n = 53$ Fälle vor, die den Demografie-Fragebogen ausgefüllt sowie am Pre- und Post-Test teilgenommen haben, und $n = 65$, die die Klausur am ersten und/oder zweiten Termin mitgeschrieben haben. Anfängliche Server-Probleme, die für die Hauptstudie allesamt behoben werden konnten, haben zeitweise die Bearbeitung stark eingeschränkt. Dies scheint einige Studierende ausreichend demotiviert zu haben, die freiwillige semesterbegleitende Bearbeitung der Online-Lernaufgaben

²¹ Detaillierte Skalendokumentation siehe Anhang 1

zu quittieren. An der herkömmlichen Präsenzübung haben diese Studierenden dennoch teilgenommen. Die Zugehörigkeit zur Interventionsgruppe (IG1-IG4) wurde nachträglich an die Regel geknüpft, dass mindestens fünf der zehn zu diesem Zeitpunkt bestehenden Aufgabensets mit mindestens zwei Aufgaben bearbeitet worden sein mussten. Probanden, die diese Bedingungen nicht erfüllten, wurden der Kontrollgruppe zugeschrieben. 26 Probanden aus der KG schrieben die Klausur am ersten und/ oder zweiten Termin mit und lieferten daher eine Grundlage für einen Treatment-Check. Vergleicht man die KG hinsichtlich des Klausurergebnisses (50 Punkte nötig zum Bestehen) mit denjenigen, die die Online-Aufgaben hinreichend bearbeitet haben, so stellt man fest, dass die KG ($M = 33.36$, $SD = 20.09$) im Mittel deutlich schlechter abschneidet als die IG ($M = 51.75$, $SD = 21.33$), die im Mittel die Klausur bestanden hat ($t(71) = -3.69$, $p \leq .001$, $d = 0.88$). Einschränkend muss man hier erwähnen, dass die Probanden der KG, die die Klausur mitgeschrieben haben, ein signifikant schlechteres Chemie Fachwissen zum Pre-Messzeitpunkt aufwiesen ($t(59.044) = -2.62$, $p \leq .05$).

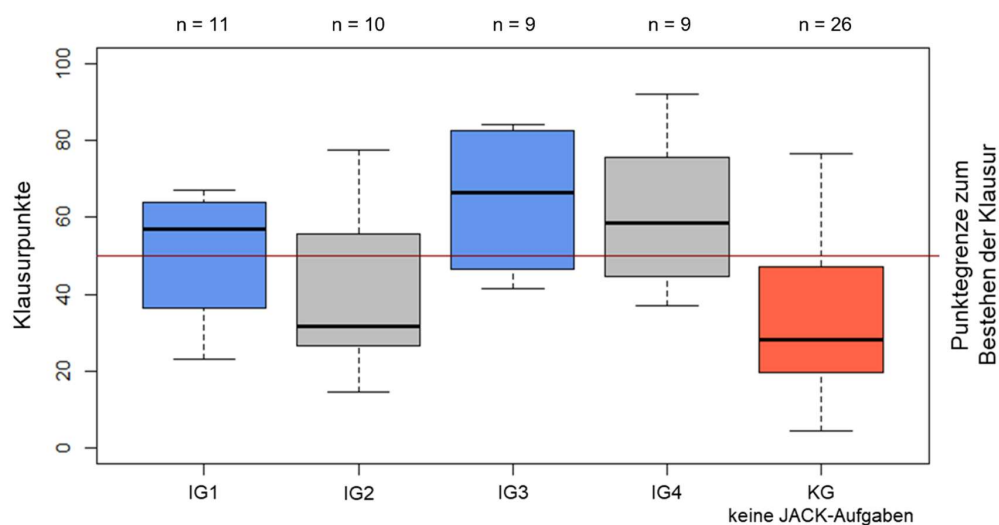


Abbildung 21: Klausurpunkte je Interventionsgruppe, Pilotierung WiSe '18

Des Weiteren lässt sich über die gesamte Stichprobe von Pre ($M = 0.01$, $SD = 0.67$) zu Post ($M = 0.78$, $SD = 0.87$) in einem t-Test ein signifikanter Wissenszuwachs in der Allgemeinen Chemie ($t(102.88) = 5.30$, $p \leq .001$, $d = 1.00$) finden. Mangels Stichprobengröße ($n_{(LK)} = 17$, $n_{(kein LK)} = 22$) für aktive JACK®-Nutzer/-innen ließ sich in weitergehenden Analysen kein Vorteil eines der beiden Algorithmen ausmachen, der über den deskriptiven Status hinausging (siehe Abbildung 22).

Kapitel 8: Ergebnisse der durchgeführten Studien

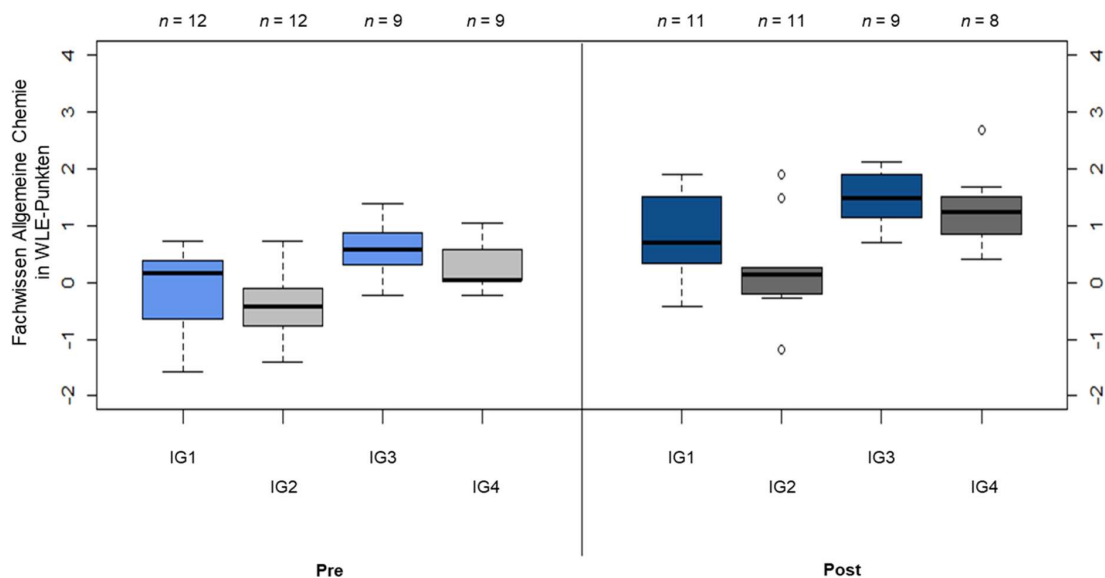


Abbildung 22: Entwicklung des Chemiewissens nach Interventionsgruppe Pilotierung WiSe '18

Eine Poweranalyse mit G*Power (Faul, Erdfelder, Lang & Buchner, 2007) ermittelt zum Nachweis mittlerer Effektstärken ($f^2 = 0.15$, $\alpha = 0.05$, $Power = 0.80$) inklusive Korrektur für hierarchische Datenstrukturen eine notwendige Stichprobe von 42 Probanden pro Gruppe (LK/kein LK), also eine minimale Gesamtstichprobengröße von $N = 94$ Probanden, die es im Rahmen der Hauptstudie anzustreben galt.

8.1.2 Implikationen für die Hauptstudie

Durch den regelmäßigen informellen Austausch mit den Studierenden während der Bearbeitungsphase der Aufgaben konnte ein Aspekt des EST-Feedback-Algorithmus, der im Rahmen einer vorangegangenen Masterarbeit bereits angeschnitten wurde (Trauten, 2018), weitergehend beleuchtet werden: Die Chemie im ersten Semester hält teilweise sehr komplexe Aufgaben für die Studierende bereit. Das führte im Rahmen der Pilotstudie dazu, dass sich einige Studierende vollkommen ohne Anhaltspunkte einem Fehler gegenübergestellt sahen. Da aber ohne eine Antwort grundsätzlich kein Weiterkommen im Algorithmus möglich ist, sahen sich die Lernenden teilweise dazu genötigt, irgendetwas zu verbessern, was unvermeidbar auch zu willkürlichen Antwortmodifikationen führte. Anstelle also zum ersten Fehler eine Rückmeldung zu bekommen, erfolgte diese zu einem wohlmöglich willkürlichen Fehler. Aus diesem Grund wurde der Algorithmus modifiziert und (siehe Abbildung 23).

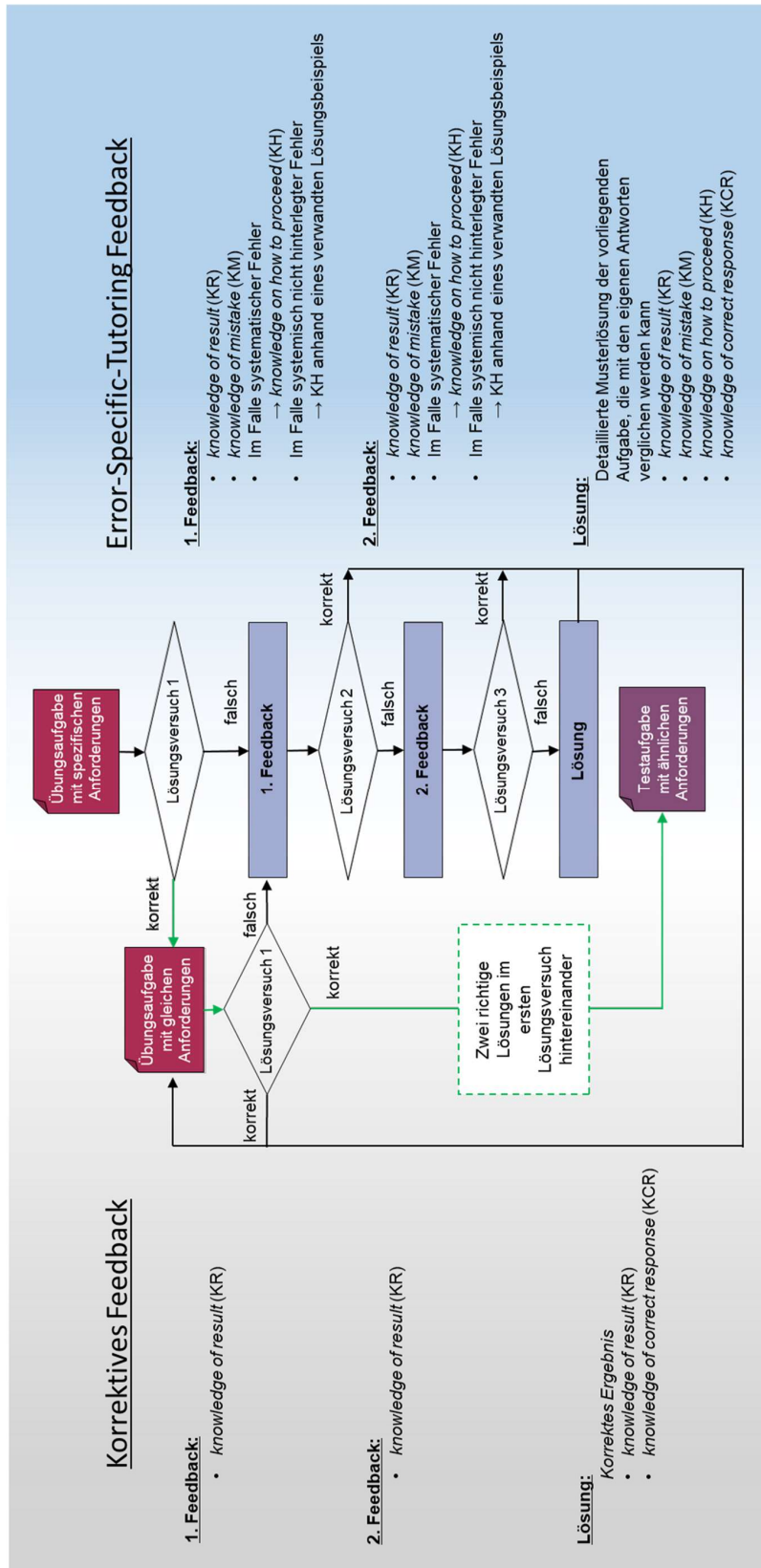


Abbildung 23: Gegenüberstellung des korrekativen Feedback Algorithmus mit dem Error-Specific-Tutoring-Feedback Algorithmus (Trauten et al., 2023)

Kapitel 8: Ergebnisse der durchgeführten Studien

Im überarbeiteten EST-Feedback-Algorithmus wurde nach dem ersten Lösungsversuch fehlerspezifisches elaboriertes Feedback statt korrektivem Feedback präsentiert, um Fehler zu adressieren, die unmittelbar im ersten Lösungsversuch entstanden sind. Damit wurde den Studierenden zwar die Möglichkeit genommen, selbst Fehler zu erkennen und diese zu korrigieren (Schuhmacher, 2008; O. Chen et al., 2017). Ihnen wurde aber auch die Chance gegeben, zielgerichteter falsche Antworten zu überarbeiten. Letzteres Vorgehen erschien auf Basis der Antworten der Studierenden aus der Pilotstudie sinnvoller, denn das korrektive Feedback hatte mehrheitlich dazu geführt, dass Fehler nicht erkannt und berichtigt wurden, bzw. dass andere Fehler hinzukamen.

8.2 Hauptstudie

Für die Hauptstudie im WiSe 2019/2020 (HS 1) wurden mit dem Demografie-Fragebogen N = 92 Probanden erfasst. Regressionsanalysen für die Sub-Gruppen (IG1-IG4) ergaben einen signifikanten Unterschied für mathematische Fähigkeiten ($F(3, 85) = 3.787, p \leq 0.05.; R^2 = 0.087$) mit einem Vorteil für Studierende, die in der Schule einen Leistungskurs belegt haben ($\beta = 0.395, p \leq 0.01$). Des Weiteren ließ sich erwartungskonform ein signifikanter Vorteil im Fachwissen Allgemeine Chemie zum Pre-Messzeitpunkt ($\beta = 0.530, p \leq 0.001$) für Studierende aus IG3 und IG4 ($F(3, 79) = 7.558, p \leq 0.01.; R^2 = 0.152$). Ebenfalls zeigt sich ein signifikanter Unterschied hinsichtlich der am Pre-Messzeitpunkt abgefragten figuralen Skala des Tests für kognitive Fähigkeiten ($F(3, 82) = 3.975, p \leq 0.05.; R^2 = 0.095$) mit einem signifikanten Vorteil ebenfalls zugunsten der Studierenden mit Leistungskurswahl im Fach Chemie ($\beta = 0.373, p \leq 0.05$).

Für die Erweiterungskohorte (HS2; WiSe 2020/2021) zeigt sich bei den kognitiven Fähigkeiten kein signifikanter Vorteil für einen Leistungskurs ($F(3, 41) = 1.426, p = 0.249.; R^2 = 0.0.028$). Ein signifikanter Vorteil im Fachwissen (Pre) für Studierende mit Chemie Leistungskurs ($\beta = 0.604, p \leq 0.001$) findet sich allerdings auch in dieser Kohorte ($F(3, 70) = 7.289, p \leq 0.001.; R^2 = 0.205$).

Die Hauptstudie ist in ihrer Konzeption als Fortführung der Pilotierungsstudie zu sehen. Bis auf geringfügige Änderungen im Algorithmus und in der maximal möglichen Aufgabenmenge sind keine Änderungen zwischen den Jahrgängen vorgenommen worden. Aufgrund der in Tabelle 10 (S. 63) gezeigten Unterschiede hinsichtlich der HS 2 können nicht alle Analysen an allen Probanden durchgeführt werden.

8.2.1 Arbeitsverhalten in der Lernumgebung

Im Vergleich zur Pilotstudie zeigt sich, dass die getroffene Entscheidung, mit dem EST-Algorithmus bereits nach dem ersten Lösungsversuch fehlerspezifisches Feedback anzubieten, sinnvoll gewesen ist. Unter der Voraussetzung, dass eine Aufgabe nicht im ersten Versuch gelöst werden konnte, zeigt sich für das Wintersemester 18/19, dass ein Studierender im Mittel 1,6 der bearbeiteten Aufgaben nach dem ersten Feedback (korrekatives Feedback wie vorgesehen) zur korrekten Lösung bringen und im Mittel nur 1,2 nach dem zweiten Feedback gelöst werden konnte (elaboriertes Feedback, das aber möglicherweise durch willkürliche Verbesserung entstandene Fehler adressiert). Im Wintersemester 19/20 wurde mit dem EST-Algorithmus bereits

Kapitel 8: Ergebnisse der durchgeführten Studien

nach dem ersten Lösungsversuch elaboriertes Feedback angeboten. Studierende waren dadurch in der Lage im Mittel 6,1 der bearbeiteten Aufgaben nach dem ersten – nun fehlerspezifischen – Feedback und 1,8 mithilfe des zweiten fehlerspezifischen Feedbacks im dritten Lösungsversuch eigenständig zu bearbeiten, sodass diese korrekt gelöst wurden. Dieser Trend zeigt sich auch für das WiSe 20/21 (6,3 im zweiten Lösungsversuch und 1,6 im dritten Lösungsversuch).

In den beiden HS-Erhebungen wurden die JACK®-Aufgaben semesterbegleitend eingesetzt. Auffällig ist hier, dass sich bereits nach den ersten drei Wochen im Semester ein erheblicher Schwund bei der aktiven Bearbeitung der Aufgaben zeigt (siehe Abbildung 24). Dieser Schwund, der sich auch in den Reihen der Präsenzübung bzw. den Teilnahmezahlen der Zoom-Übungssitzungen bemerkbar machte, sei aber laut Aussage der Dozierenden „normal“ und „typisch“ für die ersten beiden Wochen.

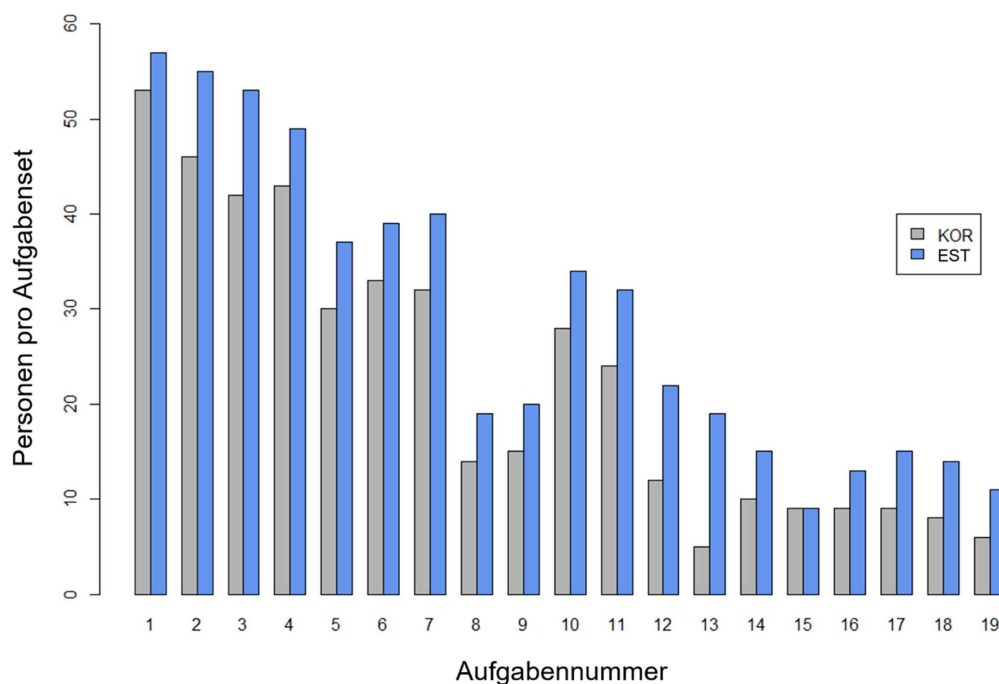


Abbildung 24: Bearbeitungshäufigkeiten der Aufgabensets pro Gruppe (WiSe '19)

Dieser Trend lässt sich auch für das insgesamt deutlich schlechter besuchte Corona-Semester (WiSe 20) beobachten (siehe Abbildung 25). Hinsichtlich der Bearbeitungshäufigkeiten bei den Aufgabensets lässt sich entlang des Semesters auch kein Unterschied für einen der Algorithmen ermitteln. Nähere Auskunft über den Eindruck, den die Lernenden von den Aufgaben haben, geben die Zwischenerhebungen, die für die Hauptstudie in den Moodle-Kurs eingebaut wurden.

