

Lohe, Viviane [Hrsg.]; Lindl, Alfred [Hrsg.]; Kirchhoff, Petra [Hrsg.]
Unterrichtsqualität in schulischen Fremdsprachen. Theoretische Ansätze und empirische Ergebnisse aus den Fachdidaktiken

Münster ; New York : Waxmann 2024, 230 S.



Quellenangabe/ Reference:

Lohe, Viviane [Hrsg.]; Lindl, Alfred [Hrsg.]; Kirchhoff, Petra [Hrsg.]: Unterrichtsqualität in schulischen Fremdsprachen. Theoretische Ansätze und empirische Ergebnisse aus den Fachdidaktiken. Münster ; New York : Waxmann 2024, 230 S. - URN: urn:nbn:de:0111-pedocs-321589 - DOI: 10.25656/01:32158; 10.31244/9783830999201

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-321589>

<https://doi.org/10.25656/01:32158>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-Licence: <http://creativecommons.org/licenses/by/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

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

Viviane Lohe, Alfred Lindl,
Petra Kirchhoff (Hrsg.)

Unterrichtsqualität in schulischen Fremdsprachen

Theoretische Ansätze und empirische Ergebnisse
aus den Fachdidaktiken

Englisch

Französisch

Italienisch

Latein

Portugiesisch

Russisch

Spanisch

WAXMANN

Viviane Lohe, Alfred Lindl,
Petra Kirchhoff (Hrsg.)

Unterrichtsqualität in schulischen Fremdsprachen

Theoretische Ansätze und empirische
Ergebnisse aus den Fachdidaktiken



Waxmann 2024
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

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

Print-ISBN 978-3-8309-4920-6
E-Book-ISBN 978-3-8309-9920-1
<https://doi.org/10.31244/9783830999201>

Waxmann Verlag GmbH, Münster 2024
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Pleßmann Design, Ascheberg
Satz: Roger Stoddart, Münster

Dieses E-Book ist unter der Lizenz CC BY 4.0 open access verfügbar.



Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Zitate etc.) sind von der CC-Lizenz ausgenommen und für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen.

Die freie Verfügbarkeit der E-Book-Ausgabe dieser Publikation wurde ermöglicht durch den Fachinformationsdienst Erziehungswissenschaft und Bildungsforschung, gefördert durch die Deutsche Forschungsgemeinschaft und ein Netzwerk wissenschaftlicher Bibliotheken zur Förderung von Open Access in der Erziehungswissenschaft, Bildungsforschung und Fachdidaktik.

Bibliothek der Berufsakademie Sachsen
Bibliothek der Hochschule für Bildende Künste Braunschweig
Bibliothek für Bildungsgeschichtliche Forschung des DIPF Berlin
Bibliotheks- und Informationssystem (BIS) der Carl von Ossietzky Universität
Oldenburg
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Evangelische Hochschule Dresden
FernUniversität in Hagen - Universitätsbibliothek
Freie Universität Berlin – Universitätsbibliothek
Gottfried Wilhelm Leibniz Bibliothek - Niedersächsische Landesbibliothek Hannover
Hochschulbibliothek der Pädagogischen Hochschule Karlsruhe
Hochschule für Bildende Künste Dresden
Hochschule für Grafik und Buchkunst Leipzig
Hochschule für Musik Dresden
Hochschule für Musik und Theater Leipzig
Hochschule für Technik, Wirtschaft und Kultur Leipzig
Hochschule für Technik und Wirtschaft Dresden
Hochschule Mittweida
Hochschule Zittau / Görlitz
Humboldt-Universität zu Berlin Universitätsbibliothek
Landesbibliothek Oldenburg
Leibniz-Institut für Bildungsmedien | Georg-Eckert-Institut Braunschweig
Medien- und Informationszentrum / Leuphana Universität Lüneburg
Pädagogische Hochschule Freiburg
Pädagogische Hochschule Heidelberg
Palucca-Hochschule für Tanz Dresden
Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
Staats- und Universitätsbibliothek Hamburg
Technische Informationsbibliothek (TIB) Hannover
Technische Universität Berlin / Universitätsbibliothek
Technische Universität Chemnitz
Technische Universitätsbibliothek Hamburg
Universitätsbibliothek Leipzig
Universitäts- und Landesbibliothek Bonn
Universitäts- und Landesbibliothek Münster
Universitäts- und Stadtbibliothek Köln
Universitätsbibliothek Augsburg
Universitätsbibliothek Bielefeld
Universitätsbibliothek Bochum
Universitätsbibliothek der LMU München

Universitätsbibliothek der Technischen Universität Hamburg
Universitätsbibliothek der TU Bergakademie Freiberg
Universitätsbibliothek Duisburg-Essen
Universitätsbibliothek Erlangen-Nürnberg
Universitätsbibliothek Gießen
Universitätsbibliothek Hildesheim
Universitätsbibliothek Johann Christian Senckenberg / Frankfurt a.M.
Universitätsbibliothek Kassel
Universitätsbibliothek Leipzig
Universitätsbibliothek Mainz
Universitätsbibliothek Mannheim
Universitätsbibliothek Marburg
Universitätsbibliothek Passau
Universitätsbibliothek Potsdam
Universitätsbibliothek Regensburg
Universitätsbibliothek Rostock
Universitätsbibliothek Vechta
Universitätsbibliothek Wuppertal
Universitätsbibliothek Würzburg
Westfälische Hochschule Zwickau

Inhalt

<i>Alfred Lindl, Viviane Lohe und Petra Kirchhoff</i> Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen Grundlagen und Forschungsansätze	9
<i>Armin Jentsch und Petra Kirchhoff</i> Was bedeutet Fachspezifität? Überlegungen und Beispiele zu fächerübergreifender sowie fachdidaktisch konkretisierter Unterrichtsqualitätsforschung	39
<i>Joel Guttke und Raphaela Porsch</i> Inhaltsvalidierung von Fragebogenitems zu kognitiver Aktivierung im Englischunterricht der Grundschule mittels Expert:innenbefragung	57
<i>Manuela Franke</i> Wie setzen Lehrkräfte das Lehrwerk im Französisch- und Spanischunterricht ein? Eine Studie zu Unterrichtsqualität in der Sekundarstufe I mit Fokus auf Sozialform, Binnendifferenzierung und Feedback	85
<i>Maria Hofrichter, Alfred Lindl, Mara Rader und Annika Krämer</i> Was macht guten Fremdsprachenunterricht aus? Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern Englisch und Latein	109
<i>Daniel Reimann</i> Unterrichtsqualität in den romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen (Französisch, Spanisch, Italienisch, Portugiesisch).	139
<i>Sonja Brunsmeier</i> Zusammenspiel von generischen und fachspezifischen Aspekten Ein Konzept für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn	157
<i>Dennis Tark</i> Der schulische Fremdsprachenunterricht Russisch: Unterrichtsqualität und Fachlichkeit	183

Jens-Folkert Folkerts

Wer Ohren hat zu hören

Qualitätskriterien für guten Hörverstehensunterricht

aus der Perspektive von Englischfachleitungen 201

Autor:innen des Bandes 229

Alfred Lindl, Viviane Lohe und Petra Kirchhoff

Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen

Grundlagen und Forschungsansätze

Abstract

In the introductory article by Alfred Lindl, Viviane Lohe and Petra Kirchhoff “Subject-specific instructional quality in foreign languages at school – Basics and research approaches,” the authors first address the concept of instructional quality as an interdisciplinary research endeavour in both educational science and subject-matter didactics. In addition to the existing theoretical approaches, which consider the subject-specific characteristics of the respective foreign languages discussed in the individual contributions, the article also introduces the perspective of educational science. Basic concepts of research on instructional quality, its background and state of the art as well as central models and methods are explained. This is followed by an outline of how this field of research has been received in foreign language didactics so far, with a focus on the subject of English. On this basis, some further research questions and gaps with regard to foreign language teaching are formulated with a view to the future. The chapter concludes with a brief overview of the contributions in this volume.

1 Unterrichtsqualität als interdisziplinäre Forschungsaufgabe

Was macht qualitätvollen Unterricht aus? Die Leitfragestellung dieses Bandes verweist auf eine zentrale Forschungsaufgabe von Bildungswissenschaft wie Fachdidaktik, deren Bewältigung eines regen Austauschs und engen Zusammenwirkens dieser beiden Disziplinen bedarf (z. B. Praetorius & Gräsel, 2021; Reusser & Pauli, 2021). Das fruchtbare Potenzial zur Weiterentwicklung dieses wichtigen Forschungsbereichs, das eine derartige Kooperation und ein wechselseitiger Erkenntnistransfer birgt, wie auch die Erfordernis zu dessen kontinuierlicher theoretischer und methodischer Innovation angesichts unterschiedlicher gesellschaftlicher Herausforderungen, vor die Lehrkräfte im Unterricht täglich gestellt sind, dokumentieren zahlreiche Publikationen und Sonderhefte in einschlägigen Fachzeitschriften, die hierzu in den letzten Jahren erschienen sind (z. B. Charalambous & Praetorius, 2018; Praetorius & Gräsel, 2021; Praetorius & Kleickmann, 2022; Praetorius & Nehring, 2020). Die darin enthaltenen Beiträge lassen jedoch eine einseitige Schwerpunktsetzung der Unterrichtsqualitätsforschung auf Mathematik und Naturwissenschaften erkennen (Helmke, 2022; Praetorius et al., 2018, 2020a;

Reusser, 2020). Die sprachlichen Schulfächer (z. B. Deutsch) und insbesondere die schulischen Fremdsprachen (z. B. Englisch, Französisch, Latein, Spanisch) stehen hingegen vergleichsweise selten im Fokus der Betrachtung und sind bislang wenig untersucht (z. B. Lindl & Krämer, 2022; Lindl et al., eingereicht; Wilden, 2021; Wiprächtiger-Geppert et al., 2021). Die Gründe hierfür sind vielfältig und sowohl in den bildungspolitischen und -pragmatischen Umständen wie einer prävalenten Beteiligung mathematisch-naturwissenschaftlicher Fächer an internationalen Vergleichsstudien als auch fachlichen Eigenheiten und Forschungstraditionen der Sprachfächer zu suchen, die sich von der oftmals quantitativ ausgerichteten, vorwiegend empirisch arbeitenden Bildungsforschung unterscheiden (z. B. Helmke, 2022; Wilden, 2021).