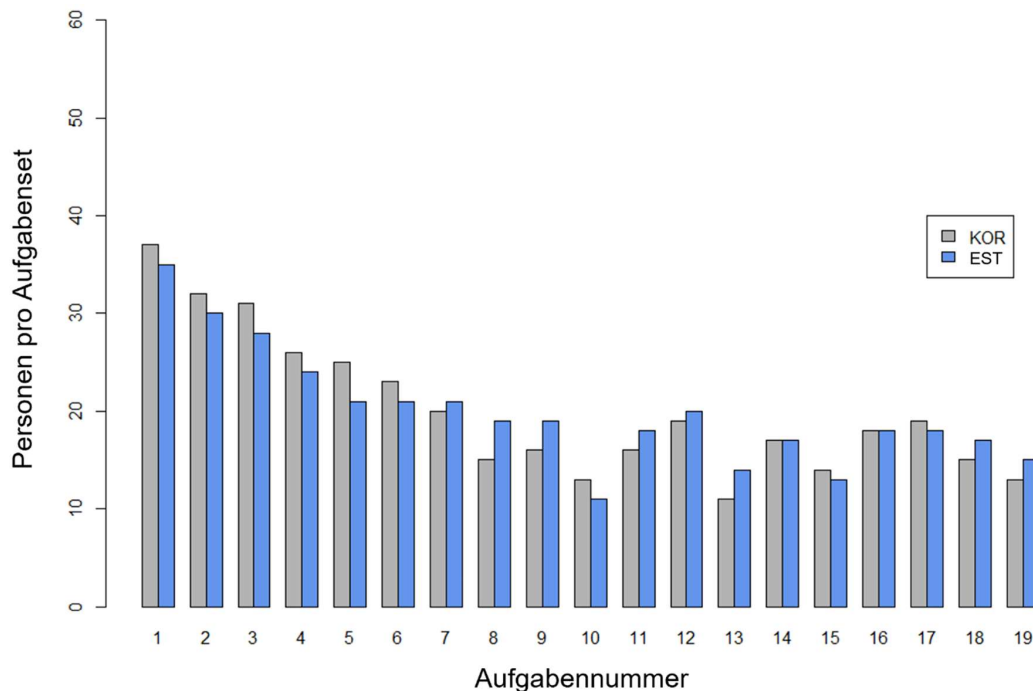


Abbildung 25: Bearbeitungshäufigkeiten der Aufgabensets pro Gruppe (WiSe '20)

Hinsichtlich der Nutzerfreundlichkeit der Aufgaben zeigt sich für beide Interventionsgruppen ein zufriedenstellendes Ergebnis (siehe Abbildung 26), wenngleich auch die Studierenden aus der EST-Gruppe ($M = 2.07$, $SD = 0.9$) die Aufgabenoberfläche signifikant intuitiver einschätzen ($t(1508.2) = 4.129$, $p \leq .001$) als Studierende der KOR-Gruppe ($M = 2.26$, $SD = 0.92$).

Während die Aufgaben seitens der EST-Gruppen ($M = 1.96$, $SD = 0.77$) signifikant leichter empfunden wurden ($t(1468.6) = 2.9435$, $p \leq .01$) als von den KOR-Gruppen ($M = 2.08$, $SD = 0.83$), wurde die kognitive Auslastung durch die Aufgaben bei beiden Gruppen als eher gering eingeschätzt ($M = 2.20$, $SD = 0.82$). Des Weiteren wurde die automatisierte Rückmeldung in den Aufgaben von den Studierenden bewertet, wodurch sich ebenfalls ein deutlicher Vorteil ($t(1457.5) = 12.875$, $p \leq .001$) für die Studierenden in den EST-Gruppen ($M = 2.26$, $SD = 0.98$) ergab. Studierende aus der KOR-Gruppe ($M = 2.94$, $SD = 1.08$) empfanden die korrektive Rückmeldung im Mittel als eher nicht nützlich. Der Anteil an Studierenden, die es als gar nicht nützlich empfanden, ist jedoch recht hoch (siehe Abbildung 27).

Kapitel 8: Ergebnisse der durchgeführten Studien

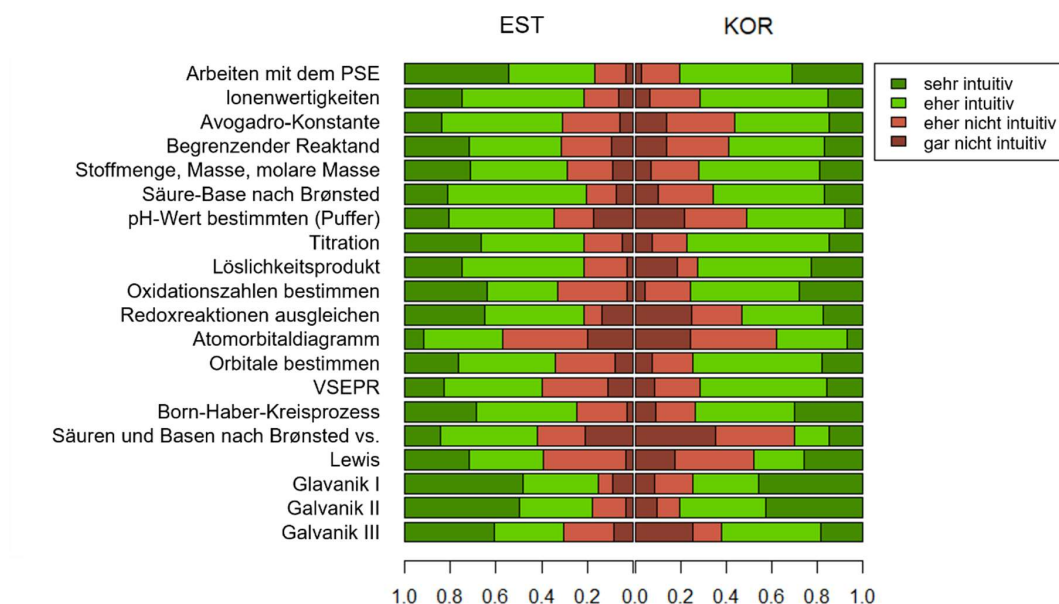


Abbildung 26: Bewertung des Aufgabeninterface je Aufgabe

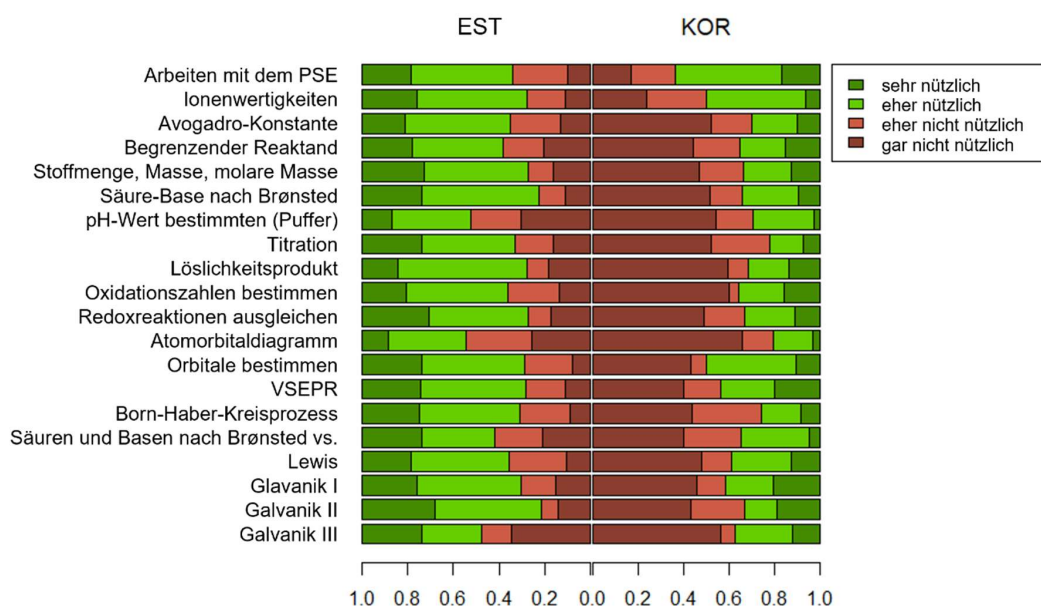


Abbildung 27: Bewertung des Feedback-Nutzens je Aufgabe

Aus deskriptiver Sicht sprechen also bereits einige Aspekte für den EST-Feedback-Algorithmus. Der Frage, ob sich auch hinsichtlich des Studienerfolgs ein Vorteil für EST-Feedback ausmachen lässt, widmen sich die folgenden Abschnitte.

8.2.2 Einfluss des Feedback-Typs auf das Fachwissen (Post)

Zur Beantwortung der Forschungsfrage, in welcher Studienerfolg als Lernstand hinsichtlich des Fachwissens (Post) und als Klausurerfolg operationalisiert wurde, wurden die „aktiven JACK®-Nutzer/-innen“ zunächst bezogen auf das Fachwissen betrachtet, da dieses ein unabhängiges, reliables und standardisiertes Performance-Maß darstellt. Folglich ergaben sich hinsichtlich der COVID-19 Pandemie keine gravierenden Einschränkungen für die so erhobenen Daten, sodass die gesamte Stichprobe ($N_{Post} = 122$) zusammengesetzt aus Studierenden der WiSe 2018, 2019 und 2020 zu Analysezwecken herangezogen werden konnte. Bereits die Verteilung in Abbildung 28 lässt vermuten, dass die Lernaufgaben mit fehlerspezifischem Feedback im Mittel zu einem Angleichen der Wissensstände von Studierenden aus Grund- und Leistungskursen im Laufe des Semesters zu führen scheinen (siehe Abbildung 28).

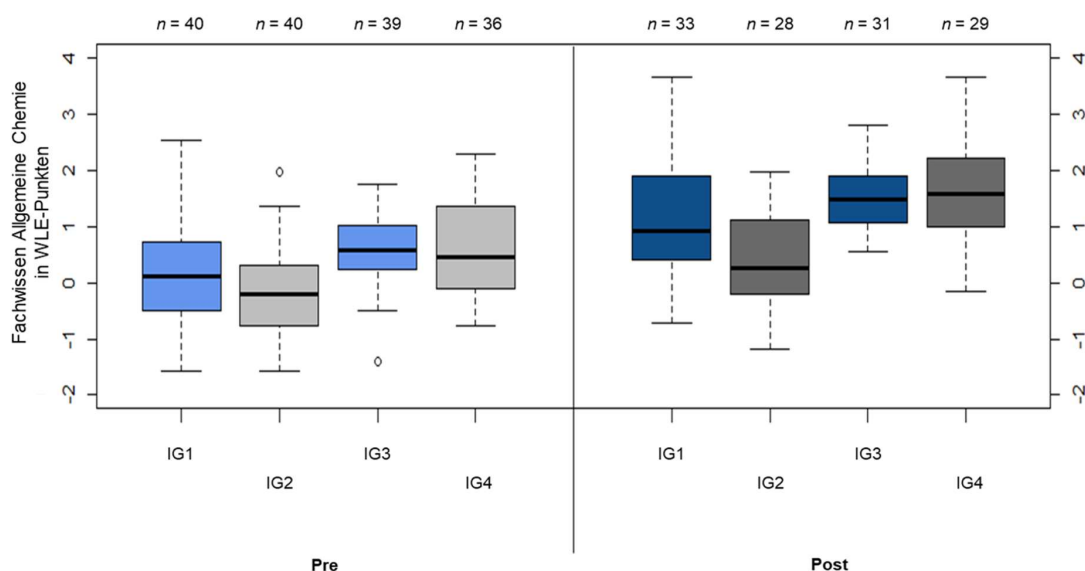


Abbildung 28: Entwicklung des Chemiewissens nach Interventionsgruppe (WiSe '18, '19 und '20)

Kapitel 8: Ergebnisse der durchgeführten Studien

Um zu ermitteln, welchen Einfluss die Interventionsbedingungen auf den Fachwissensstand zum Ende des Semesters haben, wurden multiple lineare Regressionsanalysen unter Kontrolle des Fachwissens bei Studienbeginn durchgeführt. Die Ergebnisse in Tabelle 15 zeigen deutlich, dass die Berücksichtigung der Interaktion aus Kurswahl (0 = kein LK, 1 = LK) und Feedback-Gruppe (0 = KOR, 1 = EST) zu einem signifikant besseren Modell führt ($\Delta R^2 = .014$; $p = .036$) und so 69.4 % Varianzanteil aufklärt.

Laut Datenlage ist es also für den Wissensstand am Ende des Semesters von Vorteil, entweder einen Chemie-Leistungskurs (IG4) in der Oberstufe besucht zu haben oder „aktiv“ mit Aufgaben, die Bug-Related-Tutoring Feedback bereitstellen, gelernt zu haben (IG1). Treffen beide dieser Bedingungen zu, wie z. B. bei der IG3, so zeigt der negative Interaktionsterm, dass der vorteilhafte Einfluss der Feedback-Gruppe verschwindet, da sich der Kurswahl-Effekt und Interaktionseffekt fast vollständig aufheben.

Tabelle 15: Ergebnisse hierarchischer linearer Regressionsanalysen zur Vorhersage des Fachwissens (Post)

Variable	Modell 1			Modell 2		
	B	β	SE	B	β	SE
Konstante	.806 ***		.060	.533 ***		.110
Allg. Ch. Fachw. (pre)	.916 ***	.826	.057	.859 ***	.775	.060
Kurswahl				.430 **	.205	.163
Feedback-Gruppe				.385 *	.183	.152
Feedback-Gruppe x Kurswahl				-.464 *	-.193	.214
korr. R^2	.680			.694		
ΔR^2				.014 *		

Anmerkungen: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

Wie bereits die Boxplots in Abbildung 28 (S. 91) deskriptiv verdeutlichen, zeigt sich auch an den Regressionsgeraden im Streudiagramm in Abbildung 29 sehr deutlich,

dass für Studierende der IG2 das Fachwissen am Ende des Semesters vergleichsweise gering ausfällt.

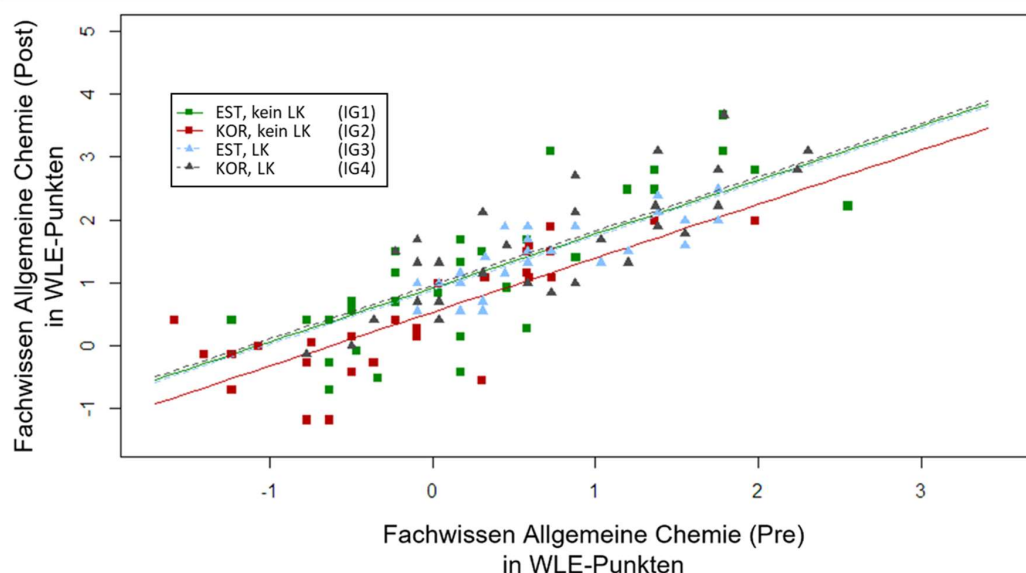


Abbildung 29: Streudiagramm mit gruppenspezifischen Regressionsgeraden des 2. Modells zur Vorhersage des Fachwissens (siehe Tabelle 15)

Auffällig ist, dass eine Regression, die anstelle der in Modell 2 (vgl. Tabelle 15) verwendeten Interaktion aus Feedback-Gruppe und Kurswahl als Prädiktor für Vorwissen, die Interaktion zwischen Feedback-Gruppe und dem Fachwissen zum Beginn des ersten Semesters untersucht, nicht einmal Tendenzen, die auf einen Interaktionseffekt zwischen Vorwissen und Feedbackbedingung hindeuten, enthält. Es findet sich auch für die folgenden Analysen keine Interaktion zwischen dem gemessenen Vorwissen (Fachwissen, Pre) und der Effektivität der beiden Feedback-Typen in Hinblick auf den Studienerfolg. Vielmehr zeigt sich hinsichtlich des gemessenen Vorwissens aus dem Fachwissenstest (Pre) ein davon unabhängiger positiver Effekt für EST-Feedback.

Der Fachwissenstest, der in seiner Konzeption auf Freyer (2013) zurückgeht, wurde zwar hinsichtlich des ESNaS-Modells nach Walpuski et al. (2010) bezüglich der Komplexitätsachse des Aufgabeninhalts gerated, die kognitiven Prozesse konnten aber aufgrund der Konzeption der Testaufgaben nicht erschlossen werden. Der überarbeitete Fachwissenstest von Averbek (2020) besteht zudem zu einem Drittel aus

Kapitel 8: Ergebnisse der durchgeführten Studien

Items, die Faktenwissen und übergeordnete Konzepte und zu zwei Dritteln Zusammenhänge abfragen und von der inhaltlichen Komponente scheinen die 35 Items auch überwiegend Aspekte zu prüfen, die grundlegend für die Allgemeine Chemie sind und sich in den Lehrplänen von Grundkurs und Leistungskurs gleichermaßen wiederfinden.

Lernende, die vormals einen Chemie Leistungskurs besucht haben, starten allerdings mit einem deutlichen Vorteil an Lernzeit in das Studium – diese entspricht der gesamten Lernzeit eines Grundkurses in der Einführungsphase. Dieser Zeitvorteil kann dazu genutzt werden, durch mehr Zeit für kumulatives Lernen ein besseres Wissensnetz aufzubauen. Studierenden, die vormals einen Grundkurs besucht haben, fehlt diese Zeit. Sich im Rahmen eines Tests nicht an Teilaspekte erinnern zu können, sagt vielleicht etwas darüber aus, wie verfügbar und wie sicher jemand weiß – wissen „wie“ oder wissen „dass“ (Vorwissen nach Renkl, 1996) – doch es wird sicherlich einen Unterschied darstellen, ob die Lernenden sich den Lerninhalt durch Anleitung in einer Lernsituation ins Gedächtnis rufen können oder noch nie von diesem Wissensselement gehört haben. Zwar weist der Test gute Kennwerte hinsichtlich der Gütekriterien auf (siehe Kapitel 5), fraglich ist aber, ob er das gesamte Vorwissen abbilden kann, das sich durch die Wahl eines Leistungskurses ergibt.

Die geringere verfügbare Lernzeit bietet im Grundkurs deutlich weniger Möglichkeiten, Zusammenhänge in demselben Umfang wie in einem Leistungskurs zu vermitteln. Das EST-Feedback präsentiert aber gerade diese Zusammenhänge anwendungsbezogen u. a. in Form von Lösungsbeispielen. Da sich diesbezüglich aber nur für Studierende ein signifikanter Vorteil zeigt, die früher einen Grundkurs Chemie belegt haben, ist es möglich, dass die Aufgaben mit EST-Feedback dabei helfen, bei gleicher Lernzeit fehlendes Zusammenhangswissen aufzuholen bzw. vorhandenes (Insel-)Wissen miteinander zu verknüpfen, was sich in einer höheren Test- und Klausurperformance bei IG1 - verglichen mit anderen Studierenden mit Grundkurs in der Oberstufe - zeigt. Es ist also wahrscheinlich, dass *Vorwissen* als latentes Konstrukt nicht vollständig durch den Fachwissenstest abgebildet werden kann, was den zusätzlichen Erklärungswert der Kurswahl in der Oberstufe erklären könnte. Insbesondere die Interaktion zwischen Feedback-Typ und Kurswahl lässt sich bei Verwendung des weiter gefassten Verständnisses von Vorwissen auf Basis der Hypothesen gut erklären.

8.2.3 Lernzuwachs in Abhängigkeit von der Kurswahl

Abbildung 30 betrachtet im Folgenden, wie der Lernzuwachs je Interventionsgruppe ausfällt.

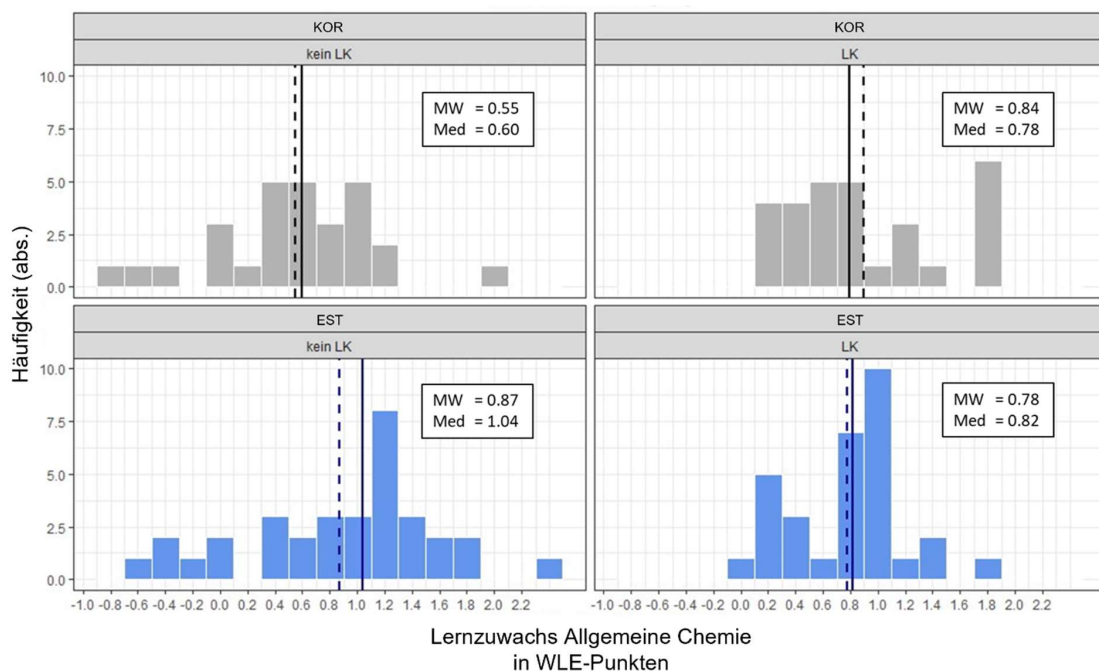


Abbildung 30: Histogramm für den Lernzuwachs nach Interventionsgruppe (WiSe '18, '19 u. '20)

Dabei findet sich die breiteste Streuung der Lernzuwächse beim EST-Feedback von Studierenden, die keinen Leistungskurs besucht haben, jedoch befindet sich ein großer Anteil an Studierenden deutlich oberhalb des Subgruppen-Mittelwerts (Median = 1.04). Die Abweichung des Medians vom Mittelwert ist hier größer als in jeder anderen Gruppe. Der mittlere Lernzuwachs dieser Subgruppe ($M = 0.87$) liegt zudem deutlich über dem Gesamtmittel über alle Gruppen ($M_{\text{Total}} = 0.76$). Bei den Studierenden, die einen Leistungskurs belegt haben und korrekatives Feedback bekommen haben, zeigt sich hingegen eine Verteilung, die nahelegt, dass vermehrt auch weniger als das Subgruppen-Mittel dazugelernt wurde (Median = 0.79 < Mittelwert = 0.87). Bei den anderen beiden Gruppen liegt der Median nahezu auf dem Mittelwert.

Eine Regressionsanalyse zur Vorhersage des Lernzuwachses ($F(5, 86) = 4.955$, $p \leq 0.001$), die zusätzlich zur Feedback-Gruppe (0 = KOR, 1 = EST) und der Kurswahl im Fach Chemie (0 = kein LK, 1 = LK) die Rechenfertigkeit und das Fachwissen

Kapitel 8: Ergebnisse der durchgeführten Studien

(Pre) in der Allgemeinen Chemie berücksichtigt, gibt möglicherweise Aufschluss über potentielle Wirkzusammenhänge (siehe Tabelle 16).

Tabelle 16: Ergebnisse der multiplen linearen Regressionsanalyse zur Vorhersage des Lernzuwachses

Variable	B	β	SE
Konstante	.707 ***		.133
Allg. Ch. Fachw. (pre)	-.242 *	-.298	.103
Rechenfähigkeit	.216 ***	.481	.054
Kurswahl	.373 *	.305	.185
Feedback-Gruppe	.353 *	.289	.172
Feedback-Gruppe x Kurswahl	-.499 *	-.359	.236
korr. R^2	.179		

Anmerkungen: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

Leider wurde im Wintersemester 2020 der Rechentest nicht mit als Kontrollvariable miterhoben, sodass für diese Analyse nur ein $n = 92$ Probanden zur Verfügung stand. Auch wenn das Regressionsmodell in Tabelle 16 die gefundenen Zusammenhänge aus Tabelle 15 (Modell 2) am Lernzuwachs überprüft, muss erwähnt werden, dass dieses Modell nicht signifikant besser als ein sparsames Modell misst, dass lediglich die Rechenfähigkeit und das Fachwissen (Pre) als Prädiktor für den Lernerfolg berücksichtigt ($F(2, 89) = 9.381$, $p \leq 0.001$; $R^2 = .156$). Da sich die Stichprobengröße allerdings noch unter dem laut Power-Analyse notwendigen Umfang (siehe Kapitel 8.1.1), für mittlere Effektstärken bewegt, ist es nicht unwahrscheinlich, dass bei einer größeren Stichprobe der Modellunterschied signifikant geworden wäre. Dies wäre durch weitere Studien zu prüfen.

Es ergibt sich unter Kontrolle des Fachwissens (Pre) für die Rechenfähigkeit in diesem Modell die stärkste prädiktive Kraft auf den Lernzuwachs. Dass die Performance im Mathematiktest auch den Lernzuwachs bedingt, lässt sich mit einem Blick auf den Aufgabeninhalt des Instruments plausibel erklären. Denn der eingesetzte Test zur Rechenfähigkeit prüft überwiegend die in der Schule erlernten Rechenoperationen ab. So finden sich Aufgaben zu den Grundrechenarten und Bruchrechnung in dem

Test. Die Beherrschung dieser Operationen ist essentiell für das Aufstellen von Reaktionsgleichungen. Des Weiteren finden sich auch Aufgaben zu komplexeren Rechenoperationen, wie z. B. das Lösen von Gleichungen mit mindestens einer Unbekannten sowie Exponentialrechnung und Logarithmen, deren Beherrschung z. B. für die pH-Wert Bestimmungen im Rahmen des Säure-Base-Konzepts nach Brønsted oder auch für Reaktionskinetik essentiell sind.

Es bleibt weiterhin ein positiver Einfluss der Kurswahl in der Oberstufe bestehen, der aber durch fehlerspezifisches Feedback kompensiert werden kann. Dass das Vorwissen als Prädiktor für den Lernzuwachs ein negatives Vorzeichen aufweist, weist darauf hin, dass ab einem gewissen Eingangsniveau an Chemiewissen im Rahmen des ersten Semesters nicht weiter dazugelernt werden kann. Darüber hinaus zeigen sich für die Variable Kurswahl und Feedback-Gruppe sowie den Interaktionsterm aus den beiden ein ähnliches Bild wie bei der Regressionsanalyse im vorherigen Kapitel 8.2.2. Hier ist der Interaktionseffekt gerade negativ genug, um den Vorteil des EST-Feedbacks in Kombination mit einer früheren Leistungskurswahl in der Oberstufe aufzuheben. Verrechnet man den Nachteil des Interaktionsterms von Kurswahl und Feedback-Gruppe mit dem Vorteil durch das EST-Feedback, so verbleibt ein Malus von $B = - .146$, die Studierenden mit früherer Leistungskurswahl die mit EST-Feedback gelernt haben, im Vergleich zu Studierenden mit früherer Leistungskurswahl, die KOR-Feedback bekamen. Es handelt sich hierbei allerdings um das komplexeste Regressionsmodell bei gleichzeitig geringster Fallzahl. Wie oben bereits beschrieben ist daher die Aussagekraft dieser Regression beschränkt.

8.2.4 Einfluss des Feedback-Typs auf das Klausurergebnis

Den zweiten Aspekt zur Beantwortung der Forschungsfrage bildet die Klausurleistung der Studierenden. Durch mindestens 50 erzielte Punkte bei einer Teilnahme an mindestens einem der zwei Klausur-Termine, die in der vorlesungsfreien Zeit nach Semesterende innerhalb von vier Wochen zur Verfügung standen, konnte die Klausur bestanden werden. Der erste Klausurtermin diente dabei den Studierenden als Freiversuch. Das heißt, dass die Prüfungen, die am ersten Termin angetreten und nicht bestanden wurden, nicht als Prüfungsversuch gewertet wurden. Ferner konnte trotz bestandener Klausur am ersten Termin der zweite zur Notenverbesserung genutzt werden. Der Klausurerfolg wird daher als die beste Punktzahl angesehen, die über die ersten beiden Klausur-Termine je Proband erreicht wurde.

Kapitel 8: Ergebnisse der durchgeführten Studien

Die Daten aus der zweiten Hauptstudie im Wintersemester 2020/2021 wurden bei der Analyse nicht betrachtet ($N_{Post} = 93$). Aufgrund der Covid-19 Pandemie und den damit verbundenen Veränderungen hinsichtlich des Prüfungsformats waren die Klausurbedingungen der resultierenden Online-Klausur nicht mit den vorherigen Klausuren zu vergleichen. In der Pilotstudie und Hauptstudie 1 wurden an den jeweiligen Klausurterminen jeweils die gleichen Klausurfragen eingesetzt, sodass ein Vergleich zwischen den Jahren möglich ist. Über eine einfaktorielle ANOVA ($F(3, 89) = 5.384$, $p \leq 0.01$; $\eta^2 = 0.154$) konnten für den Klausurerfolg im Post-Hoc Test nach Tukey²² signifikante Unterschiede zwischen den Interventionsgruppen ermittelt werden (siehe Abbildung 31). Dabei blieben Studierende der IG2, die keinen Leistungskurs Chemie in der Oberstufe belegt haben und KOR-Feedback erhielten ($M = 39.36$, $SD = 20.20$), signifikant hinter IG3 und IG4 zurück.

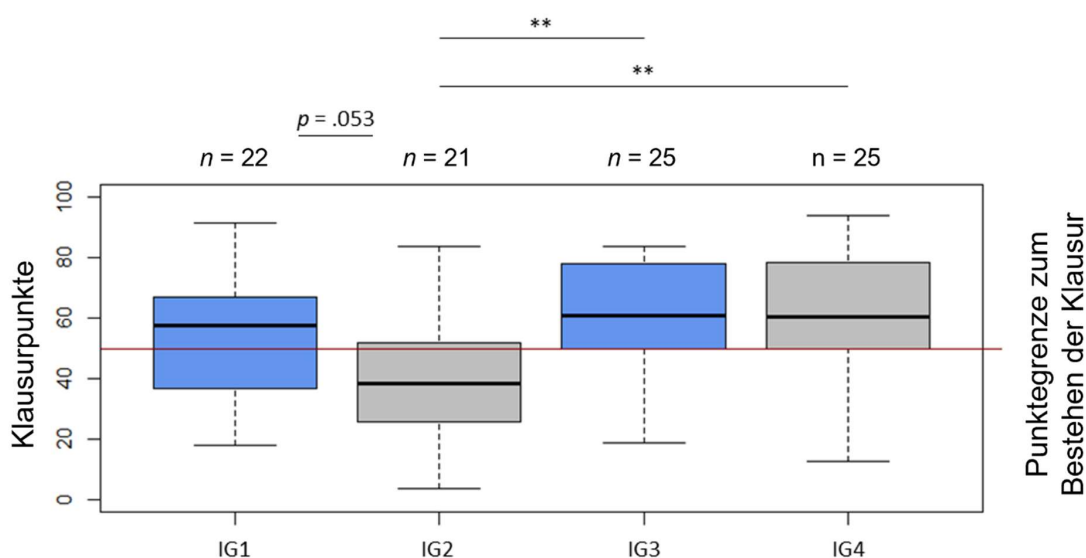


Abbildung 31: Klausurpunkte in Abhängigkeit von der Interventionsgruppe WiSe '18 u. '19 (** $p \leq .01$)

Zwischen IG3 ($M = 60.52$, $SD = 17.44$) und IG4 ($M = 61.00$, $SD = 22.23$) findet sich kein signifikanter Unterschied. Studierende ohne Chemie Leistungskurs in der Oberstufe (IG1) profitieren hinsichtlich des Klausurerfolgs am meisten vom EST-Feed-

²² Die Stichprobenumfänge sind nahezu gleich groß, der Shapiro-Wilk Test weist zudem Varianzhomogenität hin. Daher wurde hier die Post-Hoc Methode nach Tukey gewählt (Field, 2018).

back ($M = 55.61$, $SD = 21.98$), da sie nicht mehr von den Leistungen der Studierenden aus IG3 oder IG4 unterscheidbar sind und sich zudem ein Trend hinsichtlich eines Vorsprungs zu Studierenden der IG2, ebenfalls ohne Chemie Leistungskurs in der Oberstufe, abzeichnet. Prüft man die Wirkzusammenhänge, die sich in Modell 2 (Tabelle 15) zeigen, hinsichtlich des Klausurerfolgs, so zeigen sich in der entsprechenden Regression ($F(4, 88) = 14.63$, $p \leq 0.001$) ähnliche Zusammenhänge (siehe Tabelle 17). Für die Analyse wurden die Daten der „aktiven“ JACK®-Nutzer/-innen der Jahrgänge 2018 und 2019 berücksichtigt. Wie in Modell 2 (Tabelle 15) wurden das Fachwissen (pre), die Kurswahl im Fach Chemie (0 = kein LK, 1 = LK) und die Feedback-Gruppe (0 = KOR, 1 = EST) als Prädiktoren in das Regressionsmodell aufgenommen. Zwar hat das Vorwissen erwartungskonform weiterhin den größten Einfluss auf den Klausurerfolg, jedoch ergibt sich ein signifikanter Vorteil für das EST-Feedback, während in diesem Modell die Kurswahl als signifikanter Einflussfaktor verschwindet. Bei einer größeren Stichprobe sind allerdings ähnliche Zusammenhänge zu vermuten wie in Tabelle 15 hinsichtlich des Fachwissens (Post) erhalten wurden. So deutet der nicht signifikante Interaktionsterm darauf hin, dass sich unter Kontrolle des Vorwissens bei einer größeren Stichprobe nur ein Vorteil des EST-Feedbacks für Studierende ergibt, die keinen Leistungskurs in der Oberstufe gehabt haben. So ergibt sich unter Kontrolle des Vorwissens eine Datenlage, die in eine ähnliche Richtung deutet, wie es sich schon in der ANOVA und in Abbildung 31 abzeichnete.

Tabelle 17: Ergebnisse der multiplen linearen Regressionsanalyse zur Vorhersage des Klausurerfolgs

Variable	B	β	SE
Konstante	44.561 ***		3.892
Allg. Ch. Fachw. (pre)	15.973 ***	.560	2.662
Kurswahl	9.052	.207	5.557
Feedback-Gruppe	11.608 *	.266	5.360
Feedback-Gruppe x Kurswahl	-13.401	-.272	7.254
korr. R^2	.372		

Anmerkungen: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

Bedingt durch Änderungen in der Klausur im Jahr 2021 kann ein Teil der Probanden für diese Evaluation nicht genutzt werden. Die Fallzahl der Probanden fällt daher deutlich geringer aus als ursprünglich geplant. Dass die Kurswahl als in der Analyse eingeschlossener Prädiktor keinen signifikanten Einfluss auf das Klausurergebnis hat und in zweiter Instanz auch die erwartete Interaktion zwischen Feedback-Gruppe und Kurswahl ausbleibt, kann auch an der mit $n = 93$ Probanden zu geringen Fallzahl liegen, denn deskriptiv liegen IG1, die Gruppe ohne Leistungskurs mit EST-Feedback, gleich auf mit beiden Leistungskursgruppen. Der Effekt der Feedback-Gruppe erweist sich für das Klausurergebnis allerdings als resistent genug auch bei geringer Fallzahl signifikant zu bleiben.

8.2.5 Kohortenvergleich mit der Vergleichskohorte aus ALSTER I

Um final zu untersuchen, ob durch das neue Übungsformat mit automatisiertem Feedback in digitalen Lernaufgaben signifikante Leistungsverbesserungen im Fachwissenstest (Post) am Ende des Semesters erhalten werden können, wurden die Fachwissensdaten der Studierenden des WiSe 2016/2017 ($N = 105$) aus der ersten Förderphase der DFG-Forschergruppe ALSTER (FOR 2242) für einen Kohortenvergleich herangezogen. Die Kontrollkohorte erweist sich hinsichtlich Alter und Verteilung in der Gender-Variable als vergleichbar ($M_{\text{Alter}} = 20.43$ Jahre, 58.9 % männlich, 41.1 % weiblich) und es lässt sich kein signifikanter Unterschied zwischen den Subgruppen mit gleicher Kurswahl hinsichtlich des Vorwissens in der Allgemeinen Chemie ($\text{ANOVA}_{(\text{kein LK})}: F(2, 112) = 1.152, p = 0.320$; $\text{ANOVA}_{(\text{LK})}: F(2, 73) = 0.127, p = 0.881$) und der mathematischen Fähigkeiten ($\text{ANOVA}_{(\text{kein LK})}: F(2, 112) = 1.611, p = 0.204$; $\text{ANOVA}_{(\text{LK})}: F(2, 73) = 0.152, p = 0.860$) zwischen Kontrollkohorte und Interventionskohorte feststellen. Der Einfluss der Interventionsbedingungen bei sonst gleicher Lehre auf den Fachwissensstand am Ende des Semesters wurde mithilfe einer multiplen linearen Regression untersucht ($F(4, 224) = 106.8, p \leq 0.001$). Neben dem Fachwissen Chemie zu Beginn des Semesters und der Kurswahl in der Oberstufe wurden zwei weitere Variablen hinzugezogen. Durch zwei Dummy-Variablen EST-Feedback (0 = nein, 1 = ja) und KOR-Feedback (0 = nein, 1 = ja) wird der Vergleich beider Feedback-Typen mit der Kontrollkohorte ermöglicht. Zusätzlich zu der erwartungskonform hohen prädiktiven Kraft des Fachwissens zu Beginn des Semesters ($\beta = .740, p \leq 0.001$) und ebenfalls guten Vorhersagewerten für die Kurswahl ($\beta = .103, p \leq 0.05$) trägt der Prädiktor, ob EST-Feedback erhalten wurde ($\beta = .107, p \leq 0.05$), signifikant zur Varianzaufklärung bei. Ob KOR-Feedback erhalten wurde

oder nicht ist hingegen unerheblich für den Fachwissensstand am Ende des Semesters. Für diese Regression lässt sich daher festhalten, dass Studierende, die EST-Feedback erhalten haben, im Vergleich mit der Kontrollgruppe signifikant profitieren. Dass unabhängig von der Kurswahl profitiert wird, wie das Regressionsmodell möglicherweise suggeriert, ist vorsichtig zu betrachten, denn innerhalb der Intervention konnte bereits gezeigt werden, dass es eine signifikante Interaktion für EST-Feedback als Treatment-Bedingung und die Kurswahl gibt. Ob die Studierenden unabhängig von ihrer Kurswahl vom EST-Feedback profitieren, wurde in einem weiteren Modell überprüft. Das zweite Modell (siehe Tabelle 18), welches die Interaktion aus Feedback und Kurswahl in der Oberstufe berücksichtigt, zeigt wie erwartet einen signifikanten negativen Effekt für die Interaktion aus EST-Feedback und Kurswahl ($\beta = -.134$, $p \leq .05$). Da dieses Modell nicht signifikant besser misst als Modell 1 ($\Delta R^2 = .004$, $p = .089$) und der Interaktionseffekt von KOR-Feedback und Kurswahl in der Oberstufe mit $\beta = .004$ ($p = .170$) vernachlässigbar gering ausfällt, wurde ein weiteres Modell mit einem Interaktionsterm aus EST-Feedback und der Chemie Kurswahl berechnet (vgl. Modell 3, Tabelle 18). Es zeigt sich eine signifikante Verbesserung der Varianzaufklärung im Vergleich zu Modell 1 bei sonst gleichbleibenden Wirkzusammenhängen ($\Delta R^2 = .01$, $p \leq .001$). Der Vorteil von EST-Feedback ($\beta = .184$, $p \leq .01$) wird durch den Interaktionsterm nahezu aufgehoben ($\beta = -.136$, $p \leq .05$).

Wurde in der Oberstufe ein Leistungskurs besucht, wird der Vorteil des EST-Feedbacks durch den Interaktionsterm aus Kurswahl und EST-Feedback nahezu aufgehoben. Dass folglich sich nur ein Vorteil für diejenigen Studierenden ergibt, die keinen Chemie Leistungskurs in der Oberstufe besucht haben, zeigt sich auch in Abbildung 32 (S. 103) deskriptiv. Studierende ohne Leistungskurs, die aber EST-Feedback bekommen haben ($M = 1.18$, $SD = 1.16$), sind bzgl. des Fachwissens am Ende des Semesters nahezu gleichauf mit den Studierenden der Kontrollgruppe, die einen Leistungskurs besucht haben ($M = 1.28$, $SD = 1.23$). Die entsprechende ANOVA, die den Fachwissensstand zum Ende des Semesters zwischen den Interventionsgruppen vergleicht ($F(5, 223) = 11.26$, $p \leq 0.001$) zeigt im post-hoc Test (Games-Howell²³) keinen signifikanten Unterschied zwischen der Interventionsgruppe ohne Leistungskurs mit EST-Feedback zu allen Subgruppen, die einen Leistungskurs in der

²³ Der Games-Howell Post Hoc Test wird dann genutzt, wenn Varianzhomogenität nicht gegeben ist. Zudem liefert dieser Post Hoc Test auch zuverlässige Ergebnisse, wenn die Größen der zu vergleichenden Gruppen voneinander abweichen (Field, 2018).