Dieses Forschungsdesiderat und diese -distanz gerade mit Blick auf die schulischen Fremdsprachen aufzuheben, war eine zentrale Zielsetzung des interdisziplinären Symposiums „Unterrichtsqualität in schulischen Fremdsprachen“, das am 26./27. September 2022 an der Universität Erfurt stattfand (Lohe et al., 2022). Daran nahmen Vertreter:innen aus Bildungswissenschaft, -administration und Fachdidaktik, und zwar nicht nur der Fremdsprachen Englisch, Französisch, Italienisch, Latein, Portugiesisch, Russisch und Spanisch, sondern auch weiterer Fächer wie Deutsch und Mathematik teil. Zur Diskussion stand unter anderem, was qualitätsvollen Unterricht in den Fremdsprachen ausmacht und welche neuen Denkansätze und Impulse der fortgeschrittene Forschungsstand, die theoretische Modellbildung und die unterschiedlichen Untersuchungsmethoden der Bildungswissenschaft bezüglich der Analyse von Unterrichtsqualität in den Fremdsprachen bieten können. Eine besondere Rolle spielten dabei auch eine kritische Reflexion und empirische Überprüfung existierender theoretischer Ansätze unter Berücksichtigung von Fachkultur und -spezifik der jeweiligen Fremdsprache. Dies ermöglichte es, auf domänenspezifische Merkmale von Unterrichtsqualität in einer bestimmten oder mehreren Fremdsprachen aufmerksam zu machen, diese sorgfältig zu identifizieren sowie eingehender zu analysieren und somit das Verhältnis zwischen generischen und fachspezifischen Merkmalen von Unterrichtsqualität vor allem aus fremdsprachendidaktischer Sicht auszuloten. Diese intensive Auseinandersetzung ließ drängende und künftige Aufgaben der empirischen Fremdsprachendidaktik deutlich hervortreten, um das wichtige Forschungsgebiet Unterrichtsqualität eigenständig zu erschließen, dauerhaft zu etablieren und voranzubringen. Dies verspricht wesentliche Erkenntnisse über (fach-)spezifische Qualitätsmerkmale und Möglichkeiten zur Optimierung des Lehrens und Lernens von Fremdsprachen in der Schule und eine auf eigener Forschungsaktivität und -position gründende Anschlussfähigkeit zum bildungswissenschaftlichen Diskurs.

Als kleinen Beitrag zu diesem ehrgeizigen Ziel möchte dieser Sammelband die Vielfalt an Forschungsansätzen mit Bezug zu unterschiedlichen Aspekten von Unterrichtsqualität in schulischen Fremdsprachen dokumentieren. Um den inspirierenden Facettenreichtum der Präsentationen und Ergebnisse der oben erwähnten Tagung wiederzugeben, versammelt er heterogene theoretische wie auch empirische Zugänge der verschiedenen Fachdidaktiken unterrichtlicher Fremd-

sprachen, so unter anderem des Englischen, des Lateinischen, des Russischen und der romanischen Sprachen, umfasst aber auch ein Impulsreferat aus der Mathematikdidaktik (Jentsch & Kirchhoff, in diesem Band). Durch den bewusst gewählten Perspektivwechsel deutet dieses die fachübergreifende Relevanz des Forschungszweigs an, weitet es die gedankliche Bezugsnorm und bietet es wertvolle Untersuchungsanregungen, -erfahrungen und -befunde. Das vorliegende Einleitungskapitel bringt zusätzlich die Perspektive der Bildungswissenschaft ein, indem es Grundbegriffe der Unterrichtsqualitätsforschung, deren Hintergründe und aktuellen Stand sowie darin zentrale Modelle und Methoden erläutert (Abschnitte 2 und 3). Daran knüpft eine Skizze über deren bisherige Rezeption in der Fremdsprachendidaktik mit einem Schwerpunkt auf dem Fach Englisch an, auf deren Basis einige weiterführende Forschungsfragen und -desiderate für die schulischen Fremdsprachen formuliert werden (Abschnitte 4.1 und 4.2). Ein summarischer Überblick über die Beiträge in diesem Band schließt das einleitende Kapitel ab (Abschnitt 5). Damit möchte dieses allen Leserinnen und Lesern, die mit Ansätzen und Vorgehensweisen bildungswissenschaftlicher Forschung zu Unterrichtsqualität noch etwas weniger vertraut sind, eine kurze Einführung in die Thematik sowie eine theoretische und methodische Grundlage und Orientierung bieten, die zu einer präziseren Verortung, angemesseneren Beurteilung und einem besseren Verständnis der nachstehenden Beiträge verhilft.

2 Bildungswissenschaftliche Modellierung und Merkmale von Unterrichtsqualität

2.1 Grundbegriffe der Unterrichtsqualitätsforschung

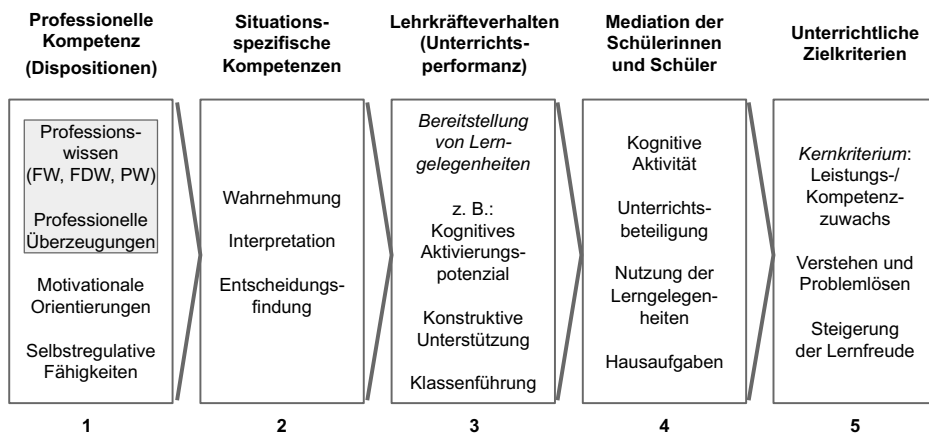
Für eine Untersuchung der Frage, was Unterricht qualitativ macht, gibt es unterschiedliche Ansätze, die oftmals davon abhängen, was unter Qualität zu verstehen ist, welche Prozesse betrachtet und schließlich welche Perspektive und Zielkriterien angelegt werden. Im Anschluss an die theoretische Differenzierung nach Berliner (2005) wird Unterricht bzw. werden dessen Merkmale in der Bildungswissenschaft dann als *qualitätsvoll* erachtet, wenn sie *gut* und *effektiv* sind (Drechsel & Schindler, 2019; Kunter & Trautwein, 2013). Gut ist Unterricht, wenn er normativen, gesellschaftlichen, moralischen etc. Standards entspricht und diese nicht verletzt (z. B. übertriebene disziplinäre Strenge, Androhung von körperlichen Strafen). Als effektiv wird Unterricht angesehen, wenn er bei Schülerinnen und Schülern zu beobachtbaren Lernerfolgen führt, d. h. beispielsweise kognitive oder affektive Lernziele erreicht werden. Falls nur eines der beiden Kriterien zutrifft oder empirisch nachweisbar ist, kann Unterricht also per definitionem nicht als qualitativ gelten.

Betrachtung und Diskussion von Merkmalen guten Unterrichts besitzen in der (inter-)nationalen Bildungswissenschaft eine lange Tradition (z. B. Brophy, 2000; Helmke, 2006; Meyer, 2004). Im deutschsprachigen Raum gewann die-

ser Forschungsbereich allerdings gerade durch den sogenannten PISA-Schock (Programme for International Student Assessment) im Jahr 2000 mit Blick auf die daran beteiligten mathematisch-naturwissenschaftlichen Fächer an Dynamik (z. B. Baumert et al., 2001), die bis heute anhält und innerhalb der Bildungs- einen eigenen Zweig der *Unterrichtsqualitätsforschung* entstehen ließ (Schilcher et al., 2021). Denn aufgrund des mäßigen Abschneidens deutscher Schüler:innen der neunten Jahrgangsstufe im internationalen Vergleich wurden Ursachen für dieses unerwartete und gesellschaftlich nicht zufriedenstellende Ergebnis gesucht. Zudem wurde verstärkt hinterfragt, ob und welche Unterrichtsfaktoren dessen Erfolg wie beeinflussen, inwiefern die Erreichung von dessen Zielkriterien aufseiten der Schüler:innen verbessert werden kann und inwieweit bisher vorwiegend theoretisch abgeleitete Wirkungsannahmen und -ketten empirisch zu belegen sind.