Kapitel 8: Ergebnisse der durchgeführten Studien

Oberstufe belegt haben und ist zudem signifikant besser ($p \leq 0.01$) als die Kontrollgruppe ohne Leistungskurs ($M = 0.30$, $SD = 1.23$). Im Vergleich zur Interventionsgruppe ohne Leistungskurs, die KOR-Feedback bekommen (IG2: $M = 0.45$, $SD = 0.92$) hat, ist auch hier wieder nur ein möglicher Trend zu erkennen ($p = 0.082$).

Tabelle 18: Multiple lineare Regressionen zur Vorhersage des Fachwissens (Post), Kohortenvergleich

Modell	Modell 1			Modell 2			Modell 3		
Variable	B	β	SE	B	β	SE	B	β	SE
Konstante	.453 ***		.076	.410 ***		.083	.408 ***		.078
Allg. Ch. Fachw. (pre)	.930 ***	.734	.055	.920 ***	.725	.055	.920 ***	.730	.055
Kurswahl	.261 *	.107	.105	.408 *	.167	.158	.414 **	.169	.125
EST-Feedback	.285 *	.107	.116	.489 **	.184	.151	.491 **	.184	.148
KOR-Feedback	.165	.060	.120	.129	.047	.159	.135	.049	.120
Kurswahl x EST-Feedback				-.469 *	-.134	.235	-.474 *	-.136	.214
Kurswahl x KOR-Feedback				.015	.004	.242			
korr. R^2	.646			.650			.656		
ΔR^2				.004			.01 *		

Anmerkungen: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

Verglichen mit einem klassischen Übungsbetrieb (vgl. $KG_{(\text{kein LK})}$ und $KG_{(\text{LK})}$) bietet es sich also an, diese Lernaufgaben mit EST-Feedback als zusätzliches Lernangebot einzusetzen, da aktives Lernen mit den Aufgaben zu einem signifikant besseren Fachwissensstand am Ende des ersten Semesters führt. Um von wirklichen Nachteilen für Studierende mit einem Leistungskurs in der Oberstufe sprechen zu können, wird die Interaktion zwischen Feedback (EST) und Kurswahl nicht negativ genug. Der Vorteil des Terms „EST-Feedback“ müsste mehr als nur aufgehoben werden. Dass der Interaktionsterm Feedback (EST) und Kurswahl den Term „EST-Feedback“ nur nahezu aufhebt, lässt darauf schließen, dass Studierende mit einem Leistungskurs in der Oberstufe weder Vor- noch Nachteile aus dieser Interventionsmaßnahme erhalten.

Kapitel 8: Ergebnisse der durchgeführten Studien

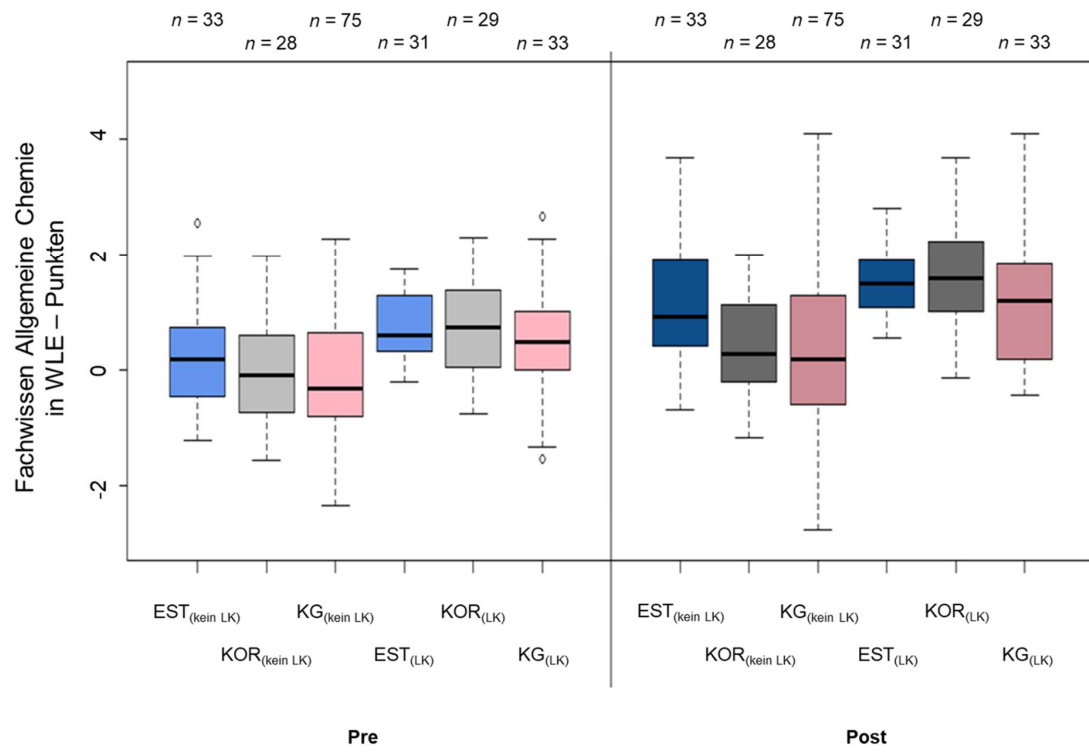


Abbildung 32: Entwicklung des Chemie Fachwissens nach Interventionsgruppe (Kohortenvergleich)

Kapitel 9: Zusammenfassung und Diskussion

Für die hohen Abbruchquoten in Chemiestudiengängen in Deutschland ließen sich zwei korrespondierende Problembereiche herleiten. Das konstatierte Performance Problem leitet sich vor allem aus den Leistungsproblemen in Chemiestudiengängen ab. Dies ist der meist genannte und für viele befragte Studienabbrechende auch der entscheidende Grund für eine Abbruchsentscheidung (Heublein et al., 2017; Heublein et al., 2020). Das Feedback-Problem lässt sich unter anderem aus dem Student-Integration Modell (Tinto, 1988) herleiten. Weil die Studieneingangsphase neben der Notwendigkeit besagte Leistungsprobleme zu überwinden zugleich auch von Integrationsprozessen in dieses neue Sozietop Hochschule geprägt ist, steht für die Überwindung fachlicher Defizite zu Beginn des ersten Semesters eben nicht die gesamte Bandbreite an sozialen Ressourcen zur Verfügung. Die Kontaktaufnahme zu anderen Studierenden und Dozierenden wird damit zu einem entscheidenden Faktor für den Verbleib an der Hochschule und den Studienerfolg (Heublein et al., 2017; Sarcletti & Müller, 2011). Gleichmaßen ist diese Kontaktaufnahme eine notwendige Voraussetzung, um in Interaktionssituationen Feedback zu erhalten. Bleibt bereits die Kontaktaufnahme aus, ergibt sich auch ein potentieller Mangel an Feedback-Situationen. Da Feedback als einer der stärksten Einflussfaktoren auf Lern- und Motivationsprozesse identifiziert werden konnte (Hattie & Timperley, 2007), ist in so einem Fall mit negativen Einfluss auf den Studienerfolg zu rechnen. Da diese Ressource aber als Konsequenz erst anlaufender Sozialisierungsprozesse zu Studienbeginn bestenfalls sehr eingeschränkt zur Verfügung steht, wird hier von einem potentiellen Mangel und damit von einem Feedback-Problem ausgegangen.

9.1 Zusammenfassung der Ergebnisse

In der hier vorgestellten Studie wurde den Erstsemesterstudierenden daher eine personenunabhängige Feedbackquelle zur Verfügung gestellt und dabei zugleich der Einfluss des Vorwissens auf die Wirksamkeit unterschiedlicher Feedbackkomponenten untersucht. Denn trotz des Wissens um die enorme Wirkmächtigkeit von Feedback zeigen Studien, die den Einfluss von Vorwissen auf die Wirksamkeit unterschiedlicher Feedback-Komponenten untersuchen, ein uneindeutiges und zum Teil auch kontroverses Bild von Wirkzusammenhängen. Narciss (2006) liefert in diesem Zusammenhang einige Modelle, die die Wirkweise und den Informationsgehalt von Feedback beschreiben, und definiert Feedback-Komponenten nach inhaltlichen Gesichtspunkten (s.

Tabelle 3, S. 28), sodass auf Basis dieser Klassifikation mit einheitlicher Nomenklatur Feedback-Typen miteinander verglichen werden können. Daran angelehnte Studien lassen aber designbedingt keine eindeutigen Schlüsse zu (Narciss & Huth, 2004). Weitere Studien, die ähnliche Zusammenhänge überprüfen, differenzieren die untersuchten Feedback-Komponenten entweder nicht sauber voneinander (Mayer & Johnson, 2010; Moreno, 2004) oder weisen designbedingte Schwächen auf, die den Ergebnissen nur eingeschränkte Gültigkeit verleihen. Dies erklärt die große Varianz und teilweise widersprüchliche Effektstärken aber nur teilweise, da sich in allen Metastudien immer wieder widersprüchlichen Effektstärken finden (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Wisniewski et al., 2019). Um diese fortbestehende Forschungslücke weiter schließen zu können, wurde sich im Rahmen dieser Arbeit mithilfe einer Studie im Pre-Post-Vergleichsgruppendesign der Beantwortung der folgenden Forschungsfrage gewidmet:

FF: Wie wirken sich die beiden gegensätzlichen Feedback-Maßnahmen (Error-Specific-Tutoring-Feedback vs. KORrektives-Feedback) in Abhängigkeit vom fachspezifischen Vorwissen auf den Studienerfolg von Erstsemesterstudierenden in der Allgemeinen Chemie aus?

Vor dem Hintergrund der lernförderlichen Wirkung mehrstufiger fehlerspezifischer ITF-Algorithmen in Studien zum Kompetenzerwerb im Bereich Physik und Mathematik (z. B. Albacete & VanLehn, 2000; Narciss & Huth, 2004) wurde in Kapitel 4 vermutet, dass ein EST-Feedback-Algorithmus für Novizen bzw. vorwissensschwache Studierende zu einem höheren Studienerfolg führen würde.

H1: Studierende mit geringem fachspezifischem Vorwissen profitieren in ihrem Studienerfolg stärker durch ein Training mit einem EST-Feedback als Studierende mit hohem Vorwissen.

Es konnte über alle Regressionsanalysen, sowohl im direkten Vergleich der Interventionsgruppen (WiSe 2018-2020), als auch im Vergleich mit der Kontrollkohorte aus ALSTER I (WiSe 2016) ein signifikanter Vorteil für EST-Feedback nachgewiesen werden (vgl. Tabelle 15, Tabelle 16, Tabelle 17, Tabelle 18). Ein direkter Zusammenhang zwischen Vorwissen, welches mit dem Vorwissenstest (Averbeck, 2020) erhoben wurde, und Interventionsbedingung konnte am vorliegenden Datensatz über $N = 122$ Studierende und auch im Vergleich mit der Kontrollkohorte nicht nachgewiesen werden.

Anhand der Vorwissensdefinition von Renkl (1996) und in Anbetracht des Vorteils an Lernzeit, mit dem eine Leistungskurswahl in der Oberstufe einher geht, wurde dargelegt, dass *Vorwissen* als latentes Konstrukt vermutlich nicht vollständig durch den Fachwissenstest abgebildet werden kann, sodass die Kurswahl in der Oberstufe als zusätzlicher Indikator für Vorwissen mit einbezogen wird. Neben dem Erhalt von EST-Feedback zeigt sich unter Kontrolle des Vorwissens ein positiver Einfluss für den Faktor Kurswahl in der Oberstufe, wenn ein Leistungskurs belegt wurde. Der positive Einfluss des EST-Feedbacks lässt sich sowohl hinsichtlich des Klausurerfolgs als auch in Bezug auf den Fachwissensstand am Ende des Semesters ausmachen. Für das Fachwissen (Post) und den Fachwissenszuwachs zeigt sich aber, dass, wenn EST-Feedback erhalten *und* in der Oberstufe ein Leistungskurs besucht wurde, sich dieser zusätzliche Vorteil aufgrund der Interaktion aus Feedback-Bedingung und Kurswahlverhalten aufhebt, wie man an Abbildung 29 (S. 93) sehen kann.

Die hierarchischen Regressionen zur Vorhersage des Fachwissensstands am Ende des Semesters (Tabelle 18, S. 102) zeigen – insbesondere der Vergleich mit der Kontrollkohorte – noch einmal differenzierter, dass der Interaktionseffekt hauptsächlich auf die Interaktion zwischen EST-Feedback und Chemie-Leistungskurs zurückzuführen ist. Der erste Interaktionsterm (Kurswahl x EST-Feedback; siehe Tabelle 18, Modell 2) zeigt, dass der Vorteil für EST-Feedback nur für Studierende ohne Leistungskurs (IG1) eintritt. Wurde hingegen ein Leistungskurs Chemie in der Oberstufe belegt und mit EST-Feedback gelernt (IG3), so äußert sich in der Interaktion ein schwacher Expertise Reversal Effekt, der den Vorteil des EST-Feedbacks nahezu aufhebt. Unter der Annahme, dass das Konstrukt Vorwissen als Kombination aus dem Fachwissenstest (Pre) und der Chemie-Kurswahl in der Oberstufe beschrieben werden kann, lässt sich Hypothese 1 also bestätigen.

Bei zunehmender Expertise wurde hingegen erwartet, dass der Nutzen von elaborem Feedback sinkt bzw. sich umkehrt (Expertise Reversal Effect, siehe Kapitel 3.2; Kalyuga, 2014), sodass für Experten auch Nachteile durch zu umfangreiches Feedback entstehen können (z. B. Smits et al., 2008) und somit KOR-Feedback zu einer besseren Performance am Ende des Semesters führt.

H2: Studierende mit hohem fachspezifischem Vorwissen profitieren in ihrem Fachwissenserwerb bzw. Studienerfolg stärker durch ein Training mit einem KOR-Feedback als Studierende mit geringem Vorwissen.

Auf Basis dessen war für IG4 ein Vorteil für das Lernen mit KOR-Feedback erwartet worden. Der Interaktionsterm aus Leistungskurswahl und KOR-Feedback (Kurswahl x KOR-Feedback; siehe Tabelle 18, Modell 2) weist zwar konform zu Hypothese 2 ein positives Vorzeichen auf, doch ist er weder signifikant noch von nennenswerter Effektstärke ($\beta = .004$ $p = .170$)

Wenngleich sich ein schwindender Vorteil der Unterstützung in der Interaktion aus EST-Feedback und Leistungskurswahl für IG3 zeigt, lassen sich keine Indizien für Hypothese 2 finden, weshalb sie verworfen werden muss. Die sehr komplexen Aufgaben, die die Allgemeine Chemie bereithält, können allerdings der Grund sein, warum Studierende der IG 3 keine Nachteile durch Lernen mit dem EST-Feedback erfahren, sondern sich nur ein signifikanter Rückgang des Vorteils im Interaktionsterm zeigt. Damit lässt sich angelehnt an Hypothese 2 immerhin ein milder Expertise Reversal Effekt finden, da vermutet wird, dass für Studierende mit Leistungskurs Chemie in der Oberstufe und damit verbunden formal ein Schuljahr mehr Lernzeit, zu mehr Expertise und damit zu einer geringeren Notwendigkeit instruktionaler Hilfestellung führt (Kalyuga, 2014). Damit liefert diese Studie hinsichtlich elaborierten Feedbacks vergleichbare Ergebnisse wie Fyfe et al. (2012).

Verglichen mit einem klassischen Übungsbetrieb (vgl. $KG_{(\text{kein LK})}$ und $KG_{(\text{LK})}$) bietet es sich also an, diese Lernaufgaben mit EST-Feedback als zusätzliches Lernangebot einzusetzen, da aktives Lernen mit den Aufgaben zu einem signifikant besseren Fachwissensstand am Ende des ersten Semesters führt. Der Vergleich in der Klausurperformance (Kapitel 8.2.4) zeigt außerdem, dass die Gruppe der Studierenden, die früher einen Grundkurs besucht haben und mit EST-Feedback gelernt haben, nicht signifikant schlechter als die ehemaligen Leistungkursstudierenden ungeachtet der Feedbackbedingung sind. Die Gruppe der Studierenden, die früher einen Grundkurs besucht haben und mit KOR-Feedback gelernt haben, weist hingegen eine signifikant schlechtere Klausurperformance im Vergleich zu Studierenden, die früher einen Leistungskurs Chemie belegt hatten, ungeachtet der Feedbackbedingung auf.

9.2 Theoretischer und praktischer Ertrag

Mit einem Blick auf die beschriebene Forschungslage im Sektor Feedbackforschung liefert diese Studie einen systematischen Vergleich zweier Feedbackmaßnahmen untereinander und mit einer Kontrollkohorte, die gar kein Feedback erhielt. Die meis-

ten der betrachteten Studien vergleichen Feedbackmaßnahmen untereinander, oftmals nicht trennscharf voneinander differenziert. Selten findet sich ein Treatment-Check mit einer Kontrollgruppe oder -Kohorte. Ferner konnte das Review von Cavalcanti et al. (2021) zeigen, dass naturwissenschaftliche Fächer und insbesondere die Chemie als Forschungsbereich in der Feedback Forschung unterrepräsentiert beziehungsweise hinsichtlich der Beforschung der Wirksamkeit von automatisiertem Feedback nicht vorhanden sind. Über die verschiedenen Metastudien, exemplarisch seien hier Wisniewski et al. (2019) und Hattie und Timperley (2007) genannt, zeigt sich im Mittel ein positiver Effekt für (elaboriertes) Feedback. So konnte auch diese Arbeit zeigen, dass automatisiertes (elaboriertes) Feedback grundsätzlich einen positiven Effekt auf die Performance hat, der mit steigender Expertise abzunehmen scheint. So konnten hier im Rahmen der Allgemeinen Chemie ähnliche Ergebnisse gefunden werden wie Fyfe et al. (2012) im Bereich Mathematik finden konnten. Bezogen auf den positiven Trend für automatisiertes Feedback in digitalen Lernumgebungen, der sich bei 32 der von Cavalcanti et al. (2021) ermittelten Artikel abzeichnet, liefert diese Arbeit damit nun auch einen empirischen Nachweis für positive Effekte automatisierten Feedbacks auf die Performance von Nutzerinnen und Nutzern für die Chemie und damit ebenfalls einen guten Ansatzpunkt für etwaige Replikationsstudien.

Diese Arbeit liefert zudem nicht nur Erkenntnisse hinsichtlich des Einflusses von Vorwissen auf die Wirkweise unterschiedlicher Feedback-Typen, sondern bildet durch die entwickelten Lernaufgaben und deren Einsatz im Feld ein Beispiel für einen webbasierten Übungsbetrieb in der Allgemeinen Chemie, wie einen möglichen Ausgangspunkt für andere Disziplinen. Des Weiteren ist im Rahmen dieser Arbeit das Tool JACK[®] um einige Funktionen erweitert worden. Der Formeleditor für chemische Formeln beinhaltet zusätzlich zu den Funktionen, die in den Aufgaben realisiert wurden, bereits die Möglichkeit auch Aggregatzustände in Reaktionsgleichungen mit einzubeziehen oder auch Zerfallsgleichungen in domänenspezifischer Schreibweise zu formulieren. Diese Funktionen konnten aber im Rahmen dieses Projektes nicht mehr im Feld auf Zuverlässigkeit getestet und letztlich in entsprechende Aufgaben implementiert werden. Nicht nur deshalb birgt JACK[®] für den webbasierten Übungsbetrieb in chemiespezifischen Veranstaltungen weiterhin Potential²⁴. Da gezeigt werden konnte, dass im ersten Semester die Arbeit mit solchen digitalen Lernaufgaben,

²⁴ Daran knüpft langfristig auch die Möglichkeit für digitales Prüfen an. Denn sobald für Lernaufgaben ein zuverlässiger Betrieb einstellt, können auch Abschlussklausuren in digitaler Form mit daran ange-

die elaborierte Hilfestellung und fehlerspezifisches Feedback bereitstellen, offensichtlich helfen kann, Vorwissensdefizite (teilweise) aufzuholen, muss für einen Ausbau des Angebots an so oder ähnlich beschaffenen Übungsaufgaben plädiert werden. Der Algorithmus liefert zudem performancebasiert eine adaptive Aufgabemenge bei fehlerinduzierter differenter Hilfestellung und klassifiziert sich damit als binnendifferenzierendes Förderangebot.

lehnten Aufgabenformaten konzipiert werden, was nicht nur den Druck von Klausuren und die Bereitstellung von Klausurbögen überflüssig macht, sondern sich über automatisierte Aufgabenbewertung auch Personalkosten einsparen ließen.