Eine solche idealisierte Wirkungsstruktur beschreibt das (auf Forschungsarbeiten zu Mathematik fußende, aber in allen Fächern anschlussfähige) Kaskadenmodell von Krauss et al. (2020), das in Abbildung 1 dargestellt ist, in komplexitätsreduzierter Weise, indem es einen Fokus auf die von Lehrkräften initiierten Effekte hinsichtlich Unterrichtsqualität und -erfolge legt. Es integriert zum einen verschiedene theoretische Ansätze der Forschung zu (professionellen) Kompetenzen und zu Expertise von Lehrkräften (Säule 1; z. B. Baumert & Kunter, 2006; Bromme, 1992; Krauss et al., 2017; Krauss, 2020; Shulman, 1986), zu deren situationsspezifischen Wahrnehmungs-, Interpretations- und Urteilsfähigkeiten (Säule 2; z. B. Blömeke et al., 2015; Blömeke & Kaiser, 2017) und deren daraus resultierender Unterrichtsperformanz (Säule 3). Zum anderen berücksichtigt das Kaskadenmodell (Abb. 1) auch die mit dem Verhalten der Lehrkräfte einhergehenden Lerngelegenheiten und das daraus entstehende Unterrichtsangebot, das Schüler:innen abhängig von weiteren Einflussfaktoren wie Aufmerksamkeit, Motivation, Vorwissen etc. nutzen können (Säule 4). Im Falle erfolgreicher Mediation kann dies zum Erreichen unterrichtlicher Zielkriterien wie Leistungs- und Kompetenzzuwachs oder Steigerung der Lernfreude führen (Säule 5). Damit bilden die letzten Säulen vereinfacht das ab, was allgemein unter dem Begriff Prozess-(Mediations-)Produkt-Paradigma gefasst (z. B. Krauss, 2020; Lindl et al., 2023) oder vor allem im deutschsprachigen Raum als Angebots-Nutzungsmodell bezeichnet wird (z. B. Helmke, 2022; Ufer & Praetorius, 2022; Vieluf et al., 2020). Während die Aspekte der Nutzung nicht zuletzt aufgrund forschungsmethodischer Limitationen erst seit den letzten Jahren etwas mehr Beachtung erhalten (Säule 4; vgl. z. B. das Sonderheft der Zeitschrift *Unterrichtswissenschaft*; Praetorius & Kleickmann, 2022), fokussiert die Unterrichtsqualitätsforschung traditionell die dritte Säule, nämlich die Bereitstellung eines möglichst qualitätsvollen Unterrichtsangebots durch die Lehrkraft (ohne damit wechselseitige Effekte zwischen Schülerinnen und Schülern sowie Lehrkräften, die zweifelsohne existieren, auszuschließen; z. B. Charalambous & Praetorius, 2020; Helmke, 2022; Vieluf et al., 2020). Ansätze zur Untersuchung, Modellierung und Messung eben dieses Angebots stehen auch in allen Beiträgen des vorliegenden Bandes im Vordergrund.

Das Kaskadenmodell



Säule(n):

Kompetenz- und Kausalmodelle:

- 1: COACTIV-Modell professioneller Lehrkräftekompetenz (grau hervorgehoben: Expertiseparadigma)
- 2: PID: Perception, Interpretation, Decision-Making
- 1, 2, 3: Modell von Kompetenz als Kontinuum
- 3, 4: Kern von Angebots-Nutzungs-Modellen der Unterrichtsqualität
- 3, (4), 5: Prozess-(Mediation-)Produkt-Paradigma

Abbildung 1: Das Kaskadenmodell (Krauss et al., 2020; übers. und zit. nach Lindl et al., 2023)

Zur Beschreibung zentraler Merkmale des Unterrichtsangebots (Säule 3, Abb. 1) wird gemäß bildungswissenschaftlichem Forschungsstand eine weitere wichtige Unterscheidung vorgenommen, nämlich in *Sicht-* und *Tiefenstrukturen* (z. B. Decristan et al., 2020; Drechsel & Schindler, 2019; Kunter & Trautwein, 2013). Elemente der *Sicht-* alias *Oberflächenstruktur* von Unterricht sind (relativ eindeutig und sofort) bei einem kurzen Blick ins Klassenzimmer erkennbar. Hierzu gehören beispielsweise Methoden (z. B. direkte Instruktion, Projektarbeit), Sozial- (z. B. Partner- oder Gruppenarbeit) oder Organisationsformen (z. B. Klassen-, Förderunterricht). Das Erkennen unterrichtlicher *Tiefenstrukturen* oder *Prozessmerkmale* erfordert hingegen ein längeres Verweilen und eine analytische Beobachtung (z. B. Decristan et al., 2020; Drechsel & Schindler, 2019; Kunter & Trautwein, 2013). Eingängige Schlagworte, die mit tiefenstrukturellen Merkmalen assoziiert werden, sind unter anderem (*effektive*) *Klassenführung*, *konstruktive Unterstützung* oder *kognitive Aktivierung* (s. Abschnitt 2.2). Während Oberflächenstrukturen *keine* oder *kaum* Einfluss auf Leistungs- und Kompetenzzuwächse als Kernzielkriterien schulischen Unterrichts besitzen, erweisen sich hierfür Prozessmerkmale der Tiefenebene als besonders relevant (z. B. Hattie, 2009). Dies betrifft insbesondere fachspezifische Lehr- und Lernprozesse (z. B. Blömeke et al., 2022; Kunter et al., 2013; Reusser, 2020; Seidel & Shavelson, 2007), die in diesem Band im Mittelpunkt stehen sollen.

2.2 Die drei „Basisdimensionen“

Die empirische Identifikation derartiger Qualitätsmerkmale gelang erstmals Klieme et al. (2001) mittels (Faktoren-)Analysen auf Basis der Videodaten der TIMS-Studie (Trends in International Mathematics and Science). Die Autor:innen schlussfolgerten daraus drei Beschreibungskategorien für die Qualität von (Mathematik-)Unterricht, die die weitere Forschung in diesem Bereich prägten: (*effektive*) *Klassenführung*, *konstruktive Unterstützung* und *kognitive Aktivierung* (z. B. Hochweber & Vieluf, 2016; Klieme & Rakoczy, 2008; Kunter et al., 2011; zusammenfassend Praetorius et al., 2018; Reusser, 2020). Sie avancierten zu den drei sogenannten *Basisdimensionen von Unterrichtsqualität*, fanden Eingang in die theoretischen Konzeptualisierungen verschiedener Studien (zusammenfassend Praetorius et al., 2018; Rothland, 2024) und dienten als Grundlage weiterentwickelter Modelle zu Unterrichtsqualität (s. u.). So bildeten sie unter anderem den theoretischen Hintergrund der – auch über die Mathematikdidaktik hinaus – viel beachteten und einflussreichen COACTIV-Studie (Cognitive Activation in the Classroom; Kunter et al., 2011; vgl. Säule 3 des Kaskadenmodells in Abb. 1). Darin werden sie wie folgt konzeptualisiert (Kunter & Voss, 2011):

- *Effektive Klassenführung* umfasst einerseits ein souveränes Zeitmanagement, das eine optimale Nutzung der verfügbaren Lernzeit garantiert und Zeitverluste durch nicht lernbezogene Aktivitäten vermeidet. Andererseits beinhaltet sie alle Maßnahmen, die der proaktiven Störungsprävention und unmittelbaren Intervention dienen, um einen möglichst reibungslosen Unterrichtsablauf zu gewährleisten.
- Unter dem Konzept *kognitive Aktivierung* wird nach Kunter und Voss (2011, S. 88) „Unterricht aus einer konstruktivistischen Sicht dann als erfolgreich bezeichnet, wenn ein tiefgehendes Verständnis der Schüler:innen für die fachlichen Inhalte gefördert wird. Lernumgebungen sollten somit so ausgestaltet sein, dass Lernende zum vertieften Nachdenken und zur aktiven mentalen Auseinandersetzung mit den Unterrichtsgegenständen angeregt werden. Aus lernpsychologischer Sicht bedeutet dies, dass bestehende Wissensstrukturen verändert, erweitert, vernetzt, umstrukturiert oder neu gebildet werden sollten.“
- Um einen hohen Grad an kognitiver Aktivierung zu erreichen, ist nach Kunter und Voss (2011) eine *konstruktive Unterstützung* der Schüler:innen unabdingbar, um diesen Autonomie- und Selbstwirksamkeitserfahrungen zu ermöglichen, aber auch beispielsweise bei Verständnisschwierigkeiten oder heterogenen Lernvoraussetzungen eine individuelle, adaptive, differenzierende und wertschätzende Lernumgebung und -unterstützung zu bieten, wozu auch eine aufrichtige Interaktion und vertrauensvolle Beziehung zwischen Lehrenden und Lernenden gehört.