Kapitel 10: Limitationen und Ausblick

Durch die Pandemie, die sich im Frühjahr 2020 bereits andeutete, ergaben sich auch im Rahmen dieses Projekts Einschränkungen, allerdings nur für die zweite Hauptstudie. Das Umdisponieren hinsichtlich der Testformate (siehe Tabelle 10, S. 63) führte zu uneinheitlichen Testbedingungen (offline/online) und Lücken bei den Kontrollvariablen sowie auch hinsichtlich des ausschließlich online abgehaltenen Wintersemesters 2020/2021 zu größeren Veränderungen der Randbedingungen. Da nicht alle Kontrollvariablen in dem online Semester erhoben worden sind, waren einige Analysen nur bei Teilstichproben möglich. Der hohe Anteil an mathematischen Aufgabenbestandteilen in der Allgemeinen Chemie gibt berechtigte Anhaltspunkte dafür, die mathematischen Fähigkeiten mit in die Berechnung der Modelle zu integrieren. Dadurch, dass aber nicht alle Studierenden diesen Test gemacht haben, können solche Modelle leider nicht mit der nötigen Menge an Testpersonen gerechnet werden, um überhaupt Effekte in der zu erwartenden Größenordnung zu finden. Ein vorheriges Projekt (Averbeck, 2020) ermöglichte zwar den Vergleich mit Fachwissensdaten, Klausurergebnisse waren hingegen nicht verfügbar, sodass sich die Erkenntnisse aus dem Kohortenvergleich auf die Fachwissensentwicklung beschränken. Zudem wurde im Rahmen der Pilotierungsstudie die aktive Arbeit mit den Übungsaufgaben an die Regel geknüpft, dass „mindestens zwei Aufgaben aus mindestens fünf der Aufgabensets bearbeitet worden sein mussten“, um als *aktiv* lernende Person zu gelten und in den Interventionsgruppen berücksichtigt zu werden. Für die $N = 122$ erhaltenen Studierenden ergeben sich somit Schwankungen in der Menge an Aufgaben, die bearbeitet worden sind, wie auch Abbildung 24 und Abbildung 25 (S. 88 f.) zeigen. Dass sich hier dennoch bereits gruppenspezifische Effekte finden lassen, obwohl die Stichprobe unwesentlich größer ist, als das laut Power-Analyse notwendige Minimum für kleine bis mittlere Effektstärken, sowie nicht alle der potenziell möglichen Aufgabensets von jeder Person entsprechend des Algorithmus bearbeitet worden sind, deutet auf ein großes Potential so beschaffener Aufgaben hin. Zudem beschränken sich die gefundenen Effekte nur auf einen Hochschulstandort in Deutschland. Hier bedarf es weiterer Studien, diese Ergebnisse an anderen Hochschulstandorten zu replizieren.

Weiterer Forschungsbedarf ergibt sich aus dem Umstand, dass keine Interaktion mit dem Fachwissenstest gefunden werden konnte. Die Interaktion mit der Leistungskurswahl in der Oberstufe wurde als Vorteil an Lernzeit für kumulativen Wissensaufbau interpretiert. Unabhängig davon, dass eine Interaktion aus Kurswahl und Feed-

back gefunden worden ist, ist nicht sicher, ob diese in 7.2 vorgenommene Interpretation auch zutreffend ist. Vorwissen besteht nicht nur aus abrufbarem Wissen, sondern hängt auch von der Struktur ab, in der dieses Wissen gespeichert ist. Ein Test, korrespondierend zu einer Learning Progression, die das Schulwissen adäquat widerspiegelt, könnte zusätzlich zum vorhandenen Wissen auch Einblicke in die Wissensstruktur der Lernenden ermöglichen, wodurch das Konstrukt Vorwissen besser beschrieben werden könnte als durch den bestehenden Test von Averbek (2020). Mindestens wäre aber eine Modifikation der Items im Sinne des ESNaS-Modells (Walpuski et al., 2010) durch ergänzen eines Itemstamms nötig. Denn die kognitiven Prozesse, die bei den Testitems angestoßen werden, bleiben mangels Aufgabenstamm bislang im Dunkeln.

Wenngleich für die entwickelten Aufgaben mit fehlerspezifischem Feedback in Summe ein positiver Effekt auf den Studienerfolg (Fachwissensstand und Klausurperformance am Ende des Semesters) nachgewiesen werden konnte, muss auch hier einschränkend erwähnt werden, dass die beiden Algorithmen streng genommen zwei Unterschiede aufweisen. Zum einen enthält der EST-Feedback Algorithmus elaborierte Hilfestellung, die den Lernenden während des Lösungsprozesses bereitgestellt wird, zum anderen enthält die Musterlösung des KOR-Feedback-Algorithmus keine Hinweise zum Lösungsprozedere, auch keine unkommentierten Zwischenschritte bei der Aufgabenlösung. Damit ergibt sich zugunsten des EST-Feedback-Algorithmus aufgrund der dortigen elaborierten Musterlösung ein Informationsplus, selbst, wenn man das Feedback außer Acht lässt. Daher kann durch den Vergleich dieser beiden Algorithmen nicht vollständig trennscharf ermittelt werden, zu welchem Anteil der resultierende Vorteil hinsichtlich Klausurergebnis und Fachwissensstand letztlich auf das elaborierte Feedback oder die umfangreichere Aufgabenlösung zurückzuführen ist.

Betrachtet man den Dropout der Studie, so konnte nicht verhindert werden, dass über ein Semester ein Schwund von 58.1 %, beziehungsweise im Mittel 46.5 %, wenn man das pandemiebedingte Online-Semester aufgrund der Sonderstellung exkludiert, an Studierenden zu verzeichnen gewesen ist. Die Bearbeitung der Aufgaben wirkt sich zwar signifikant positiv auf die Wahrscheinlichkeit aus, die Klausur zu bestehen, was für eine gewisse Qualität der Aufgaben spricht, da aber im Verlauf des Semesters immer mehr Studierende hinsichtlich der Aufgabenbearbeitung wegbrechen, deutet dies sicherlich auf einen Optimierungsbedarf einiger Aufgaben hin. Welche Optimierungen hier angestrebt werden könnten, soll an einem ausgewählten Beispiel diskutiert werden:

Da die Freitext-Eingabe von Lösungen softwarebedingt nicht möglich war, ergab sich für einige Aufgaben eine Vielzahl von Fill-In Feldern. So konnte für viele Aufgaben hinsichtlich Aufgabenstamm und Lösungseingabe eine Darstellung erreicht werden, die sehr nah an der einer korrespondierenden Paper-Pencil Aufgabe liegt, wodurch sich aber Einschränkungen hinsichtlich der Nutzerfreundlichkeit ergaben (s. Abbildung 26, S. 90), sehr deutlich zu erkennen bei den „Redox“-Aufgaben. Am folgenden Aufgabenbeispiel kann die durchaus unvorteilhafte Bewertung des Aufgabeninterfaces gut erklärt werden (s. Abbildung 33). Zwar weist die Aufgabe hinsichtlich Erscheinungsbild und Art der Eingabe starke parallelen zu einer korrespondierenden Paper-Pencil Aufgabe auf, für das vorliegende Beispiel ergeben sich damit allerdings 14 notwendige (Neu-)Eingaben pro Lösungsversuch, die über jeden Lösungsversuch der Aufgabe leider softwarebedingt *neu* ausgefüllt werden mussten.

Vervollständigen Sie die folgende Reaktionsgleichung. Gehen Sie dabei wie folgt vor:

- Bestimmen Sie die Oxidationszahlen der einzelnen Atome bzw. Ionen. Verwenden Sie dazu +/- römische Ziffern (z.B. -II oder +IV) oder die 0.
- Formulieren Sie anhand der nun erkennbaren Elektronenübergänge die Teilgleichungen für Oxidation und Reduktion. Wählen Sie hierbei die stöchiometrischen Koeffizienten so **klein wie möglich**.
- Stellen Sie die Gesamtgleichung für die Redoxreaktion auf und kürzen Sie so weit wie möglich.

In Salpetriger Säure Umgebung findet folgende Reaktion statt (unausgeglichen und unvollständig):

$$\boxed{} \boxed{} \boxed{} \boxed{} + \boxed{} \boxed{} \boxed{} \boxed{} \rightleftharpoons \boxed{} \boxed{} \boxed{} \boxed{}$$

$N \ O + N \ O_3^- \rightleftharpoons N_2 \ O_4$

Oxidation: → | * (Erweiterungsfaktor)

Reduktion: → | * (Erweiterungsfaktor)

Redoxreaktion: →

Abbildung 33: Screenshot einer Aufgabe zum Aufstellen von Redox-Reaktionen

Es ist nicht unwahrscheinlich, dass es bei der Neubearbeitung einer Aufgabe, insbesondere bei der Übertragung der richtigen Ergebnisse, zu Flüchtigkeitsfehlern kommen kann, wodurch Aufgaben erneut bearbeitet werden müssen, auch wenn das Endergebnis (hier die Redox-Reaktion) richtig gewesen ist. Bei dem obigen Beispiel kommt hinzu, dass sechs der zu bearbeitenden Felder komplexe chemische Formeln beinhalten sollen. Diese Eingabe ist zusätzlich zur erhöhten Wahrscheinlichkeit von

Flüchtigkeitsfehler noch sehr zeitintensiv. Das führt rückwirkend zu der Einschätzung, dass gerade komplexere Aufgaben mit vielen und/oder komplexen Eingabefeldern schnell zu Motivationsverlust führen können.

Gleichzeitig wurden im Gegensatz zu den Einzelstudien, die die Lernwirksamkeit von differenter Hilfestellung in Abhängigkeit vom Vorwissen untersucht haben, nicht nur ein Themengebiet (Moreno, 2004; Narciss & Huth, 2004; Smits et al., 2008), sondern gleich eine Vielzahl unterschiedlicher Themenfelder mit insgesamt 20 verschiedenen Lernzielen mit Lernaufgaben versehen. Der individuelle Lernstand der Probanden kann dazu führen, dass jemand für eine Aufgabe zum Stoffumsatz eher als Experte einzustufen ist, für eine Aufgabe zu Redox-Reaktionen hingegen eher als Novize. Während bei der ersten Aufgabe ein zweiter oder dritter Lösungsversuch gerne genutzt wird, um selbstständig das richtige Ergebnis zu ermitteln, was letztlich zu Kompetenzerleben führen kann (Ryan & Deci, 2017), mag ein weiterer Lösungsversuch bei einer Aufgabe zu Redox-Reaktionen eher als störend oder gar demotivierend empfunden werden, da mit dem eigenen Scheitern gerechnet wird. Gerade, wenn die Eingabe sehr zeitaufwändig ist, kann das trotz gewissenhafter Auseinandersetzung mit den Aufgaben zusätzlich demotivierend wirken. Hier bietet sich eine Optimierung des EST-Algorithmus an (siehe Abbildung 34).

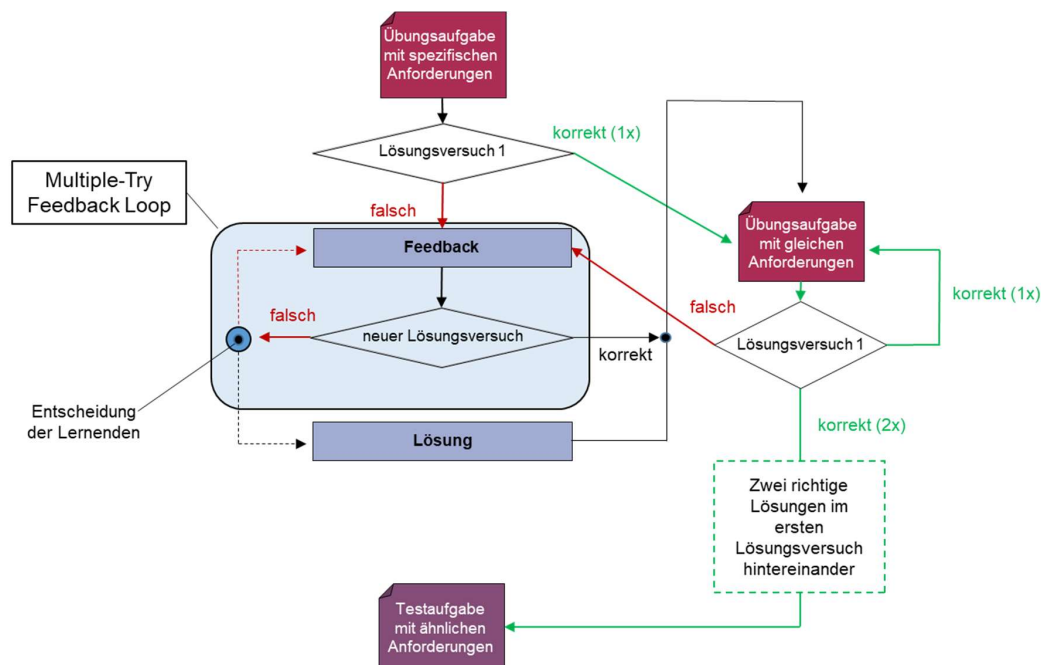


Abbildung 34: Entwurf eines Self-Determined-Multiple-Try-Tutoring-Feedback Algorithmus

Damit die Probanden nach der ersten Bearbeitung der Aufgabe und dem ersten Erhalten fehlerspezifischen tutoriellen Feedbacks nicht in der Aufgabe gefangen sind, wenn sehr viele Fehler zu berichtigen sind, wurde der Feedback-Algorithmus durch einen multiple-try Feedback-Loop so modifiziert, dass die Studierenden sich entscheiden können, entweder die Aufgabe nach einem Lösungsversuch mit Hilfe des fehlerspezifischen Feedbacks eigenständig zu lösen oder nach dem fehlerspezifischen Feedback direkt eine elaborierte Musterlösung angezeigt zu bekommen. Dies könnte zudem durch die Funktion ergänzt werden, dass die Eingabefelder bereits mit den Lösungen des vorherigen Versuchs vorausgefüllt sind, sodass nur noch die farblich als falsch gekennzeichneten Eingabefelder²⁵ berichtigt werden müssen. Mindestens eine Performanceleistung in der Aufgabe verhindert weiterhin die Feedback Presearch Availability (Bangert-Drowns, 1991), während nach dem ersten Feedback der Adressat selbst entscheiden kann, wie weiter verfahren wird.

Es geht bei diesen Lernaufgaben darum, fehlendes Wissen zu vermitteln. Das kann (unter Umständen sogar besser) direkt am konkreten Aufgabenbeispiel erfolgen. Wird bei sehr vielen Fehlern nämlich die Hilfestellung an einem verwandten Worked Example bereitgestellt, führt die Transferleistung, die notwendig ist, die Hilfestellung auf die aktuelle Aufgabe zu übertragen, zu einer kognitiven Belastung, die so umgangen werden kann. Gleichmaßen lädt diese „skip-Funktion“ natürlich dazu ein, schnell und arbeitssparsam zur Lösung zu gelangen. Denjenigen, die die Aufgaben allerdings gewissenhaft bearbeiten, könnte so eine Arbeitserleichterung zuteilwerden, denn die wichtige Hilfestellung ist ja auch in der elaborierten Musterlösung in Form eines Worked Examples enthalten. Ob ein solcher Ansatz zu einer verbesserten Effizienz der Lernaufgaben hinsichtlich der Förderung des Fachwissenserwerbs und langfristig zu einem höheren Studienerfolg führt, gilt es zu prüfen.

²⁵ Bislang wurde der Fehlerort im Feedback-Text eingegrenzt. Für Folgeversionen von JACK® ist bereits geplant, fehlerhafte Eingabefelder nach Einreichen der Aufgabenlösung farblich hinterlegen zu können.

Abbildungsverzeichnis

Abbildung 1: Studienabbruchsmodell (Heublein, 2014; Paczulla, Fischer, Sumfleth & Walpuski, 2019).....	7
Abbildung 2: Ausschlaggebende Studienabbruchmotive nach Abschlussart, unabhängig vom Studiengang (Heublein et al., 2017, S. 23)	9
Abbildung 3: Fachwissen in der Allgemeinen Chemie zu Beginn und zum Ende des ersten Semesters in Abhängigkeit zur Kurswahl in der Oberstufe (nach Avertebeck, 2020).....	13
Abbildung 4: Informationsverarbeitungsmodell des Gedächtnisses adaptiert nach Woolfolk (2014, S. 280), Hussy, Fritz und Tobinski (2018, S. 84) und Moreno und Mayer (2007, S. 314).....	20
Abbildung 5: Der Lehlalgorithmus (nach Klauer & Leutner, 2012, S. 47).....	23
Abbildung 6: Modell des Selbstregulierten Lernens (Narciss, 2006, S. 53) nach Butler und Winne (1995).....	25
Abbildung 7: Ein Modell über Feedback zur Verbesserung des Lernens (Hattie, 2013)	31
Abbildung 8: Interactive two feedback-loops Model – ITFL-Modell (Narciss, 2018).....	33
Abbildung 9: Übersicht über Determinanten des Informationswerts von Feedback (nach Narciss, 2008)	35
Abbildung 10: Dreigliedriges Spannungsfeld in Feedback präsentierenden Softwarelösungen.....	39
Abbildung 11: Auswahl und Spezifizierung von Feedbackelementen (nach Narciss und Huth, 2004).....	41
Abbildung 12: Kapazität des Arbeitsgedächtnisses (nach Sweller et al., 1998)	44
Abbildung 13: Integratives Modell der Text und Bildbearbeitung (nach Schnotz, 2002)	46
Abbildung 14: Modell zur Entwicklung digitaler Lernumgebungen mit verallgemeinerter Inhaltskomponente (nach Reinhold, 2019)	54
Abbildung 15: Überführung der Papier-Bleistift-Aufgaben in digitale Lernaufgaben unter Berücksichtigung der Einschränkungen durch die Software Lösung.....	65
Abbildung 16: Realisierung multimedialer Prinzipien am Beispiel eines BRT-Feedbackausschnitts von Aufgaben zur Redoxreaktion.....	68
Abbildung 17: Beispiel strukturierender Antwortsätze aus dem Aufgabenset zum Born-Haber-Kreisprozess.....	71
Abbildung 18: Gegenüberstellung des Korrektiven Feedback Algorithmus mit dem BRT-Feedback-Algorithmus von Narciss und Huth (2004).....	74

Abbildungsverzeichnis

Abbildung 19: Lernpfad-Regeln exemplarisch für die dritte Aufgabe des Aufgabensets „Arbeiten mit dem PSE“	76
Abbildung 20: Ausschnitt aus dem Moodle-Kurs des Online-Tutoriums WiSe 19/20	78
Abbildung 21: Klausurpunkte je Interventionsgruppe, Pilotierung WiSe '18	83
Abbildung 22: Entwicklung des Chemiewissens nach Interventionsgruppe Pilotierung WiSe '18	84
Abbildung 23: Gegenüberstellung des korrektiven Feedback Algorithmus mit dem Error-Specific-Tutoring-Feedback Algorithmus (Trauten et al., 2023)	85
Abbildung 24: Bearbeitungshäufigkeiten der Aufgabensets pro Gruppe (WiSe '19)	88
Abbildung 25: Bearbeitungshäufigkeiten der Aufgabensets pro Gruppe (WiSe '20)	89
Abbildung 26: Bewertung des Aufgabeninterface je Aufgabe	90
Abbildung 27: Bewertung des Feedback-Nutzens je Aufgabe	90
Abbildung 28: Entwicklung des Chemiewissens nach Interventionsgruppe (WiSe '18, '19 und '20).....	91
Abbildung 29: Streudiagramm mit gruppenspezifischen Regressionsgeraden des 2. Modells zur Vorhersage des Fachwissens (siehe Tabelle 15)	93
Abbildung 30: Histogramm für den Lernzuwachs nach Interventionsgruppe (WiSe '18, '19 u.'20).....	95
Abbildung 31: Klausurpunkte in Abhängigkeit von der Interventionsgruppe WiSe '18 u. '19 (** $p \leq .01$)	98
Abbildung 32: Entwicklung des Chemie Fachwissens nach Interventionsgruppe (Kohortenvergleich).....	103
Abbildung 33: Screenshot einer Aufgabe zum Aufstellen von Redox-Reaktionen	113
Abbildung 34: Entwurf eines Self-Determined-Multiple-Try-Tutoring-Feedback Algorithmus	114

Tabellenverzeichnis

Tabelle 1: Übersicht über postulierte Feedback-Funktionen	24
Tabelle 2: Sieben mögliche Reaktionen auf Informationen, die den eigenen Erwartungen, Vorstellungen oder Wissensstrukturen widersprechen (Chinn & Brewer, 1993).....	27
Tabelle 3: Klassifikation unterschiedlicher Feedback-Komponenten nach inhaltlichen Gesichtspunkten (Narciss, 2006, S. 23)	28
Tabelle 4: Grundlegende Definitionen der CTML (nach Mayer, 2014, S. 2).....	47
Tabelle 5: Drei Arten multimedialer Learning Outcomes (nach Mayer, 2014, S. 21)	48
Tabelle 6: Ausgewählte Grundprinzipien multimedialen Designs nach (Mayer, 2009, 2014).....	50
Tabelle 7: Interventionsgruppen im 2x2 Design	59
Tabelle 8: Testinstrumente zur Erfassung von Kontrollvariablen.....	61
Tabelle 9: Studiendesign für Pilotierung und Hauptstudie 1, Semesterplan	62
Tabelle 10: Eingesetzte Instrumente je Kohorte.....	63
Tabelle 11: Formulierungshilfe für Lernziele (Handlungskomponente) nach Cursio und Jahn (2015)	66
Tabelle 12: Verteilung verwertbarer Operatoren (Trauten, 2018).....	70
Tabelle 13: Stichprobenbeschreibung	79
Tabelle 14: Verteilung der „aktiv“ Studierenden pro Interventionsgruppe.....	80
Tabelle 15: Ergebnisse hierarchischer linearer Regressionsanalysen zur Vorhersage des Fachwissens (Post)	92
Tabelle 16: Ergebnisse der multiplen linearen Regressionsanalyse zur Vorhersage des Lernzuwachses.....	96
Tabelle 17: Ergebnisse der multiplen linearen Regressionsanalyse zur Vorhersage des Klausurerfolgs.....	99
Tabelle 18: Multiple lineare Regressionen zur Vorhersage des Fachwissens (Post), Kohortenvergleich	102