Diese drei Basisdimensionen sowie deren hier exemplarisch vorgestellte Konzeptualisierungen wurden in späteren Studien immer wieder aufgegriffen, angepasst,

ergänzt und erweitert bzw. neuere Modelle damit abgeglichen (z.B. Lipowsky & Bleck, 2019; Kleickmann et al., 2020; Praetorius & Charalambous, 2018; Schlesinger et al., 2018; Wisniewski et al., 2020; zusammenfassend Charalambous & Praetorius, 2018; Praetorius et al., 2018). Dies geschah nicht zuletzt, um die wirkungsvollen fachlich determinierten Lehr- und Lernprozesse zu integrieren und erfassbar zu machen. So schlagen beispielsweise Lipowsky und Bleck (2019) eine Erweiterung der drei Basisdimensionen um eine vierte vor, die sie als fachbezogen ansehen, und implementieren Schlesinger et al. (2018) in ihrer Untersuchung mathematikspezifische Merkmale von Unterrichtsqualität. Kleickmann et al. (2020) plädieren hingegen für eine Ausdifferenzierung der Kategorie der konstruktiven Unterstützung in eine emotionale und eine kognitive Unterstützung, die damit in die Nähe der kognitiven Aktivierung rückt. Wisniewski et al. (2020) fächern jede der drei Basisdimensionen weiter auf und entwickeln unter Rekurs darauf ein eigenes Modell von Unterrichtsqualität mit sieben Dimensionen (Captivation, Care, Challenge, Clarity, Conferment, Consolidation, Control; vgl. Gates Foundation, 2013). Praetorius und Charalambous (2018) sowie Praetorius et al. (2020b) identifizieren in ihrem (erweiterten) Syntheseframework, das auf einer Synthese und Weiterentwicklung von zwölf internationalen Ansätzen zur beobachtungs-basierten Erfassung von Unterrichtsqualität gründet (Ufer & Praetorius, 2022), ebenfalls sieben Qualitätsdimensionen, die bei einem Abgleich allerdings sowohl von der Einteilung bei Wisniewski et al. (2020) als auch von den drei Basisdimensionen abweichen (Praetorius et al., 2020c). Mit diesem Vorschlag möchten die Autorinnen und Autoren bisher unberücksichtigte „blinde Flecken“ (Praetorius et al., 2020c, S. 303) der vorangehenden Modelle füllen. Vor allem fordern sie, wie bereits Schlesinger et al. (2018), Lipowsky und Bleck (2019) sowie Reusser und Pauli (2021), eine stärkere Berücksichtigung domänenspezifischer Merkmale bei der Betrachtung von Unterrichtsqualität. Denn darin erkennen auch sie den dringlichsten Konkretisierungs-, Ergänzungs- und Revisionsbedarf (zusammenfassend Praetorius et al., 2020b).

2.3 Das MAIN-TEACH-Modell als Kristallisations- und Orientierungspunkt

In Fortsetzung dieser Forschungstendenz wurde auf Grundlage des (erweiterten) Syntheseframeworks (Praetorius & Charalambous, 2018; Praetorius et al., 2020b) von Charalambous und Praetorius (2020) das MAIN-TEACH-Modell (Multi-layered and integrated in conceptualizing the quality of teaching; vgl. Abb. 2; Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) entwickelt, das angesichts seines Differenzierungsgrads und seiner Komplexität gewissermaßen einen Gegenpol zum griffig beschreibbaren und empirisch gut untersuchbaren Modell der drei Basisdimensionen darstellt. Aufgrund seiner integrativen Systematisierung bezieht es nicht nur weitere bedeutsame sowie explizit generische und fachspezifische Aspekte von Unterrichtsqualität mit ein bzw. bietet hierzu – insbesondere unter ei-

ner fachlichen Perspektive – Möglichkeiten (Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022). Ein entscheidender Vorteil ist, dass das MAIN-TEACH-Modell im Unterschied zu additiven Merkmallisten (z.B. Brophy, 2000; Helmke, 2022; Meyer, 2004) unterschiedliche Dimensionen von Unterrichtsqualität hinsichtlich ihrer Funktion für den Lernprozess zueinander in Beziehung setzt und damit deren beliebiger Ausdeutung sowie Ergänzung, die unter anderem Helmke kritisiert (2022), entgegenwirkt und auch deren potenzielle Überschneidungen und Wechselwirkungen berücksichtigt (Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022).

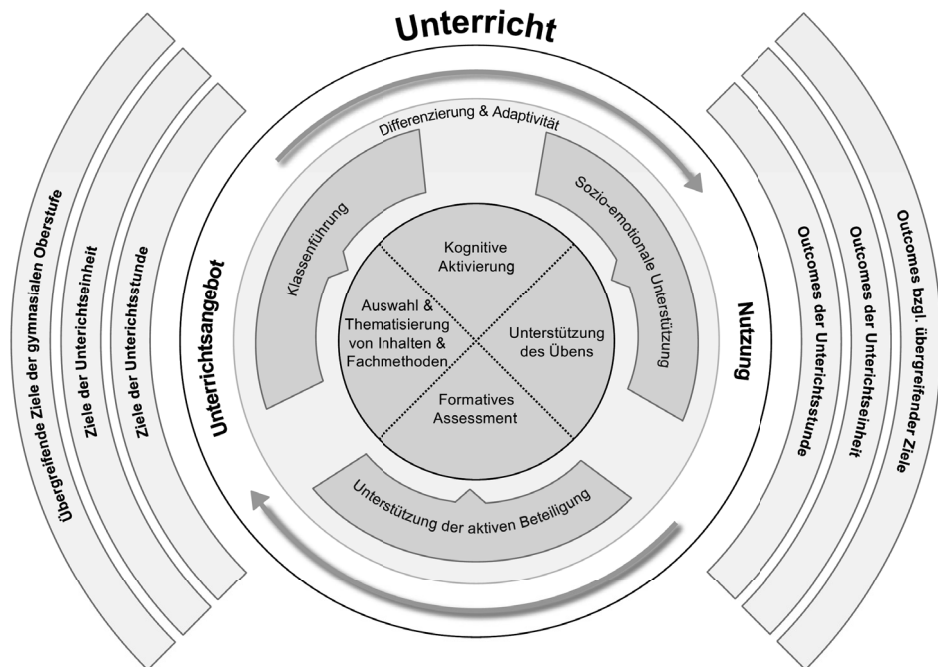


Abbildung 2: Adaptiertes MAIN-TEACH-Modell (Multi-layered and integrated in conceptualizing the quality of teaching; Charalambous & Praetorius, 2020; dt. Abb. nach Ufer & Praetorius, 2022)

Der Kern des Modells, das in Abbildung 2 wiedergegeben ist, besteht aus den drei sich überlagernden, funktional verschiedenen, grau eingefärbten kreisförmigen Ebenen. Den äußersten grauen Ring und die unterste Ebene stellen Differenzierung und Adaptivität dar, die nach Ansicht der Autorinnen und Autoren (Charalambous & Praetorius, 2020; Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) die erforderliche Basis für die beiden darüber liegenden Ebenen bilden und mit diesen eng verknüpft sind. Die mittlere Ebene nehmen die Unterrichtsdimensionen Klassenführung, sozio-emotionale Unterstützung und Unterstützung der aktiven Beteiligung ein, die auf die oberste und innerste Ebene eventuell, aber nicht notwendigerweise unterstützend wirken können. Die Klassenführung umfasst dabei (wie bereits im Modell der drei Basisdimensionen) im Wesentlichen Facet-

ten der Störungsprävention und des Verhaltensmanagements, die sozio-emotionale Unterstützung Aspekte der Beziehung der Schüler:innen untereinander wie auch zur Lehrkraft und die Unterstützung der aktiven Beteiligung Merkmale der Bereitstellung eines lernförderlichen Umfelds und der Förderung der aktiven Mitwirkung aller Schüler:innen. Die in Abbildung 2 auf oberster Ebene befindlichen, eng zusammenhängenden Dimensionen sind einem prototypischen Lehr-Lernprozess entsprechend angeordnet und fokussieren dessen direkte Unterstützung: Zur Dimension Auswahl und Thematisierung von Inhalten und Fachmethoden gehören Aspekte der Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten, von deren Motivierung und Strukturierung im Unterricht wie auch zu deren fachlicher Akkuratheit und Korrektheit. Während auf die Dimension der kognitiven Aktivierung bereits weiter oben eingegangen wurde (Abschnitt 2.2), beschreibt die Unterstützung des Übens die Festigung von Lern- und Anwendungsprozessen (z. B. durch intelligentes oder repetitives Üben) sowie einen konstruktiven Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern beim Üben. Eine klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen, eine regelmäßige Überprüfung des Verständnisses der Schüler:innen, ein qualitativ hochwertiges Feedback an die Schüler:innen und eine Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts sind hingegen Facetten der Dimension formatives Assessment.

Die voranstehend erläuterten drei Ebenen im Zentrum des MAIN-TEACH-Modells sind somit Beschreibungskategorien der Angebots-Nutzungsstruktur von Unterricht, einem reflexiv-zirkulären und kokonstruktiven Aushandlungsprozess, der durch die beiden umlaufenden Pfeile in Abbildung 2 symbolisiert wird, um das Wechselspiel zwischen unterrichtlichem Angebot auf Lehrkraftseite und dessen Nutzung aufseiten der Schüler:innen hervorzuheben und noch stärker in den Fokus der Unterrichtsforschung zu rücken (Praetorius & Gräsel, 2021; Praetorius & Ufer, 2022; Vieluf et al., 2020). Analog zur hypothetischen Wirkungskette von den Säulen 3 bis 5 im Kaskadenmodell (Abb. 1) werden zudem direkte Zusammenhänge zwischen der obersten Ebene bzw. dem innersten Kreis der Unterrichtsdimensionen und den hierdurch bei Schülerinnen und Schülern potenziell initiieren Lernprozessen angenommen. Die beiden weiteren Ebenen sind mit diesen nur indirekt über den innersten Kreis verbunden, der somit als Mediator fungiert (Charalambous & Praetorius, 2020; Ufer & Praetorius, 2022). Die angeregten Lernprozesse sollten kurz-, mittel- bzw. langfristig zu Leistungssteigerungen und -erfolgen bezüglich bestimmter Outcomes (z. B. Wissens-, Kompetenzzuwachs, affektive Komponenten) führen, die wiederum mit den intendierten Bildungszielen übereinstimmen sollten. Dies wird schließlich durch eine Rahmung mit zeitlich abgestuften Zielen und Outcomes zum Ausdruck gebracht. Hierzu erläutern Praetorius und Gräsel (2021), dass Unterricht in der Forschung nicht nur auf Wissenserwerbsziele reduziert werden, sondern die Vielfalt an bedeutsamen Unterrichtszielen (z. B. ästhetische Bildung, Interesse am Fach, Lernen lernen) größere Beachtung finden sollte, wenn diese unter Umständen auch schwer(er) zu erfassen seien (Rothgangel, 2021; Lindl et al., 2024). Zudem könnten verschied-

dene Lernziele gegebenenfalls konkurrieren, einer längeren Untersuchungsdauer als bislang üblich bedürfen oder eigene Beschreibungs- und Analysedimensionen von Unterrichtsqualität erfordern, sodass diese ohne Kenntnis der zu erreichenden Ziele möglicherweise nur eingeschränkt zu beurteilen sei. Schließlich sei über die grundlegenden Zusammenhänge zwischen Zielen und unterrichtlichen Prozessen zu deren Erreichung bislang nur wenig bekannt (Praetorius & Gräsel, 2021; Seidel et al., 2021).

Auch wenn der voranstehende summarische Überblick über Grundbegriffe, Ansätze und Entwicklungen bezüglich Merkmalen und Modellen von Unterrichtsqualität gemäß dem Charakter dieses Einführungsbeitrags lediglich einige Meilensteine des fortschreitenden bildungswissenschaftlichen Forschungsdiskurses skizzieren konnte, werden daraus aktuelle Forschungstendenzen und -aufgaben ersichtlich. Die bislang angenommene Allgemeingültigkeit der drei Basisdimensionen wird zunehmend kritisch hinterfragt und zugunsten einer größeren Spezifizierung aufgegeben. Denn zum einen ist die inhaltliche Valenz der drei Merkmale – wie sich insbesondere am Beispiel der kognitiven Aktivierung zeigen lässt – unterbestimmt, zwischen verschiedenen Studien (teils erheblich) variabel konzeptualisiert und damit oftmals wenig konkretisierbar (Praetorius et al., 2018; Rothland, 2024). Zum anderen finden gewisse generische (z. B. Unterstützung des Übens) wie auch fachspezifische Merkmale (z. B. fachliche Korrektheit, Umgang mit Angst oder Ekel beim Lernen unter Gefährdungspotenzial oder an Lebewesen) bislang zu wenig Beachtung (Praetorius et al., 2020b). Um diese Leerstellen zu identifizieren, wird ein stärkerer Einbezug unterschiedlicher Fachkulturen und der Fachdidaktiken in die Forschung zu Unterrichtsqualität gefordert (z. B. Lindl et al., eing.; Praetorius & Gräsel, 2021; Praetorius & Nehring, 2020). Vor diesem verbreiterten Disziplinspektrum sollen die bislang vorwiegend in mathematisch-naturwissenschaftlichen Kontexten entwickelten und überprüften Modelle zu (Merkmalen von) Unterrichtsqualität systematisch um (Sub-)Dimensionen ergänzt und weiterentwickelt sowie untersucht werden, über welche beobachtbaren Indikatoren sich diese empirisch erfassen lassen und welche generischen und domänenspezifischen Aspekte für qualitätsvollen Unterricht und die Erreichung von dessen unterschiedlichen Zielen relevant sind.

Das hier vorgestellte MAIN-TEACH-Modell (Abb. 2) kann als gegenwärtiger Kristallisationspunkt der bildungswissenschaftlichen Forschung zu Unterrichtsqualität der auf diesem Gebiet vergleichsweise jungen Fremdsprachendidaktik (vgl. hierzu Abschnitt 4) einen theoretischen Orientierungs- und Ansatzpunkt bieten, auf den auch alle Beiträge dieses Bandes Bezug nehmen. Die Ebenen und Dimensionen des Modells eignen sich nämlich als konzeptuelle Grundlagen für die Entwicklung von fachbezogenen Analyseansätzen und -kategorien wie auch als unterschiedliche Perspektiven auf Unterricht, die gewisse Merkmale in den Vorder- oder Hintergrund treten lassen. Hierdurch können sie ebenfalls (angehenden) Lehrkräften und allen, die in der Lehrkräftebildung tätig sind, dazu dienen, eigenen und fremden Unterricht in Hinblick auf wichtige Qualitätsdimensionen zu beschreiben und zu beurteilen (Praetorius & Gräsel, 2021).

3 Unterrichtsqualität (fach-)spezifisch denken und erfassen

3.1 Spezifität als Kontinuum

Der Ansatz, Unterrichtsqualität anhand einiger weniger generischer Merkmale wie den drei Basisdimensionen zu modellieren und zu erfassen, ist nicht nur unter pragmatischen, theoretischen, methodischen, statistischen und ökonomischen Gesichtspunkten attraktiv. Er beansprucht auch ein hohes Abstraktionsniveau und eine übergreifende Gültigkeit etwaiger Befunde beispielsweise jenseits von Schularten, Jahrgangsstufen oder Fachkulturen. Dies geht jedoch zulasten der inhaltlichen Validität, des spezifischen Auflösungsgrads sowie der kriterialen Passgenauigkeit im Einzelfall wie bei der Betrachtung einer bestimmten Unterrichtssequenz zu einem gewissen Thema oder einer konkreten Unterrichtsaktivität (z. B. Instruieren oder Erklären). Hier sind nuancenreichere, differenziertere Merkmalsausprägungen und Untersuchungsinstrumente oder zumindest deren gegenstandsbezogene Adaption erforderlich, um zu einer angemessenen Einschätzung der Qualität dieser unterrichtlichen Mikro- bis Mesoprozesse gelangen zu können (vgl. z. B. Charalambous & Praetorius, 2018; Lindl et al., 2019). Als Folge solcher Konkretisierungen und Spezifizierungen sind Tragweite und Übertragbarkeit der jeweiligen Erkenntnisse aber womöglich fraglich oder sogar nicht gegeben. In diesem offensichtlichen Spannungsfeld zwischen Spezifität und Generalisierbarkeit, das sich als Kontinuum beschreiben (vgl. Abb. 3; Charalambous & Praetorius, 2018) und in dem sich jedes Vorhaben verorten lässt (vgl. Abb. 3), bewegt sich die Forschung zu Unterrichtsqualität. Die daraus resultierenden Implikationen sind bei der Konzeption entsprechender Untersuchungsansätze, Fragestellungen wie auch Ziele notwendigerweise zu berücksichtigen.