Literaturverzeichnis

- Akçayır, M. & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11. <https://doi.org/10.1016/j.edurev.2016.11.002>
- Albacete, P. L. & VanLehn, K. A. (2000). Evaluating the Effectiveness of a Cognitive Tutor for Fundamental Physics Concepts. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 22(22). Verfügbar unter: <https://escholarship.org/uc/item/0166b7p0>
- Ashton, H. S., Beevers, C. E., Korabinski, A. A. & Youngson, M. A. (2005). Investigating the medium effect in computer-aided assessment of school Chemistry and college Computing national examinations. *British Journal of Educational Technology*, 36(5), 771–787. <https://doi.org/10.1111/j.1467-8535.2005.00501.x>
- Autorengruppe Bildungsberichterstattung; W. Bertelsmann Verlag. (2018). *Bildung in Deutschland 2018. Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung* (1. Auflage). Bielefeld: wbv Media. Verfügbar unter: <http://www.bildungsbericht.de>
- Averbeck, D. (2020). *Zum Studienerfolg in der Studieneingangsphase des Chemie-studiums* (Studien zum Physik- und Chemielernen, Bd. 308). Berlin: Logos.
- Bandura, A. (1979). *Sozial-kognitive Lerntheorie* (Konzepte der Humanwissenschaften, 1. Aufl.). Stuttgart: Klett-Cotta.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A. & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61(2), 213–238. <https://doi.org/10.3102/00346543061002213>
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U. et al. (Hrsg.). (2002). *PISA 2000 — Die Länder der Bundesrepublik Deutschland im Vergleich* (1st ed. 2002). Wiesbaden: VS Verlag für Sozialwissenschaften; Imprint VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-663-11042-2>
- Beckmann, J. F. (2010). Taming a beast of burden – On some issues with the conceptualisation and operationalisation of cognitive load. *Learning and Instruction*, 20(3), 250–264. <https://doi.org/10.1016/j.learninstruc.2009.02.024>
- Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), 4–16. <https://doi.org/10.3102/0013189X013006004>
- Bond, T. G., Yan, Z. & Heene, M. (2020). *Applying the Rasch model. Fundamental measurement in the human sciences* (Fourth edition). New York, NY: Routledge. Retrieved from <https://www.taylorfrancis.com/books/9780429030499>

- Bourdieu, P. (1987). *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft* (Suhrkamp-Taschenbuch Wissenschaft, Bd. 658, 1. Aufl., [Nachdr.]. (französisch: La distinction. Critique sociale du jugement. Paris 1979). Frankfurt am Main: Suhrkamp.
- Brossardt, B., Ehrentraut, O. (Vereinigung der Bayerischen Wirtschaft e. V., Hrsg.). (2015). *Arbeitslandschaft 2040*. Zugriff am 23.08.2021. Verfügbar unter: https://www.prognos.com/sites/default/files/2021-01/20150521_prognos_arbeitslandschaft2040-final.pdf
- Butler, D. L. & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281.
- Caduff, C., Pfiffner, M. & Bürgi, V. (2018). *Lernen. Eine Publikation der Pädagogischen Hochschule Zürich* (1. Aufl.). Didaktische Hausapotheke Band 11. Bern: hep verlag. Verfügbar unter: http://www.content-select.com/index.php?id=bib_view&ean=9783035503012
- Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, 86(3), 360–367. <https://doi.org/10.1037/0022-0663.86.3.360>
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D. et al. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, 100027. <https://doi.org/10.1016/j.caeai.2021.100027>
- Chen, O., Kalyuga, S. & Sweller, J. (2017). The Expertise Reversal Effect is a Variant of the More General Element Interactivity Effect. *Educational Psychology Review*, 29(2), 393–405. <https://doi.org/10.1007/s10648-016-9359-1>
- Chen, Q. (2013). *Remove or keep. Linking items showing item parameter drift* (Dissertations Abstracts International). Ann Arbor: ProQuest Dissertations & Theses. Retrieved from http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqm&rft_dat=xri:pqdiss:3560290
- Chen, X. (2015). STEM attrition among high-performing college students: Scope and potential causes. *Journal of Technology and Science Education*, 5(1). <https://doi.org/10.3926/jotse.136>
- Chinn, C. A. & Brewer, W. F. (1993). The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Review of Educational Research*, 63(1), 1–49. <https://doi.org/10.3102/00346543063001001>

- Cohen, P. A., Kulik, J. A. & Kulik, C.-L. C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal*, 19(2), 237–248. <https://doi.org/10.3102/00028312019002237>
- Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79(4), 347–362. <https://doi.org/10.1037/0022-0663.79.4.347>
- Cursio, M. & Jahn, D. (2015). *Formulierung kompetenzorientierter Lernziele auf Modulebene*. Erlangen: Friedrich-Alexander Universität Erlangen-Nürnberg.
- Dickmann, T. (2019). *Visuelles Modellverständnis und Studienerfolg in der Chemie. Zwei Seiten einer Medaille* (Studien zum Physik- und Chemielernen).
- Ditton, H. & Müller, A. (2014). Feedback: Begriff, Formen und Funktion. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11–28). Münster: Waxmann.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th edition). Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.
- Fischer, H. E., Glemnitz, I., Kauertz, A. & Sumfleth, E. (2007). Auf Wissen aufbauen - kumulatives Lernen in der Chemie und Physik. In E. Kircher (Hrsg.), *Physikdidaktik. Theorie und Praxis* (Springer-Lehrbuch, S. 657–678). Berlin: Springer-Verlag Berlin. Verfügbar unter: https://link.springer.com/content/pdf/10.1007/978-3-540-34091-1_22.pdf
- Freyer, K. (2013). *Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie* (Studien zum Physik- und Chemielernen, Bd. 156). Berlin: Logos.
- Fritz, A. & Tobinski, D. (2018). Lerntheorien und pädagogisches Handeln. In A. Fritz, W. Hussy & D. Tobinski (Hrsg.), *Pädagogische Psychologie. Mit ... 91 Kontrollfragen* (Basics, Bd. 3373, 3. Auflage, S. 222–246). München: Ernst Reinhardt Verlag.
- Fyfe, E. R., Rittle-Johnson, B. & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104(4), 1094–1108. <https://doi.org/10.1037/a0028389>
- Glenberg, A. M., Wilkinson, A. C. & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, 10(6), 597–602. <https://doi.org/10.3758/BF03202442>

- Goedicke, M. (2022). *E-Learning und E-Assessment: paluno – the Ruhr Institute for Software Technology*. Zugriff am 21.04.2022. Verfügbar unter: <https://s3.paluno.uni-due.de/forschung/e-learning-und-e-assessment>
- Goldhausen, I. (2015). *Mathematische Modelle im Chemieunterricht*. Dissertation. Uni-Edition GmbH.
- Hattie, J. (2013). *Lernen sichtbar machen* (Überarb. dt.-sprachige Ausg. von „Visible learning“). Baltmannsweiler: Schneider-Verl. Hohengehren.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R)*. Göttingen: Beltz Test.
- Hellmich, F. (Hrsg.). (2011). *Selbstkonzepte im Grundschulalter. Modelle, empirische Ergebnisse, pädagogische Konsequenzen* (Schulpädagogik). Stuttgart: Kohlhammer.
- Heublein, U. (2014). Student Drop-out from German Higher Education Institutions. *European Journal of Education*, 49(4), 497–513. <https://doi.org/10.1111/e-jed.12097>
- Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J. et al. (2017). *Zwischen Studierenerwartungen und Studienwirklichkeit. Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen* (Forum Hochschule). Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH.
- Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D. & Besuch, G. (2010). *Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. Ergebnisse einer Bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08*, HIS: Hochschul Information System GmbH. Zugriff am 03.11.2021. Verfügbar unter: <https://hsdbs.hof.uni-halle.de/documents/t1944.pdf>
- Heublein, U., Richter, J. & Schmelzer, R. (2020). *Die Entwicklung der Studienabbruchquoten in Deutschland*. https://doi.org/10.34878/2020.03.DZHW_BRIEF
- Heublein, U., Schmelzer, R. & Sommer, D. (2005). *Studienabbruchsstudie 2005. Die Studienabbrecherquoten in den Fächergruppen und Studienbereichen der Universitäten und Fachhochschulen*, Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH.
- Huber, L. (1991). Sozialisation in der Hochschule. In K. Hurrelmann & D. Ulich (Hrsg.), *Neues Handbuch der Sozialisationsforschung* (4., völlig neubearb. Aufl., S. 417–441). Weinheim: Beltz.

- Hussy, W. (2018). Aufmerksamkeit. In A. Fritz, W. Hussy & D. Tobinski (Hrsg.), *Pädagogische Psychologie. Mit ... 91 Kontrollfragen* (Basics, Bd. 3373, 3. Auflage, 63-79). München: Ernst Reinhardt Verlag.
- Hussy, W., Fritz, A. & Tobinski, D. (2018). Strukturen und Prozesse im Arbeitsgedächtnis. In A. Fritz, W. Hussy & D. Tobinski (Hrsg.), *Pädagogische Psychologie. Mit ... 91 Kontrollfragen* (Basics, Bd. 3373, 3. Auflage, S. 80–96). München: Ernst Reinhardt Verlag.
- IQB & KMK. (2020). *Aufgaben für die Fächer Biologie, Chemie und Physik. Einheitliche Operatorenliste*, Institut zur Qualitätsentwicklung im Bildungswesen; Kultusminister Konferenz. Zugriff am 27.04.2022. Verfügbar unter: https://www.iqb.hu-berlin.de/abitur/abitur/dokumente/naturwissenschaften/N_Einheitliche_O.pdf
- Jacobs, B. (Medienzentrum der Philosophischen Fakultät der Universität Saarbrücken, Hrsg.). (1998). *Aufgaben stellen und Feedback geben*. Zugriff am 04.07.2018. Verfügbar unter: <http://www.phil.uni-sb.de/~jakobs/wwwartikel/feedback/index.html>
- Johnson, C. I. & Priest, H. A. (2014). The Feedback Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 449–463). Cambridge: Cambridge University Press.
- Kalyuga, S. (2014). The Expertise Reversal Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 576–597). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.028>
- Kimpel, L. (2017). *Aufgaben in der Allgemeinen Chemie*. Dissertation. Universität Duisburg-Essen; Logos Verlag Berlin.
- Klauer, K. J. & Leutner, D. (2012). *Lehren und Lernen. Einführung in die Instruktionspsychologie* (2. Aufl.). Weinheim: Beltz Verlagsgruppe. Verfügbar unter: http://www.content-select.com/index.php?id=bib_view&ean=9783621279789
- Kluger, A. N. & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), 254–284. Zugriff am 19.05.2020. Verfügbar unter: https://www.researchgate.net/publication/232458848_The_Effects_of_Feedback_Interventions_on_Performance_A_Historical_Review_a_Meta-Analysis_and_a_Preliminary_Feedback_Intervention_Theory
- Kopp, B. & Mandl, H. (2014). Lerntheoretische Grundlagen von Rückmeldungen. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 29–41). Münster: Waxmann.

- Kosovich, J. J., Hulleman, C. S., Barron, K. E. & Getty, S. (2015). A Practical Measure of Student Motivation. *The Journal of Early Adolescence*, 35(5-6), 790–816. <https://doi.org/10.1177/0272431614556890>
- Krause, U.-M., Stark, R. & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, 19(2), 158–170. <https://doi.org/10.1016/j.learninstruc.2008.03.003>
- Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos (Ed.), *Criticism and the growth of knowledge* (Proceedings of the International Colloquium in the Philosophy of Science, vol. 4, S. 91–196). London: Cambridge Univ. Press. <https://doi.org/10.1017/CBO9781139171434.009>
- Lehmenkühler, A., Roscher, A. & Theis, W. (1976). Feedback: Anmerkungen zu Funktion und Form. In M. Sader (Hrsg.), *Verbesserung von Interaktion durch Gruppendynamik. Psychologische Grundlagen und pragmatische Strukturierungen* (Arbeiten zur sozialwissenschaftlichen Psychologie, Bd. 5). Münster: Aschendorff.
- Mayer, R. E. (1992). Cognition and instruction: Their historic meeting within educational psychology. *Journal of Educational Psychology*, 84(4), 405–412. Zugriff am 09.11.2021.
- Mayer, R. E. (2009). *Multimedia Learning*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Mayer, R. E. (2014). Introduction to Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. & Johnson, C. I. [Cheryl I.]. (2010). Adding Instructional Features That Promote Learning in a Game-Like Environment. *Journal of Educational Computing Research*, 42(3), 241–265. <https://doi.org/10.2190/EC.42.3.a>
- McLeod, D. B., Cronbach, L. J. & Snow, R. E. (1978). Aptitudes and Instructional Methods: A Handbook for Research on Interactions. *Journal for Research in Mathematics Education*, 9(5), 390. <https://doi.org/10.2307/748778>
- Moreno, R. (2004). Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science*, 32(1/2), 99–113. <https://doi.org/10.1023/B:TRUC.0000021811.66966.1d>
- Moreno, R. & Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, 19(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>

- Mortimer, C. E. & Müller, U. (2003). *Chemie. Das Basiswissen der Chemie ; 125 Tabellen* (8., komplett überarb. und erw. Aufl.). Stuttgart: Thieme.
- Müller, J., Stender, A., Fleischer, J., Borowski, A., Dammann, E., Lang, M. et al. (2018). Mathematisches Wissen von Studienanfängern und Studienerfolg. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 183–199. <https://doi.org/10.1007/s40573-018-0082-y>
- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 56). Dresden, Techn. Univ., Habil.-Schr., 2005. Münster: Waxmann.
- Narciss, S. (2008). Feedback strategies for interactive learning task. In J. M. Spector (Ed.), *Handbook of research on educational communications and technology* (Education, 4th ed., S. 125–144). New York, NY: Springer.
- Narciss, S. (2012). Feedback in Instructional Contexts. In N. M. Seel (Hrsg.), *Encyclopedia of the Sciences of Learning* (S. 1285–1292). Boston, MA: Springer US.
- Narciss, S. (2014). Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 43–82). Münster: Waxmann.
- Narciss, S. (2018). Feedbackstrategien für interaktive Lernaufgaben. In H. M. Niegemann & A. Weinberger (Hrsg.), *Handbuch Bildungstechnologie. Konzeption und Einsatz digitaler Lernumgebungen* (Springer reference, S. 1–24). Berlin: Springer. https://doi.org/10.1007/978-3-662-54373-3_35-1
- Narciss, S. & Huth, K. (2004). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction*, 16(4), 310–322. <https://doi.org/10.1016/j.learninstruc.2006.07.003>
- Neugebauer, M., Heublein, U. & Daniel, A. (2019). Studienabbruch in Deutschland: Ausmaß, Ursachen, Folgen, Präventionsmöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, 22(5), 1025–1046. <https://doi.org/10.1007/s11618-019-00904-1>
- Organisation for Economic Co-operation and Development. (2008). *Encouraging Student Interest in Science and Technology Studies*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264040892-en>
- Paas, F. G. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86(1), 122–133. <https://doi.org/10.1037/0022-0663.86.1.122>

- Paczulla, B., Fischer, V., Sumfleth, E. & Walpuski, M. (2019). Development and Evaluation of a Chemistry Test for Higher Education. In O. Levrini & Tasquier Giulia (Hrsg.), *The Beauty and Pleasure of Understanding: Engaging with Contemporary Challenges Through Science Education* (S. 1969–1975).
- Pobel, S. & Striewe, M. (2019). Domain-Specific Extensions for an E-Assessment System. In M. A. Herzog, Z. Kubincová, P. Han & M. Temperini (Hrsg.), *Advances in Web-Based Learning – ICWL 2019* (Bd. 11841, S. 327–331). Cham: Springer International Publishing.
- Portele, G. (1985). Habitus und Lernen. Sozialpsychologische Überlegungen. *Neue Sammlungen*, 25(3), 298–313.
- Proske, A., Körndle, H. & Narciss, S. (2012). Interactive Learning Tasks. In N. M. Seel (Hrsg.), *Encyclopedia of the Sciences of Learning* (S. 1606–1610). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1100
- QUA-LiS NRW (Käuser, R., Hrsg.). (2022). *Schulentwicklung NRW - Lehrplannavigator S II - Gymnasiale Oberstufe - Chemie - Hinweise und Beispiele - Beispiel eines schulinternen Lehrplans für die gymnasiale Oberstufe im Fach Chemie*. Zugriff am 03.06.2022. Verfügbar unter: <https://www.schulentwicklung.nrw.de/lehrplaene/lehrplannavigator-s-ii/gymnasiale-oberstufe/chemie/hinweise-und-beispiele/schulinterner-lehrplan/schulinterner-lehrplan.html>
- Rau, M. A. (2017). Conditions for the Effectiveness of Multiple Visual Representations in Enhancing STEM Learning. *Educational Psychology Review*, 29(4), 717–761. <https://doi.org/10.1007/s10648-016-9365-3>
- Reinhold, F. (2019). *Wirksamkeit von Tablet-PCs bei der Entwicklung des Bruchzahlbegriffs aus mathematikdidaktischer und psychologischer Perspektive. Eine empirische Studie in Jahrgangsstufe 6* (Studien zur theoretischen und empirischen Forschung in der Mathematikdidaktik). Wiesbaden: Springer Fachmedien Wiesbaden; Imprint Springer Spektrum. <https://doi.org/10.1007/978-3-658-23924-4>
- Renkl, A. (1996). Vorwissen und Schulleistung. In J. Möller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 175–190). Weinheim: Beltz PsychologieVerl.-Union.
- Renkl, A. (2014a). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, 38(1), 1–37. <https://doi.org/10.1111/cogs.12086>
- Renkl, A. (2014b). The Worked Examples Principle in Multimedia Learning. In R. E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 391–412). Cambridge: Cambridge University Press.
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs-Hug, P. Füglistner, K.