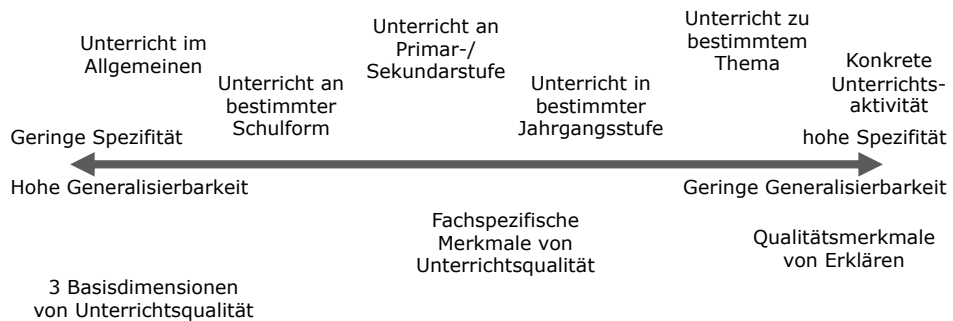


Abbildung 3: Spezifität als Kontinuum (mit exemplarisch benannten Abstufungen)

In der Bildungswissenschaft fokussierte die Forschung zu Unterrichtsqualität bislang vorwiegend die linke Hälfte von Abbildung 3 unter der Zielsetzung einer tendenziell hohen Generalisierbarkeit der Befunde (bei geringer Spezifität); in den Fachdidaktiken überwog ein vernehmlich differenziertes mikro- bis mesoskopisches Vorgehen mit eher hohem Auflösungsgrad (bei evtl. eingeschränkter Gene-

realisierbarkeit) auf der rechten Seite von Abbildung 3. Die Untersuchung von Unterrichtsqualität muss auch weiterhin aus beiden Blickwinkeln erfolgen, an allen Stellen dieses Kontinuums von Spezifität ansetzen und wichtige weiterführende Fragestellungen aufgreifen. Eine gemeinsame Hauptaufgabe von Bildungswissenschaft und Fachdidaktik besteht allerdings darin, die jeweiligen Perspektiven, die unterschiedlichen Abstufungen und Schwerpunktsetzungen von Ansätzen beider Seiten sowie die darauf fußenden Befunde in ein kohärentes Gesamtbild zu überführen, zu integrieren und wechselseitig nutzbar zu machen (Wilden, 2021). Hinzu kommt, dass in der Lehrkräftebildung nur in der Gesamtschau eine kohärente Vorstellung von Unterrichtsqualität über institutionelle Fächergrenzen hinweg vermittelt werden kann. Zu einem solchen fruchtbaren Brückenschlag, an dem es laut Helmke (2022) insbesondere im Bereich Fremdsprachen mangelt, möchte der vorliegende Band einen Beitrag leisten, indem er auf der Mesoebene fachspezifischer Unterrichtsqualität in Fremdsprachen die Forschungsperspektiven unterschiedlicher Vorhaben und Projekte versammelt.

3.2 Spezifität als methodische Gestaltungskomponente

Konzeptualisierung und Operationalisierung. Wurde bislang vor allem der inhaltliche Ansatzpunkt von Spezifität betrachtet, so wird diese auch durch Elemente des Studiendesigns beeinflusst und festgelegt. Dies trifft erstens auf die Konzeptualisierung, d. h. die theoretische Setzung und Beschreibung der zu untersuchenden Konstrukte, zu, die generisch oder fachspezifisch vorgenommen werden kann. Während eine fachspezifische Konzeptualisierung nur mit einer eben solchen Operationalisierung einhergehen kann, ist es zweitens jedoch denkbar und möglich, eine generische Konzeptualisierung beispielsweise jahrgangsstufen-, fach-, themenbezogen etc. messbar zu machen, indem die Testaufgaben, Fragebogenitems o. Ä. entsprechend formuliert werden. Im ersten Fall bleiben die Befunde wiederum auf den spezifischen Untersuchungsgegenstand beschränkt und ihre Generalisierbarkeit unklar; demgegenüber bietet der Ansatz einer generischen Konzeptualisierung und spezifischen Operationalisierung den wesentlichen Vorteil, Ergebnisse auf einer abstrakteren Ebene vergleichen und verallgemeinern zu können (z. B. über Themen, Unterrichtsprozesse oder Fächer hinweg). Zugleich ist die Messung dem im speziellen Studienkontext gewünschten Auflösungsgrad gemäß inhaltlich konkretisiert, angemessen adaptiv und ausreichend valide. Tabelle 1 gibt abschließend einen Überblick über Kombinationsmöglichkeiten generischer und spezifischer Konzeptualisierung und Operationalisierung sowie deren Vor- und Nachteile.

Tabelle 1: Kombinationsmöglichkeiten sowie Vor- und Nachteile generischer und spezifischer Konzeptualisierung und Operationalisierung von Unterrichtsqualität

Operationalisierung	Konzeptualisierung	
	generisch	spezifisch
generisch	<ul style="list-style-type: none"> + gute Vergleichbarkeit + hohe Verallgemeinerbarkeit + dekontextualisierte Theoriebildung - kontextlose Erfassung kaum umsetzbar - reduzierte inhaltliche Aussagekraft - evtl. eingeschränkte Validität 	(nicht möglich)
spezifisch	<ul style="list-style-type: none"> + Vergleichbarkeit gegeben + Verallgemeinerungen möglich + für Theoriebildung abstraktionsfähig + i. d. R. gegenstandsbezogen angemessene Erfassung - Überprüfung der konstruktturen Konkretisierung erforderlich - Aussagegehalt von umgesetzter Operationalisierung abhängig 	<ul style="list-style-type: none"> + konstruktidentische Operationalisierung + klare inhaltliche Aussagekraft + gegenstandsbezogene Validität - geringe Vergleichbarkeit - niedrige Verallgemeinerbarkeit - für Theoriebildung kaum abstraktionsfähig

Beobachtungsperspektiven. Zur Erfassung von Unterrichtsqualität werden in der bildungswissenschaftlichen Forschung unterschiedliche methodische Zugänge verwendet, die in der Regel nach den drei Perspektiven ihrer Akteurinnen und Akteure unterschieden werden. Erstens kann die Lehrkraft selbst Auskunft über die empfundene Qualität ihres Unterrichts geben (z. B. in Fragebögen, Interviews, Portfolios, Tagebüchern etc.), zweitens können die Schüler:innen als die direkten Adressatinnen und Adressaten mittels Fragebögen über die wahrgenommene Qualität des Unterrichts befragt werden und drittens kann Unterricht oder auch nur darin als wesentliches Steuerungselement eingesetztes Lehr-Lernmaterial von (externen) Beobachterinnen und Beobachtern in Präsenz und simultan oder auch asynchron mittels Videotechnik unter Einbezug von Kriterienrastern, Beobachtungsbögen etc. betrachtet und analysiert werden. Alle drei Blickwinkel besitzen dabei spezifische Vorteile, aber auch einige Nachteile, die in Tabelle 2 zusammengefasst sind (vgl. z. B. Clausen, 2002; Helmke, 2022; Herbert & Schweig, 2021; Kleickmann et al., 2019).

Tabelle 2: Vor- und Nachteile verschiedener Beobachtungsperspektiven bezüglich Unterrichtsqualität (nach Kleickmann et al., 2019)

Perspektive	Vorteile	Nachteile
Lehrkräfte	<ul style="list-style-type: none"> - durch ihre Ausbildung i. d. R. auf Beurteilung von Unterrichtsqualität vorbereitet - besitzen fachdidaktische und pädagogische Expertise - Einsicht in verfolgte Ziele und in Kontextinformationen zur Klasse - Langzeitperspektive einnehmbar - ökonomisch durchzuführen 	<ul style="list-style-type: none"> - Involviertheit beeinflusst Wahrnehmung - soziale/subjektive Erwünschtheit (eigener blinder Fleck) - verzerrte Wahrnehmung sogar von Oberflächenstrukturen (wie z. B. Redezeit)
Schüler:innen	<ul style="list-style-type: none"> - Adressat:innen des Unterrichts (z. B. Verständlichkeit, sozio-emotionale Unterstützung) - Kenntnis der Kontextinformationen - variabler, längerer Beurteilungszeitraum möglich - ökonomisch durchführbar und hohe Zuverlässigkeit auf Klassenebene 	<ul style="list-style-type: none"> - keine fachdidaktische oder pädagogische Expertise (z. B. kognitive Aktivierung) - Involviertheit in unterrichtliches Geschehen - (im Primärbereich: Verständnisschwierigkeiten und Fehlinterpretationen)
Beobachter:innen	<ul style="list-style-type: none"> - nicht selbst involviert, kriterienfokussiert, trainiert, distanziertere Wahrnehmung - Vergleichsmöglichkeiten und Maßstäbe aufgrund Erfahrung - bei Videoverfahren: Standardisierung, Wiederholbarkeit, hohe Reliabilität und Validität 	<ul style="list-style-type: none"> - kurzer (verzerrter) Beobachtungszeitraum (z. B. bei kog. Aktivierung) - „Show-Stunden“ (kaum bedeutsam für Basisdimensionen), Reaktivität - keineswegs fehlerfrei bzw. unverzerrt - wenig Kenntnis der Kontextinformationen (insb. bei Videobeobachtung) - zeit- und kostenintensiv

Untersuchungen zu Zusammenhängen zwischen diesen drei Perspektiven legen des Weiteren nahe, dass derselbe Unterricht je nach Blickwinkel teils ganz unterschiedlich eingeschätzt werden kann (z. B. Clausen, 2022; Fauth et al., 2014; Kunter & Baumert, 2006; Waldis et al., 2010). Dabei vermag keine der drei Sichtweisen Unterricht so abzubilden, wie er wirklich ist, und somit für sich eine höhere Authentizität zu beanspruchen. Vielmehr machen bereits Kunter und Baumert (2006) deutlich, dass jede der drei Perspektiven ihre ganz eigene perspektivenspezifische Validität und Dignität besitzt. Dies gilt insbesondere in Bezug auf bestimmte Unterrichtsmerkmale, da nämlich Schüler:innen beispielsweise für die Adressatenspezifität und die Verständlichkeit von Unterricht, Lehrkräfte selbst für

die intendierten didaktischen Zielsetzungen bzw. externe Beobachter:innen für Aspekte der kognitiven Aktivierung und Auswahl von Lerngegenständen Expertinnen und Experten sind. Sofern forschungsökonomisch also umsetzbar, sollte die Beobachtungsmethode möglichst passend zum Untersuchungszweck und Erkenntnisziel gewählt, bestenfalls mehrere Perspektiven einbezogen und schließlich eine Triangulation der heterogenen Blickwinkel angestrebt werden (Clausen, 2002; Desimone et al., 2010; Göllner et al., 2016, 2020; Kleickmann et al., 2019; Kunter & Baumert, 2006; Praetorius et al., 2017; zusammenfassend Fauth et al., 2020). Unter den Einzelperspektiven gilt aber weiterhin die Beobachtung durch externe Urteilende als Goldstandard der Unterrichtsqualitätsforschung (z. B. Helmke, 2022), zumal darauf die meisten Befunde zur prognostischen Validität von Qualitätsmerkmalen bezüglich Lernerfolgen von Schülerinnen und Schülern zurückgehen (z. B. Fauth et al., 2014; Kleickmann et al., 2019; Praetorius et al., 2018). Anzumerken, aber noch unzureichend geklärt, ist außerdem der Einfluss von Vorerfahrung und Fachkenntnissen auf die Einschätzung von Unterrichtsqualität, was für die Gruppen von beispielsweise fachkundigen bzw. -fremden Lehrkräften sowie Schülerinnen und Schülern ebenso gilt wie für externe Beobachtende mit oder ohne fachlicher, fachdidaktischer bzw. pädagogischer Vorbildung, wie Dreher und Leuders (2021) verdeutlichen. So können bei hochinferenten Ratings zum Teil bis zu 41% (!) der Varianz auf divergierende Wahrnehmungs- und Deutungsprozesse der Beobachter:innen zurückgeführt werden und liegen damit nicht im fokussierten Merkmal selbst begründet (Clausen, 2002; Kleickmann et al., 2019).

Dauer der Beobachtung. Auch wenn die diesbezügliche Methodenforschung noch am Anfang steht, legen erste Befunde nahe, dass je nach Merkmal von Unterrichtsqualität unterschiedlich lange Zeiträume zu dessen reliabler Erfassung erforderlich sind. Während Klassenführung und konstruktive Unterstützung relativ stabile Merkmale darstellen, die sich bereits an einer Unterrichtsstunde festmachen lassen (Praetorius, 2014), erscheint eine Bewertung des kognitiven Aktivierungspotenzials erst ab neun beobachteten Unterrichtsstunden zuverlässig möglich (Praetorius, 2014). Dieser Herausforderung begegnen aktuelle Studien einerseits unter anderem mit einem Einbezug von Lehr- und Lernmaterialien zur Analyse des Angebotspotenzials von Unterricht, um verlässlichere Ergebnisse zu erhalten (z. B. Jordan et al., 2006; Herbert & Schweig, 2021).

Andererseits reduziert der Thin-Slices-Ansatz den mit dieser Problematik einhergehenden forschungsökonomischen Aufwand gerade durch eine gegenteilige Herangehensweise, indem die Kosten-Nutzen-Relation optimiert und Unterricht anhand eines Ausschnitts von 30 Sekunden beurteilt wird, der einer videograferten Unterrichtsstunde zufällig entnommen wird (Begrich et al., 2017; Dorfner et al., 2017). Die Präzision dieses „Ersteindrucksratings“ ist in vielerlei Hinsicht bemerkenswert und der Ansatz stellt insbesondere aufgrund seiner prognostischen Validität, die von derjenigen systematischer Ratings nicht allzu sehr abweicht, eine erwägenswerte Alternative zur ökonomischen Untersuchung von Unterrichtsqualität dar.

4 Unterrichtsqualität in der Fremdsprachendidaktik

Da Wilden (2021) und Guttke (2023) jüngst umfassende Übersichten über den Forschungsstand zu Unterrichtsqualität in der Fremdsprache Englisch bzw. insbesondere zur Dimension der kognitiven Aktivierung im schulischen Fremdsprachenunterricht vorgelegt haben, beschränken sich die nachstehenden Ausführungen auf eine kurze Skizze der Grundzüge dieses Forschungsbereichs sowie auf einige ausgewählte, zentrale Studien (z. B. DESI, TEPS, TIOS). Hierzu finden sich in diesem Band auch weitere Darstellungen bei Brunsmeier, Folkerts, Hofrichter et al. sowie Guttke und Porsch für das Fach Englisch, bei Hofrichter et al. für Latein, bei Tark für Russisch sowie bei Franke und Reimann für die romanischen Fremdsprachen. Bei einer zusammenfassenden Gegenüberstellung mit den übrigen Sprachfächern ist der Forschungsstand in Englisch als vergleichsweise fortgeschritten zu bezeichnen, weshalb im Folgenden näher auf diesen eingegangen wird.

4.1 Forschung zu Unterrichtsqualität im Fach Englisch

Bisherige Forschungsarbeiten zur Qualität von Englischunterricht beschäftigen sich vorwiegend mit normativen Prinzipien zu bestimmten Teilbereichen (z. B. interkulturelles Lernen, Brunsmeier, 2016; Freitag-Hild, 2010), wie bereits Wilden (2021) konstatiert. Diese können zwar gewisse Anhaltspunkte hinsichtlich der Bedeutung einzelner Mikroprozesse wie des Code-Switchings seitens der Lehrkräfte bei der Vermittlung von neuem Vokabular (Tian & Macaro, 2016) oder des formelhaften Sprachgebrauchs für das produktive Sprechen in der Grundschule (Kostka, 2020) geben; sie bieten wertvolle Impulse für fachspezifische Merkmale von Unterrichtsqualität und verleiten zu bestimmten Hypothesen bezüglich deren Effekten auf Lernerfolge von Schülerinnen und Schülern. Umfassende quantitative Untersuchungen, die generische oder gar domänenspezifische Qualitätsaspekte modellieren und deren Wirksamkeit in Bezug auf unterrichtliche Zielkriterien wie Kompetenzzuwachs oder Steigerung der Lernfreude überprüfen, sind bislang jedoch eher selten (Wilden, 2021). Vier Studien, die hierzu erste Ergebnisse vorlegen, werden im Folgenden kurz in chronologischer Reihenfolge vorgestellt.

Deutsch Englisch Schülerleistungen International (DESI). In einer bundesweiten Erhebung während des Schuljahrs 2003/2004 wurden Zuwächse in der rezeptiven und produktiven Kompetenz im Englischen (Hörverstehen, Leseverstehen, Schreiben, Textrekonstruktion, Sprachbewusstheit, Sprechen, interkulturelle Kompetenz) von knapp 11.000 Schülerinnen und Schülern aller Schulformen am Anfang und Ende der neunten Jahrgangsstufe untersucht. Auch deren Lehrkräfte, Schulleitungen und Eltern wurden befragt (Helmke et al., 2008; Klieme et al., 2008; Nold & Rossa, 2008). In einer angegliederten Videostudie (Helmke et al., 2008), die Lehr-Lern-Prozesse zur „Gewinnung verhaltensnaher Indikatoren der Unterrichtsqualität“ (Helmke et al., 2008, S. 345) fokussierte, wurde analysiert,

welche der erfassten fachspezifischen Unterrichtsmerkmale Zusammenhänge mit einem Zuwachs an Hörverstehens- und Textrekonstruktionsfähigkeit als Globalindikatoren von Sprachkompetenz besitzen. Am Ende des Schuljahres korrelierten folgende Prozessmerkmale substanziell mit positiven Ergebnissen beim Hörverstehen, das laut Klieme (2007) besonders stark von der Lernumgebung beeinflusst wird und daher mit der Qualität von Englischunterricht in Verbindung gebracht wird:

- Langsamkeitstoleranz, d.h. Wartezeit auf Antworten von Schülerinnen und Schülern,
- hoher zeitlicher Anteil der Sprechzeit von Schülerinnen und Schülern an der Gesamtredezeit,
- Verwendung der englischen anstatt der deutschen Sprache im Unterricht,
- Aufgabenorientierung und Strukturiertheit,
- Aktivierung von Schülerinnen und Schülern, d.h. Einbezug möglichst aller Schüler:innen in die Kommunikation,
- mehr längere Äußerungen von Schülerinnen und Schülern („sustained speech“), weniger „Ein-Wort-Sätze“ (Helmke et al., 2008).

Die Ergebnisse der DESI-Studie weisen somit zum einen auf wichtige fachspezifische Prozessmerkmale des Englischunterrichts hin, können aber zum anderen auch zahlreiche generische Aspekte von Unterrichtsqualität bestätigen wie die „Integration der Hausaufgaben in den Unterricht, eine effiziente Klassenführung, Nutzung von Fehlern als Lerngelegenheiten und die ausgeprägte Verständlichkeit der Lehreraußerungen“ (Klieme, 2007, S. 7).