- Reusser, H. Wyss & F. Achtenhagen (Hrsg.), *Didaktik auf psychologischer Grundlage. Von Hans Aeblis kognitionspsychologischer Didaktik zur modernen Lehr- und Lernforschung* (1. Auflage). Bern: hep der Bildungsverlag.
- Rost, F. (2018). *Lern- und Arbeitstechniken für das Studium* (Lehrbuch, 8., vollständig überarbeitete und aktualisierte Auflage). Wiesbaden: Springer VS.
- Ryan, R. M. & Deci, E. L. (2017). *Self-determination theory. Basic psychological needs in motivation, development, and wellness*. New York, London: The Guilford Press.
- Saracetti, A. & Müller, S. (2011). Zum Stand der Studienabbruchforschung. Theoretische Perspektiven, zentrale Ergebnisse und methodische Anforderungen an künftige Studien. *Zeitschrift für Bildungsforschung*, 1(3), 235–248. <https://doi.org/10.1007/s35834-011-0020-2>
- Schlusche, C., Schnaubert, L. & Bodemer, D. (2021). Perceived Social Resources Affect Help-Seeking and Academic Outcomes in the Initial Phase of Undergraduate Studies. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.732587>
- Schnotz, W. (2002). Towards an Integrated View of Learning From Text and Visual Displays. *Educational Psychology Review*, 14(1), 101–120. <https://doi.org/10.1023/A:1013136727916>
- Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts: SESSKO*. Göttingen: Hogrefe.
- Schuhmacher, R. (2008). Der produktive Umgang mit Fehlern. Fehler als Lerngelegenheit und Orientierungshilfe. In R. Caspary (Hrsg.), *Nur wer Fehler macht, kommt weiter. Wege zu einer neuen Lernkultur* (Herder-Spektrum, Bd. 5892, Orig.-Ausg, S. 49–72). Freiburg im Breisgau: Herder.
- Schwedler, S. (2017). Was überfordert Chemiestudierende zu Studienbeginn? *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 165–179.
- Smits, M. H., Boon, J., Sluijsmans, D. M. & van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: effects on learning as a function of prior knowledge. *Interactive Learning Environments*, 16(2), 183–193. <https://doi.org/10.1080/10494820701365952>
- Spinath, B., Stiensmeier-Pelster, J., Schöne, C. & Dickhäuser, O. (2002). *Skalen zur Erfassung von Lern- und Leistungsmotivation: SELLMO*. Göttingen: Hogrefe.
- Striwe, M. (2016). An architecture for modular grading and feedback generation for complex exercises. *Science of Computer Programming*, 129, 35–47. <https://doi.org/10.1016/j.scico.2016.02.009>

- 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., Ayres, P. & Kalyuga, S. (2011). *Cognitive load theory* (Explorations in the learning sciences, instructional systems and performance technologies, vol. 1). New York, NY: Springer. <https://doi.org/10.1007/978-1-4419-8126-4>
- Sweller, J. & Cooper, G. A. (1985). The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. *Cognition and Instruction*, 2(1), 59–89. https://doi.org/10.1207/s1532690xci0201_3
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296. Zugriff am 26.05.2020. Verfügbar unter: <https://link.springer.com/content/pdf/10.1023/A:1022193728205.pdf>
- Thomas, L. (2002). Student retention in higher education: the role of institutional habitus. *Journal of Education Policy*, 17(4), 423–442. <https://doi.org/10.1080/02680930210140257>
- Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), 89. <https://doi.org/10.2307/1170024>
- Tinto, V. (1988). Stages of Student Departure: Reflections on the Longitudinal Character of Student Leaving. *The Journal of Higher Education*, 59(4), 438. <https://doi.org/10.2307/1981920>
- Trauten, F. (2018, 18. Juli). *Entwicklung und Evaluation von feedbackgestützten Online-Aufgaben im Fach Chemie*. Masterarbeit. Universität Duisburg-Essen, Essen. Zugriff am 26.04.2022.
- Trauten, F., Eitemüller, C. & Walpuski, M. (2020). Evaluation adaptiven Feedbacks in Online-Aufgaben in der Chemie. In S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Wien 2019* (S. 884–887). Essen: Gesellschaft für Didaktik der Chemie und Physik.
- Trauten, F., Eitemüller, C. & Walpuski, M. (2023). Lernaufgaben mit automatisiertem Feedback in einem digitalen Hochschultutorium zur Vorlesung der Allgemeinen Chemie als Angebot zur Binnendifferenzierung. *ZeHf – Zeitschrift für empirische Hochschulforschung*, 7(1-2023), 13–33. <https://doi.org/10.3224/zehf.v7i1.03>
- Trautwein, C. & Bosse, E. (2017). The first year in higher education—critical requirements from the student perspective. *Higher Education*, 73(3), 371–387. <https://doi.org/10.1007/s10734-016-0098-5>

- Walpuski, M., Kauertz, A., Kampa, N., Fischer, H. E., Mayer, J., Sumfleth, E. et al. (2010). ESNaS. Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I. In A. Gehrmann (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 171–184). Bad Heilbrunn: Klinkhardt.
- Walter, O. (2005). *Kompetenzmessung in den PISA-Studien. Simulationen zur Schätzung von Verteilungsparametern und Reliabilitäten*. Kiel, Univ., Diss., 2005. Lengerich: Pabst.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450. <https://doi.org/10.1007/BF02294627>
- Watzlawick, P., Bavelas, J. B. & Jackson, D. D. (1967). *Pragmatics of human communication. A study of interactional patterns, pathologies, and paradoxes*. New York: Norton.
- Weedon, E. & Riddell, S. (2016). Widening Participation. In M. Shah, A. Bennett, E. Southgate & A. Badarneh (Eds.), *Widening higher education participation. A global perspective* (1st ed., S. 49–61). Amsterdam, Netherlands: Chandos Publishing.
- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, (15), 185–200.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636. <https://doi.org/10.3758/bf03196322>
- Winter, M. (2009). *Das neue Studieren. Chancen, Risiken, Nebenwirkungen der Studienstrukturenreform: Zwischenbilanz zum Bologna-Prozess in Deutschland (HoF-Arbeitsbericht 1/2009)*, Institut für Hochschulforschung. Zugriff am 08.09.2021. Verfügbar unter: https://www.hof.uni-halle.de/dateien/ab_1_2009.pdf
- Wisniewski, B., Zierer, K. & Hattie, J. (2019). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10, 1–14. <https://doi.org/10.3389/fpsyg.2019.03087>
- Wissenschaftsrat. (2015). *Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt. Zweiter Teil der Empfehlung zur Qualifizierung von Fachkräften vor dem Hintergrund des demographischen Wandels (Drs. 4925-15)*.
- Woolfolk, A. (2014). *Pädagogische Psychologie* (12., aktualisierte Auflage). Pearson. Verfügbar unter: <https://www.pearson-studium.de/drm/reader/fr/usr/28084/isbn/9783863265328>
- Wu, H.-K. & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, 88(3), 465–492. <https://doi.org/10.1002/sce.10126>

Literaturverzeichnis

Young, K. R., Schaffer, H. E., James, J. B. & Gallardo-Williams, M. T. (2021). Tired of Failing Students? Improving Student Learning Using Detailed and Automated Individualized Feedback in a Large Introductory Science Course. *Innovative Higher Education*, 46(2), 133–151. <https://doi.org/10.1007/s10755-020-09527-5>

Anhang

1. Skalendokumentation der eingesetzten Tests

1.1. Mathematisches Wissen

Skalen (Itemanzahl)	MAT_wle_t1 (23/25*) MAT_wle_t3 (26)
Label	MAT_wle_t1: Personenfähigkeit Mathe MZP1 Gesamtstichprobe MAT_wle_t3: Personenfähigkeit Mathe MZP3 Gesamtstichprobe
Kategorien	0: Falsch/Fehlend 1: Richtig
Quelle	Müller, J., Stender, A., Fleischer, J., Borowski, A., Dammann, E., Lang, M., & Fischer, H. E. (2018). Mathematisches Wissen von Studienanfängern und Studienerfolg. <i>Zeitschrift für Didaktik der Naturwissenschaften</i> , 24(1), 183-199.

Variable	Skalenwerte			
	Lösungshäufigkeit T1	Lösungshäufigkeit T3	Itemparameter T1	Itemparameter T3
MAT_S01	0.83		-3.359 ¹	
MAT_S02	0.84		-3.149 ¹	
MAT_S03	0.68	0.81	-1.408 ¹	²
MAT_S04	0.65		-1.705 ¹	
MAT_S05	0.20	0.34	0.874 ¹	²
MAT_S06	0.39		-0.013 ¹	
MAT_S07	0.28	0.42	0.710 ¹	²
MAT_S08	0.38	0.58	-0.065 ¹	²
MAT_S09	0.32	0.47	-0.097 ¹	²
MAT_S10	0.24	0.41	1.021 ¹	²
MAT_S11	0.43	0.55	-0.225 ¹	²
MAT_S12	0.36	0.45	-0.026 ¹	²
MAT_S13	0.79	0.82	-2.792 ¹	-1.899
MAT_S14	0.63	0.69	-1.769 ¹	
MAT_S15	0.55	0.47	-1.112 ¹	²
MAT_S16	0.32	0.28	0.150 ¹	²
MAT_S17	0.15	0.20	1.277 ¹	²
MAT_S18	0.11	0.94	1.949 ¹	²
MAT_S19	0.83	0.20	-3.102 ¹	²

Anhang

MAT_S20	0.08	0.26	2.303	²
MAT_S21	0.09	0.24	2.100 ¹	²
MAT_S22	0.11	0.53	1.971 ¹	²
MAT_S23	0.33	0.65	0.237 ¹	²
MAT_P07	0.418	0.80	-1.269 ¹	²
MAT_P16	0.621	0.39	-1.574 ¹	²
MAT_P17	0.230	0.08	1.447 ¹	²
MAT_P21	0.022	0.14	3.233 ¹	²
MAT_V04	0.025	0.09	3.869	2.617 ²
MAT_V05	0.006	0.14	5.239	3.183
MAT_V15	0.039	0.81	3.604	

2.617

Skalenwerte WLE Reliability = 0.842
MAT_wle_t1 EAP Reliability = 0.860
Item INFIT (Range) = 0.808 – 1.241
WLE M = -0.872
WLE Var = 2.512
WLE min = -5.54, WLE max = 3.47

Skalenwerte WLE Reliability = 0.860
MAT_wle_t3 EAP Reliability = 0.872
Item INFIT (Range) = 0.762 – 1.199
WLE M = 0.173
WLE Var = 2.348
WLE min = -4.74, WLE max = 4.41

Anmerkungen. * Die Items MAT_S01, MAT_S02, MAT_S04, MAT_S06, MAT_S14 wurden zu T1 nur von Studierenden mit Studienbeginn WiSe 2018/19 bearbeitet. Die Items MAT_P07, MAT_P16, MAT_P17, MAT_P21, MAT_V04, MAT_V05 und MAT_V15 wurden zu T1 nur von Studierenden mit Studienbeginn WiSe 2019/20 bearbeitet. Dementsprechend bearbeiteten Studierende mit Studienbeginn WiSe 2018/19 zu T1 23 Items und Studierende mit Studienbeginn 2019/20 zu T1 25 Items.

¹ Wurde auf in Pilotierung in früheren Projektphase bestimmte Schwierigkeit fixiert.

² Wurde auf in MZP1 bestimmte Schwierigkeit fixiert.

1.2. Fachwissen Chemie

Skalen	(Itemanzahl)	ALC_wle_t1 (35)
		ALC_wle_t3 (35)
Label		ALC_wle_t1: Personenfähigkeit Chemie MZP1Gesamtstichprobe
		ALC_wle_t3: Personenfähigkeit Chemie MZP3 Gesamtstichprobe
Kategorien		0: Falsch/Fehlend
		1: Richtig
Quelle		Daniel Averbeck (2021). Zum Studienerfolg in der Studieneingangsphase des Chemiestudiums - Der Einfluss kognitiver und affektiv-motivationaler Variablen. Berlin: Logos

Variable	Skalenwerte			
	Lösungshäufigkeit	Lösungshäufigkeit	Itemparameter T1	Itemparameter T3
	T1	T3		
ALC01_C_k	0.836	0.874	-1.958 ¹	-1.336
ALC02_C_k	0.596	0.734	-0.433 ¹	²
ALC03_k	0.433	0.601	0.603 ¹	²
ALC04_C_k	0.325	0.506	1.251 ¹	²
ALC05_C_k	0.363	0.446	0.393 ¹	²
ALC06_C_k	0.475	0.705	-0.219 ¹	²
ALC07_C_k	0.593	0.598	-0.341	0.376
ALC08_k	0.196	0.327	1.567 ¹	²
ALC09_EC_k	0.308	0.690	0.824	²
ALC10_C_k	0.597	0.765	-0.447 ¹	²
ALC11_E_k	0.455	0.789	0.168	²
ALC12_k	0.758	0.918	-1.305	-1.846
ALC13_C_k	0.173	0.305	2.044 ¹	²
ALC14_C_k	0.545	0.707	-0.013 ¹	²
ALC15_C_k	0.530	0.639	-0.231	²
ALC16_C_k	0.299	0.322	0.117 ¹	1.537
ALC17_k	0.679	0.850	-0.567 ¹	²
ALC18_C_k	0.305	0.429	0.649 ¹	²
ALC19_k	0.681	0.918	-1.409 ¹	²
ALC20_C_k	0.522	0.659	-0.113 ¹	²
ALC21_C_k	0.328	0.406	0.151 ¹	1.139
ALC22_k	0.578	0.763	-0.214 ¹	²
ALC23_EC_k	0.543	0.828	-0.131	²
ALC24_C_k	0.769	0.877	-1.060 ¹	²
ALC25_C_k	0.331	0.392	0.290 ¹	1.241
ALC26_C_k	0.497	0.722	-0.374 ¹	²
ALC27_EC_k	0.372	0.562	0.439	²
ALC28_E_k	0.821	0.903	-1.702	²

Anhang

ALC29_k	0.264	0.496	0.521 ¹	²
ALC30_k	0.679	0.828	-0.658 ¹	²
ALC31_C_k	0.337	0.637	0.854 ¹	²
ALC32_k	0.690	0.748	-0.627 ¹	²
ALC33_k	0.691	0.901	-0.658 ¹	-1.545
ALC34_C_k	0.751	0.714	-0.239 ¹	²
ALC35_k	0.064	0.286	2.892	1.920

Skalenwerte WLE Reliability = 0.752
 ALC_wle_t1 EAP Reliability = 0.765
 Item INFIT (Range) = 0.812 – 0.193
 WLE M = -0.004
 WLE Var = 0.705
 WLE min = -2.57, WLE max = 2.64

Skalenwerte WLE Reliability = 0.789
 ALC_wle_t3 EAP Reliability = 0.802
 Item INFIT (Range) = 0.763 – 1.275
 WLE M = 0.854
 WLE Var = 0.850
 WLE min = -1.36, WLE max = 3.67

Anmerkungen.¹ Wurde auf in Pilotierung in früheren Projektphase bestimmte Schwierigkeit fixiert.

² Wurde auf in MZP1 bestimmte Schwierigkeit fixiert.

1.3. Kognitive Fähigkeiten

Skala	MZP_KFT
Label	MZP an dem KFT erhoben wurde
Kategorien	1: MZP 1 3: MZP 3

Anmerkungen. Diese Daten wurden nicht von den Pbn ausgefüllt, sondern es handelt sich um technische Daten.

Der Erhebungszeitpunkt war abhängig von der Stichprobe.

Skalen (Itemanzahl)	KFTF_wle (25)
Label	Personenfähigkeit KFTF Gesamtstichprobe
Kategorien	0: Falsch/Fehlend 1: Richtig
Quelle	Heller, K. A. & Perleth, C. (2000). <i>Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R)</i> . Göttingen: Beltz Test.

Variable	Skalenwerte	
	Lösungshäufigkeit	Itemparameter
KFTN1_01_k	0.929	-2,979
KFTN1_02_k	0.733	-1,146
KFTN1_03_k	0.706	-1,188
KFTN1_04_k	0.872	-2.299
KFTN1_05_k	0.903	-2.467
KFTN1_06_k	0.654	-0.596
KFTN1_07_k	0.771	-1.315
KFTN1_08_k	0.737	-1.064
KFTN1_09_k	0.801	-1.517
KFTN1_10_k	0.702	-1.019
KFTN1_11_k	0.659	-0.598
KFTN1_12_k	0.706	-0.777
KFTN1_13_k	0.423	0.331
KFTN1_14_k	0.790	-1.644
KFTN1_15_k	0.637	-0.617
KFTN1_16_k	0.664	-0.642
KFTN1_17_k	0.420	0.410
KFTN1_18_k	0.531	0.083
KFTN1_19_k	0.442	0.554
KFTN1_20_k	0.337	1.034

Anhang

KFTN1_21_k	0.586	-0.033
KFTN1_22_k	0.324	1.080
KFTN1_23_k	0.379	0.967
KFTN1_24_k	0.340	1.071
KFTN1_25_k	0.186	1.976

Skalenwerte KFTF

WLE Reliability = 0.760

EAP Reliability = 0.772

Item INFIT (Range) = 0.918 – 1.157

WLE M = -0.005

WLE Var = 1.056

WLE min = -3.93, WLE max = 4.14

2. Bonuspunktregelung

Hinweise und Erläuterung zum Bonussystem (Aushang in Moodle)

Sehr geehrte Studierende,

in diesem Semester können Sie durch die Bearbeitung von Online- und Offline Aufgaben in der Allgemeinen Chemie Übung Bonuspunkte für die „Allgemeine Chemie“ Klausur sammeln und Inhalte üben und wiederholen.

Für den Erhalt von Bonuspunkten gelten die folgenden Regeln:

Sie können mit den *Online-Aufgaben* insgesamt 5 Bonuspunkte sammeln.

$$\left(\frac{\text{Anzahl korrekt gelöster Testaufgaben}}{\text{Anzahl aller Testaufgaben}} \right) \cdot 5 = \text{Online-Bonuspunkte}$$

(gerundet auf ganze Punkte)

Sie können mit den *Offline-Aufgaben* insgesamt 5 Bonuspunkte sammeln.

$$\left(\frac{\sum(\text{Ihre Punktzahl der Übungsblätter})}{\sum(\text{Gesamtpunktzahl der Übungsblätter})} \right) \cdot 5 = \text{Online-Bonuspunkte}$$

(gerundet auf ganze Punkte)

Sie können also insgesamt 10 Bonuspunkte bei dieser Übung verdienen. Diese werden zu Ihrer Punktzahl in der Allgemeinen Chemie Klausur addiert, wenn Ihre Klausur bestanden ist (d.h. dass Sie mindestens 50 von 100 Punkten selbstständig erreicht haben).

1. Online-Aufgaben

Die *Online-Aufgaben* sind im E-Assessmentsystem JACK realisiert. Die JACK-Aufgaben sind für Sie  1.1 Avogadrokonstante bequem in Moodle anzusteuern. JACK-Elemente erkennen Sie an grünen Puzzle-Teilen wie diesem:

Es öffnet sich dann ein neues Fenster oder ein neuer Tab mit der Aufgabe in JACK. Die Einstiegsseite einer Aufgabe (Abb.1) zeigt Ihnen noch einmal, um welche Aufgabe es sich handelt. Durch das Klicken von „Einreichen“ können Sie mit der jeweiligen Aufgabe beginnen.

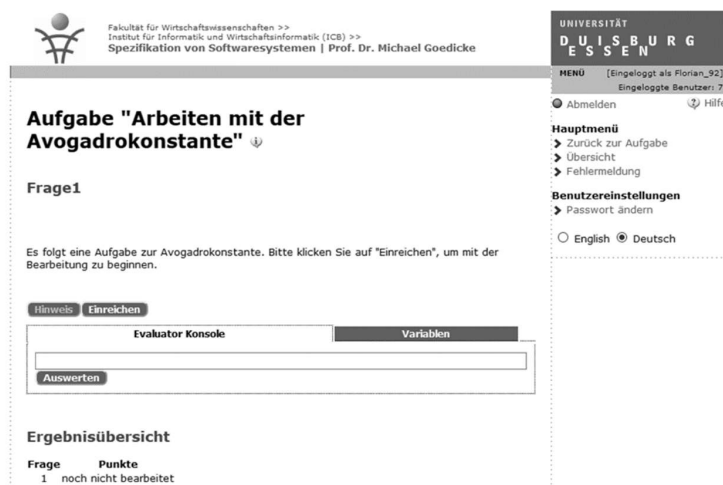


Abbildung 1: Startseite einer JACK-Aufgabe

Ein Aufgabenset in JACK enthält immer sechs sich dynamisch verändernde Lernaufgaben und eine Testaufgabe. Nach Beantworten einer Lernaufgabe ist es wichtig, dass Sie das JACK Fenster schließen und den Tab mit dem Moodle-Kurs aktualisieren (z. B. mit F5). Sie erhalten dann eine neue Lernaufgabe zu diesem Themengebiet. Wenn Sie zwei Lernaufgaben in Folge mit 50 Punkten oder eine erhebliche Anzahl an Aufgaben mit weniger Punkten bearbeiten, gelangen Sie zur Testaufgabe.

WICHTIG: Die Testaufgabe können Sie nur einmal bearbeiten! Jede Lernaufgabe können Sie beliebig oft beantworten.

2. Offline-Aufgaben

Die *Offline-Aufgaben* sind für Sie ebenfalls in Moodle hinterlegt. Diese ist ebenfalls digital einzureichen (bitte als .pdf-Datei). Jedes „Übungsblatt“ hat eine eigene Rubrik im Moodle-Kurs, in der Sie das Aufgabenblatt „Offline Übung X“, eine digitale  Abgabefunktion

und ggf. MATERIAL finden.

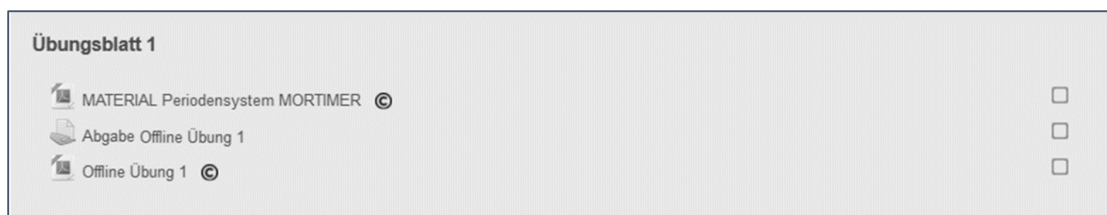


Abbildung 35: „Übungsblatt“ in Moodle

WICHTIG!

Sie haben für jede Offline-Übung **eine Woche Bearbeitungszeit**. Es wird jede Woche eine neue Übung für Sie freigeschaltet. Die **Abgabe** muss **bis spätestens 22 Uhr des darauffolgenden Donnerstags** erfolgt sein.

4. Weitere Fragen

Stellen Sie Ihre Fragen zur Berechnung der Bonuspunkte, zu einzelnen Ansichten oder zu JACK generell an Florian Trauten per Mail:

jack-chemiedidaktik@uni-due.de

Sollten Sie eine Frage haben, die potentiell weitere Studierende interessieren könnte, zögern Sie nicht, Ihre Frage im extra eingerichteten Forum in Moodle zu stellen.

Wir wünschen Ihnen produktive Übungserlebnisse und viel Erfolg!

Bisher erschienene Bände der Reihe
Studien zum Physik- und Chemielernen

ISSN 1614-8967

Vollständige Übersicht auf unserer Website



<https://www.logos-verlag.de/spcl>

Aktuelle Bände

- 300 Amany Annaggar (2020): A Design Framework for Video Game-Based Gamification Elements to Assess Problem-solving Competence in Chemistry Education
ISBN 978-3-8325-5150-6 52.00 EUR
- 301 Alexander Engl (2020): CHEMIE PUR – Unterrichten in der Natur. Entwicklung und Evaluation eines kontextorientierten Unterrichtskonzepts im Bereich Outdoor Education zur Änderung der Einstellung zu „Chemie und Natur“
ISBN 978-3-8325-5174-2 59.00 EUR (open access)
- 302 Christin Marie Sajons (2020): Kognitive und motivationale Dynamik in Schülerlaboren. Kontextualisierung, Problemorientierung und Autonomieunterstützung der didaktischen Struktur analysieren und weiterentwickeln
ISBN 978-3-8325-5155-1 56.00 EUR (open access)
- 303 Philipp Bitzenbauer (2020): Quantenoptik an Schulen. Studie im Mixed-Methods Design zur Evaluation des Erlanger Unterrichtskonzepts zur Quantenoptik
ISBN 978-3-8325-5123-0 59.00 EUR (open access)
- 304 Malte Ubben (2020): Typisierung des Verständnisses mentaler Modelle mittels empirischer Datenerhebung am Beispiel der Quantenphysik
ISBN 978-3-8325-5181-0 43.50 EUR (open access)
- 305 Wiebke Hinrike Kuske-Janßen (2020): Sprachlicher Umgang mit Formeln von LehrerInnen im Physikunterricht am Beispiel des elektrischen Widerstandes in Klassenstufe 8
ISBN 978-3-8325-5183-4 47.50 EUR (open access)
- 306 Kai Bliesmer (2020): Physik der Küste für außerschulische Lernorte. Eine Didaktische Rekonstruktion
ISBN 978-3-8325-5190-2 58.00 EUR (open access)
- 307 Nikola Schild (2021): Eignung von domänenspezifischen Studieneingangsvariablen als Prädiktoren für Studienerfolg im Fach und Lehramt Physik
ISBN 978-3-8325-5226-8 42.00 EUR (open access)

- 308 Daniel Averbek (2021): Zum Studienerfolg in der Studieneingangsphase des Chemie-
studiums. Der Einfluss kognitiver und affektiv-motivationaler Variablen
ISBN 978-3-8325-5227-5 51.00 EUR
- 309 Martina Strübe (2021): Modelle und Experimente im Chemieunterricht. Eine Video-
studie zum fachspezifischen Lehrerwissen und -handeln
ISBN 978-3-8325-5245-9 45.50 EUR
- 310 Wolfgang Becker (2021): Auswirkungen unterschiedlicher experimenteller Repräsen-
tationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR
- 311 Marvin Rost (2021): Modelle als Mittel der Erkenntnisgewinnung im Chemieunter-
richt der Sekundarstufe I. Entwicklung und quantitative Dimensionalitätsanalyse eines
Testinstruments aus epistemologischer Perspektive
ISBN 978-3-8325-5256-5 44.00 EUR (open access)
- 312 Christina Kobl (2021): Förderung und Erfassung der Reflexionskompetenz im Fach
Chemie
ISBN 978-3-8325-5259-6 41.00 EUR (open access)
- 313 Ann-Kathrin Beretz (2021): Diagnostische Prozesse von Studierenden des Lehramts.
eine Videostudie in den Fächern Physik und Mathematik
ISBN 978-3-8325-5288-6 45.00 EUR (open access)
- 314 Judith Breuer (2021): Implementierung fachdidaktischer Innovationen durch das An-
gebot materialgestützter Unterrichtskonzeptionen. Fallanalysen zum Nutzungsverhal-
ten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik
ISBN 978-3-8325-5293-0 50.50 EUR (open access)
- 315 Michaela Oettle (2021): Modellierung des Fachwissens von Lehrkräften in der Teil-
chenphysik. Eine Delphi-Studie
ISBN 978-3-8325-5305-0 57.50 EUR (open access)
- 316 Volker Brüggemann (2021): Entwicklung und Pilotierung eines adaptiven Multistage-
Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR (open access)
- 317 Stefan Müller (2021): Die Vorläufigkeit und soziokulturelle Eingebundenheit natur-
wissenschaftlicher Erkenntnisse. Kritische Reflexion, empirische Befunde und fachdi-
daktische Konsequenzen für die Chemielehrer*innenbildung
ISBN 978-3-8325-5343-2 63.00 EUR
- 318 Laurence Müller (2021): Alltagsentscheidungen für den Chemieunterricht erkennen
und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR
- 319 Lars Ehlert (2021): Entwicklung und Evaluation einer Lehrkräftefortbildung zur Pla-
nung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-7 41.50 EUR (open access)

- 320 Florian Seiler (2021): Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR (open access)
- 321 Nadine Boele (2021): Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR
- 322 Franziska Zimmermann (2022): Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR
- 323 Lars-Frederik Weiß (2021): Der Flipped Classroom in der Physik-Lehre. Empirische Untersuchungen in Schule und Hochschule
ISBN 978-3-8325-5418-7 51.00 EUR
- 324 Tilmann Steinmetz (2021): Kumulatives Lehren und Lernen im Lehramtsstudium Physik. Theorie und Evaluation eines Lehrkonzepts
ISBN 978-3-8325-5421-7 51.50 EUR
- 325 Kübra Nur Celik (2022): Entwicklung von chemischem Fachwissen in der Sekundarstufe I. Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich „Umgang mit Fachwissen“
ISBN 978-3-8325-5431-6 55.00 EUR
- 326 Matthias Ungermann (2022): Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR
- 327 Christoph Hoyer (2022): Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR
- 328 Tobias Schüttler (2022): Schülerlabore als interessefördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR
- 329 Christopher Kurth (2022): Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR
- 330 Dagmar Michna (2022): Inklusiver Anfangsunterricht Chemie. Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion
ISBN 978-3-8325-5463-7 49.50 EUR
- 331 Marco Seiter (2022): Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR

- 332 Jörn Hägele (2022): Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion
ISBN 978-3-8325-5476-7 56.50 EUR (open access)
- 333 Erik Heine (2022): Wissenschaftliche Kontroversen im Physikunterricht. Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie
ISBN 978-3-8325-5478-1 48.50 EUR (open access)
- 334 Simon Goertz (2022): Module und Lernzirkel der Plattform FLexKom zur Förderung experimenteller Kompetenzen in der Schulpraxis. Verlauf und Ergebnisse einer Design-Based Research Studie
ISBN 978-3-8325-5494-1 66.50 EUR
- 335 Christina Toschka (2022): Lernen mit Modellexperimenten. Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten
ISBN 978-3-8325-5495-8 50.00 EUR (open access)
- 336 Alina Behrendt (2022): Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR (open access)
- 337 Manuel Daiber (2022): Entwicklung eines Lehrkonzepts für eine elementare Quantenmechanik. Formuliert mit In-Out Symbolen
ISBN 978-3-8325-5507-8 48.50 EUR
- 338 Felix Pawlak (2022): Das Gemeinsame Experimentieren (an-)leiten. Eine qualitative Studie zum chemiespezifischen Classroom-Management
ISBN 978-3-8325-5508-5 46.50 EUR
- 339 Liza Dopatka (2022): Konzeption und Evaluation eines kontextstrukturierten Unterrichtskonzeptes für den Anfangs-Elektrizitätslehreunterricht
ISBN 978-3-8325-5514-6 69.50 EUR
- 340 Arne Bewersdorff (2022): Untersuchung der Effektivität zweier Fortbildungsformate zum Experimentieren mit dem Fokus auf das Unterrichtshandeln
ISBN 978-3-8325-5522-1 39.00 EUR (open access)
- 341 Thomas Christoph Münster (2022): Wie diagnostizieren Studierende des Lehramtes physikbezogene Lernprozesse von Schüler*innen?. Eine Videostudie zur Mechanik
ISBN 978-3-8325-5534-4 44.50 EUR (open access)
- 342 Ines Komor (2022): Förderung des symbolisch-mathematischen Modellverständnisses in der Physikalischen Chemie
ISBN 978-3-8325-5546-7 46.50 EUR

- 343 Verena Petermann (2022): Überzeugungen von Lehrkräften zum Lehren und Lernen von Fachinhalten und Fachmethoden und deren Beziehung zu unterrichtsnahem Handeln
ISBN 978-3-8325-5545-0 47.00 EUR (open access)
- 344 Jana Heinze (2022): Einfluss der sprachlichen Konzeption auf die Einschätzung der Qualität instruktionaler Unterrichtserklärungen im Fach Physik
ISBN 978-3-8325-5553-5 42.00 EUR (open access)
- 345 Jannis Weber (2022): Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich
ISBN 978-3-8325-5566-5 68.00 EUR (open access)
- 346 Fabian Sterzing (2022): Zur Lernwirksamkeit von Erklärvideos in der Physik. Eine Untersuchung in Abhängigkeit von ihrer fachdidaktischen Qualität und ihrem Einbettungsformat
ISBN 978-3-8325-5576-4 52.00 EUR (open access)
- 347 Lars Greitemann (2022): Wirkung des Tablet-Einsatzes im Chemieunterricht der Sekundarstufe I unter besonderer Berücksichtigung von Wissensvermittlung und Wissenssicherung
ISBN 978-3-8325-5580-1 50.00 EUR
- 348 Fabian Poensgen (2022): Diagnose experimenteller Kompetenzen in der laborpraktischen Chemielehrer*innenbildung
ISBN 978-3-8325-5587-0 48.00 EUR
- 349 William Lindlahr (2023): Virtual-Reality-Experimente. Entwicklung und Evaluation eines Konzepts für den forschend-entwickelnden Physikunterricht mit digitalen Medien
ISBN 978-3-8325-5595-5 49.00 EUR
- 350 Bert Schlüter (2023): Teilnahmemotivation und situationales Interesse von Kindern und Eltern im experimentellen Lernsetting KEMIE
ISBN 978-3-8325-5598-6 43.00 EUR
- 351 Katharina Nave (2023): Charakterisierung situativer mentaler Modellkomponenten in der Chemie und die Bildung von Hypothesen. Eine qualitative Studie zur Operationalisierung mentaler Modellkomponenten für den Fachbereich Chemie
ISBN 978-3-8325-5599-3 43.00 EUR
- 352 Anna B. Bauer (2023): Experimentelle Kompetenz Physikstudierender. Entwicklung und erste Erprobung eines performanzorientierten Kompetenzstrukturmodells unter Nutzung qualitativer Methoden
ISBN 978-3-8325-5625-9 47.00 EUR (open access)
- 353 Jan Schröder (2023): Entwicklung eines Performanztests zur Messung der Fähigkeit zur Unterrichtsplanung bei Lehramtsstudierenden im Fach Physik
ISBN 978-3-8325-5655-6 46.50 EUR
- 354 Susanne Gerlach (2023): Aspekte einer Fachdidaktik Körperpflege. Ein Beitrag zur Standardentwicklung
ISBN 978-3-8325-5659-4 45.00 EUR

- 355 Livia Murer (2023): Diagnose experimenteller Kompetenzen beim praktisch-naturwissenschaftlichen Arbeiten. Vergleich verschiedener Methoden und kognitive Validierung eines Testverfahrens
ISBN 978-3-8325-5657-0 41.50 EUR (open access)
- 356 Andrea Maria Schmid (2023): Authentische Kontexte für MINT-Lernumgebungen. Eine zweiteilige Interventionsstudie in den Fachdidaktiken Physik und Technik
ISBN 978-3-8325-5605-1 57.00 EUR (open access)
- 357 Julia Ortmann (2023): Bedeutung und Förderung von Kompetenzen zum naturwissenschaftlichen Denken und Arbeiten in universitären Praktika
ISBN 978-3-8325-5670-9 37.00 EUR (open access)
- 358 Axel-Thilo Prokop (2023): Entwicklung eines Lehr-Lern-Labors zum Thema Radioaktivität. Eine didaktische Rekonstruktion
ISBN 978-3-8325-5671-6 49.50 EUR
- 359 Timo Hackemann (2023): Textverständlichkeit sprachlich variierten physikbezogener Sachtexte
ISBN 978-3-8325-5675-4 41.50 EUR (open access)
- 360 Dennis Dietz (2023): Vernetztes Lernen im fächerdifferenzierten und integrierten naturwissenschaftlichen Unterricht aufgezeigt am Basiskonzept Energie. Eine Studie zur Analyse der Wirksamkeit der Konzeption und Implementation eines schulinternen Curriculums für das Unterrichtsfach „Integrierte Naturwissenschaften 7/8“
ISBN 978-3-8325-5676-1 49.50 EUR
- 361 Ann-Katrin Krebs (2023): Vielfalt im Physikunterricht. Zur Wirkung von Lehrkräftefortbildungen unter Diversitätsaspekten
ISBN 978-3-8325-5672-3 65.50 EUR (open access)
- 362 Simon Kaulhausen (2023): Strukturelle Ursachen für Klausurmisserfolg in Allgemeiner Chemie an der Universität
ISBN 978-3-8325-5699-0 37.50 EUR (open access)
- 363 Julia Eckoldt (2023): Den (Sach-)Unterricht öffnen. Selbstkompetenzen und motivationale Orientierungen von Lehrkräften bei der Implementation einer Innovation untersucht am Beispiel des Freien Explorierens und Experimentierens
ISBN 978-3-8325-5663-1 48.50 EUR (open access)
- 364 Albert Teichrow (2023): Physikalische Modellbildung mit dynamischen Modellen
ISBN 978-3-8325-5710-2 58.50 EUR (open access)
- 365 Sascha Neff (2023): Transfer digitaler Innovationen in die Schulpraxis. Eine explorative Untersuchung zur Förderung der Implementation
ISBN 978-3-8325-5687-7 59.00 EUR (open access)
- 366 Rahel Schmid (2023): Verständnis von Nature of Science-Aspekten und Umgang mit Fehlern von Schüler*innen der Sekundarstufe I. Am Beispiel von digital-basierten Lernprozessen im informellen Lernsetting Smartfeld
ISBN 978-3-8325-5722-5 53.50 EUR (open access)

- 367 Dennis Kirstein (2023): Individuelle Bedingungs- und Risikofaktoren für erfolgreiche Lernprozesse mit kooperativen Experimentieraufgaben im Chemieunterricht. Eine Untersuchung zum Zusammenhang von Lernvoraussetzungen, Lerntätigkeiten, Schwierigkeiten und Lernerfolg beim Experimentieren in Kleingruppen der Sekundarstufe I
ISBN 978-3-8325-5729-4 50.50 EUR (open access)
- 368 Frauke Düwel (2024): Argumentationslinien in Lehr-Lernkontexten. Potenziale englischer Fachtexte zur Chromatografie und deren hochschuldidaktische Einbindung
ISBN 978-3-8325-5731-7 63.00 EUR (open access)
- 369 Fabien Güth (2023): Interessenbasierte Differenzierung mithilfe systematisch variierten Kontextaufgaben im Fach Chemie
ISBN 978-3-8325-5737-9 48.00 EUR (open access)
- 370 Oliver Grewe (2023): Förderung der professionellen Unterrichtswahrnehmung und Selbstwirksamkeitsüberzeugungen hinsichtlich sprachsensibler Maßnahmen im naturwissenschaftlichen Sachunterricht. Konzeption und Evaluation einer video- und praxisbasierten Lehrveranstaltung im Masterstudium
ISBN 978-3-8325-5738-6 44.50 EUR (open access)
- 371 Anna Nowak (2023): Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht. Entwicklung des Reflexionsmodells REIZ
ISBN 978-3-8325-5739-3 59.00 EUR (open access)
- 372 Dominique Angela Holland (2023): Bildung für nachhaltige Entwicklung (BNE) kooperativ gestalten. Vergleich monodisziplinärer und interdisziplinärer Kooperation von Lehramtsstudierenden bei der Planung, Durchführung und Reflexion von Online-BNE-Unterricht
ISBN 978-3-8325-5760-7 47.00 EUR (open access)
- 373 Renan Marcello Vairo Nunes (2024): MINT-Personal an Schulen. Eine Untersuchung der Arbeitssituation und professionellen Kompetenzen von MINT-Lehrkräften verschiedener Ausbildungswege
ISBN 978-3-8325-5778-2 51.00 EUR (open access)
- 374 Mats Kieserling (2024): Digitalisierung im Chemieunterricht. Entwicklung und Evaluation einer experimentellen digitalen Lernumgebung mit universeller Zugänglichkeit
ISBN 978-3-8325-5786-7 45.50 EUR
- 375 Cem Aydin Salim (2024): Die Untersuchung adaptiver Lernsettings im Themenbereich „Schwimmen und Sinken“ im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-5787-4 49.00 EUR (open access)
- 376 Novid Ghassemi (2024): Evaluation eines Lehramtsmasterstudiengangs mit dem Profil Quereinstieg im Fach Physik. Erkenntnisse zu Eingangsbedingungen, professionellen Kompetenzen und Aspekten individueller Angebotsnutzung
ISBN 978-3-8325-5789-8 41.50 EUR (open access)

- 377 Martina Flurina Cavelti (2024): Entwicklung und Validierung eines Messinstruments zur Erfassung der Schülerkompetenzen im Bereich des wissenschaftlichen Skizzierens im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-5829-1 45.00 EUR (open access)
- 378 Tom Bleckmann (2024): Formatives Assessment auf Basis von maschinellem Lernen. Eine Studie über automatisiertes Feedback zu Concept Maps aus dem Bereich Mechanik
ISBN 978-3-8325-5842-0 46.50 EUR (open access)
- 379 Jana Marlies Rehberg (2024): Das physikspezifische Mindset zum Studienbeginn. Fragebogenentwicklung und Aufbau einer Online-Intervention
ISBN 978-3-8325-5850-5 59.50 EUR (open access)
- 380 Florian Trauten (2024): Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie
ISBN 978-3-8325-5859-8 46.00 EUR (open access)
- 381 Johanna Dejanovikj (2024): Reflexion als Lerngegenstand im Unterricht zur Förderung von Bewertungskompetenz
ISBN 978-3-8325-5860-4 41.00 EUR (open access)
- 382 Katharina Flieser (2024): Verständlichkeit physikalischer Sachtexte. Untersuchungen zum Wirkungsgefüge zwischen sprachlicher Textgestaltung und der Behaltensleistung sowie der Textwahrnehmung im Schulfach Physik
ISBN 978-3-8325-5858-1 44.00 EUR (open access)
- 383 Stephanie Neppl (2024): Perspektivenübernahme im Physikunterricht. Explorative Interviewstudie zu einer Seminarkonzeption mit dem Schwerpunkt Perspektivenübernahme bei der Planung von Physikunterricht
ISBN 978-3-8325-5865-9 48.00 EUR (open access)
- 384 Katja Plicht (2024): Ein Physikübungskonzept zur Förderung der Problemlösekompetenz. Entwicklung und empirische Evaluation eines Strategietrainings auf der Basis von Expertisemerkmalen
ISBN 978-3-8325-5875-8 45.00 EUR (open access)

Vollständige Übersicht unter: <https://www.logos-verlag.de/spcl>

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder telefonisch (030 - 42 85 10 90) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung im deutschsprachigen Raum.

Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Univ.-Prof. Dr. Martin Hopf
Universität Wien,
Österreichisches Kompetenzzentrum
für Didaktik der Physik,
Porzellangasse 4, Stiege 2,
1090 Wien, Österreich,
Tel. +43-1-4277-60330,
e-mail: martin.hopf@univie.ac.at

Prof. Dr. Mathias Ropohl
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen,
Tel. 0201-183 2704,
e-mail: mathias.ropohl@uni-due.de

Ein Großteil der Studienabbrüche im Fach Chemie an deutschen Hochschulen (Abbruchquote 52 %; Stand 2022) lässt sich auf Lern- und darauf folgende Leistungsprobleme zurückführen. Mit digitalen, feedbackgestützten Lernaufgaben wurde angestrebt, ein binnendifferenzierendes Förderangebot zu erstellen und zu evaluieren, um Wissensdefizite bereits im ersten Semester parallel zur Vorlesung in der Allgemeinen Chemie zu adressieren. Gleichzeitig soll ein Beitrag zur Feedbackforschung geleistet werden, die hinsichtlich des Einflusses von Vorwissen auf die Lernwirksamkeit unterschiedlicher Feedbacktypen bislang keine eindeutige Ergebnislage aufweist.

In einem Prä-Post-Vergleichsgruppendesign wurde die Wirkung zweier Feedbacktypen (fehlerspezifisches tutorielles Feedback und korrekatives Feedback) untersucht. Die beiden Gruppen ($N=122$) wurden untereinander und mit einer Kontrollkohorte ($N=105$) verglichen.

Das Lernen mit fehlerspezifischem tutoriellem Feedback ist für alle Studierenden lernwirksam. Es ermöglicht aber Studierenden mit zu Studienbeginn niedrigerem Vorwissen Wissenslücken zu schließen und so gegenüber den Studierenden mit hohem Vorwissen bis zum Semesterende aufzuholen. Rein korrekatives Feedback zeigte unter keiner Bedingung einen Vorteil gegenüber elaborem Feedback. Die Arbeit liefert damit ein Beispiel für wirksame automatisierte fehlerspezifische Hilfestellungen im webbasierten Übungsbetrieb der Hochschullehre.

Logos Verlag Berlin

ISBN 978-3-8325-5859-8