Längsschnittstudie von Maag Merki (2011). An der Längsschnittstudie zu Qualitätsmerkmalen im Englischunterricht von Maag Merki (2011) nahmen in den Jahren 2007, 2008 und 2009 insgesamt 3258 Schüler:innen von 18 hessischen und 19 Bremer Gymnasien teil. Die implizierte Fragestellung des Ländervergleichs, der als quasi-experimentelles Design diente, war, inwiefern die Substitution dezentraler durch die Einführung zentraler Abiturprüfungen die Qualität des Unterrichtsangebots in Abschlussklassen positiv beeinflusst. Hierzu wurden in Anlehnung an Klieme (2006) die Kriterien konstruktive Unterstützung und kognitive Aktivierung betrachtet und von Schülerinnen und Schülern in Grund- sowie Leistungskursen anhand von 18 generischen Fragebogenitems bewertet, die ebenso in anderen Fächern (z. B. Biologie, Deutsch, Mathematik) zum Einsatz kamen. Im Fokus standen allerdings weder die Übertragbarkeit oder genuine Passung dieser fachübergreifend angelegten Indikatoren noch die Identifikation fachspezifischer Merkmale; die Ergebnisse der Mehrebenenanalysen, die je nach Fachkontext unterschiedlich ausfallen und somit dessen Bedeutung nahelegen, verdeutlichen vielmehr eine gezielte Steuer- und Optimierbarkeit konstruktiver Unterstützungs- sowie vor allem kognitiver Aktivierungsprozesse und eine mittelfristige Stabilität dieser Effekte im Fach Englisch.

Teacher Input Observation Scheme (TIOS). Kersten et al. (2018) entwickelten ein Beobachtungsinstrument zur Messung von Unterrichtsqualität für den Pri-

mar- und Sekundarbereich. Dieses enthält 41 Items zu Lehrtechniken in der Zielsprache Englisch mit den folgenden fünf Dimensionen:

- Cognitively Stimulating Tasks/Activities,
- Verbal Input,
- Non-Verbal Input,
- Support of Learners' Output und
- Corrective Feedback.

Diese Dimensionen wurden auf Basis des kognitiv-interaktionalen Forschungsparadigmas zum Fremdsprachenerwerb nach Long (2015) operationalisiert. In einer Pilotierung mit 17 Lehrkräften sowie 218 Schülerinnen und Schülern der dritten bzw. vierten Jahrgangsstufe an regulären und bilingualen Grundschulen erwies sich TIOS als reliables Instrument (Cronbachs Alpha, $\alpha = .905$, 38 Items) zur Erfassung der oben genannten fünf Dimensionen des Englischunterrichts (Kersten, 2021). Im Bereich des Wortschatz- und Grammatikverständnisses werden die L2-Kompetenzen der Schüler:innen maßgeblich von der Qualität des fremdsprachlichen Unterrichts beeinflusst. Bedeutende Unterschiede existieren zudem zwischen herkömmlichen und bilingualen Grundschulen hinsichtlich der L2-Kontaktdauer, -intensität sowie der Einstellungen der Lehrkräfte und der angewandten Unterrichtsmethoden, die Effekte auf den L2-Erwerb, die kognitive Entwicklung und die Einstellungen der Schüler:innen zeigen (Kersten, 2020; vgl. auch Hofrichter et al., in diesem Band; Wilden, 2021).

Teaching English in Primary Schools (TEPS). In einer ersten Studie untersuchten Wilden & Porsch (2019), ob die rezeptiven Fertigkeiten (Hören und Lesen) von 269 Schülerinnen und Schülern der Primarstufe, die bereits in der ersten Jahrgangsstufe mit Englisch begannen, diejenigen der Kinder übertrafen, die ab der dritten Jahrgangsstufe Englisch lernten. Hierzu wurden die Kompetenzen der Schüler:innen und ihre Einstellungen zum Englischunterricht am Ende des vierten Schuljahres querschnittlich erfasst. Die Kinder, die früher in Englisch unterrichtet wurden, erreichten nur im Bereich Leseverstehen leicht bessere Ergebnisse als die später beginnenden Schüler:innen. Ein weiteres Ziel dieser Pilotstudie war zu überprüfen, inwiefern sich die drei Basisdimensionen unterstützendes Klima, kognitive Aktivierung und Klassenführung anhand der Perspektive von Schülerinnen und Schülern auf Englischunterricht in der Primarstufe valide abbilden lassen. Die Ergebnisse implizieren einen moderierenden Effekt von Unterrichtsqualität auf die Beziehung zwischen Einstellungen zum Englischunterricht und rezeptiven Fertigkeiten; Unterrichtsqualität beeinflusst folglich die Wirkung der Einstellungen auf die rezeptiven Fertigkeiten. In weiterführenden Analysen konnte zudem gezeigt werden, dass es signifikante Wechselwirkungen zwischen kognitiver Aktivierung, Klassenführung und positiven Einstellungen zum Englischunterricht mit Blick auf die Lesefertigkeiten gibt, die Einstellungen also umso positiver, je höher die beiden oben genannten Dimensionen waren, was wiederum zu besseren Lesefertigkeiten führte.

In einer Folgestudie mit 776 Schülerinnen und Schülern der vierten Jahrgangsstufe sowie weiterführenden Analysen fanden Wilden et al. (2020) sowie Porsch und Wilden (2021) keine Indizien dafür, dass Lernende mit frühem Lernbeginn (1. Jgst.) hinsichtlich ihrer rezeptiven Englischfähigkeiten am Ende der Grundschulzeit signifikant besser abschneiden als Lernende mit späterem Lernbeginn (3. Jgst.), und zwar auch nicht bei Berücksichtigung verschiedener Merkmale von Unterrichtsqualität bzw. der Qualifikation der Lehrkräfte. Unterstützendes Klima, d.h. das Erleben von Autonomie oder ein konstruktiver Umgang mit Fehlern, korrelierte signifikant und positiv mit den rezeptiven Englischfähigkeiten der Schüler:innen. Darüber hinaus wurde festgestellt, dass die Wahrnehmung der Unterrichtsqualität durch die Lehrkräfte positiv mit ihrer fachspezifischen Begeisterung und ihren zielsprachlichen Kompetenzen zusammenhängt (zusammenfassend Hofrichter et al., in diesem Band).

Wie aus dieser Übersicht deutlich wird, stecken die quantitativen Zugänge zur empirischen Untersuchung von Unterrichtsqualität im Fach Englisch noch in den Anfängen; weitere vielversprechende Ansätze befinden sich derzeit erst in Entwicklung (z.B. Guttke & Porsch, in diesem Band; Helsper, in Vorbereitung; Hofrichter et al., in diesem Band). Diese sind unter theoretischem Rekurs auf das (erweiterte) Syntheseframework und MAIN-TEACH-Modell insbesondere um fachspezifische Ausdifferenzierungen und Ergänzungen bei der Konzeptualisierung bzw. um Konkretisierung bei der Operationalisierung bemüht, um zusätzliche Merkmale qualitativollen Englischunterrichts empirisch zu identifizieren und zu analysieren. In diesem Zusammenhang werden unter anderem auch die (Sub-)Dimensionen kommunikative Aktivierung (Wilden, 2021) oder das Potenzial zu ästhetisch-emotionalem Lernen (Kranefeld, 2021; Lindl et al., 2024) diskutiert.

4.2 Forschungsaufgaben in Bezug auf fremdsprachliche Unterrichtsqualität

Mit Blick auf die gegenwärtigen Forschungsstände der Bildungswissenschaft und Fremdsprachendidaktik hinsichtlich Unterrichtsqualität wird eine Vielzahl an bislang ungelösten Forschungsaufgaben erkennbar, um Antworten auf die einleitend aufgeworfene Fragestellung, was qualitativollen Unterricht ausmacht, zu finden. Für einen ersten Überblick seien hier exemplarisch einige weiterführende Denkansätze aufgeführt:

- Ausdifferenzierung, Konkretisierung und Validierung von (bildungswissenschaftlichen) Modellen zu Unterrichtsqualität in Bezug auf schulische Fremdsprachen (z.B. Folkerts, Hofrichter et al., Guttke & Porsch, Reimann, Tark, je in diesem Band)

Mögliche Fragestellung: *Welche zusätzlichen fachspezifischen (Sub-)Dimensionen sollte ein Modell von Unterrichtsqualität für die Fremdsprachen enthalten? Lassen sich diese (Sub-)Dimensionen empirisch nachweisen?*

- empirische Untersuchung der Effektivität normativer Unterrichtsqualitätsmerkmale in schulischen Fremdsprachen
Mögliche Fragestellung: *Welche Qualitätsmerkmale hängen positiv mit Zielkriterien von Fremdsprachenunterricht zusammen (wie Kompetenzzuwachs, Steigerung der Lernfreude etc.)?*
- Identifikation und Überprüfung von Gemeinsamkeiten und Unterschieden zwischen verschiedenen Fremdsprachen (und anderen Schulfächern) (z.B. Hofrichter et al., Jentsch & Kirchhoff, Reimann, je in diesem Band)
Mögliche Fragestellung: *Welche Dimensionen von Unterrichtsqualität gelten fremdsprachenübergreifend? Worin unterscheiden sich die Fremdsprachen (auch von anderen Fächern)? Welche Unterschiede ergeben sich beispielsweise für den Unterricht in alten und modernen Fremdsprachen?*
- Untersuchung schulform- und jahrgangsstufenübergreifender Qualitätsdimensionen (z.B. Brunsmeier, Franke, je in diesem Band)
Mögliche Fragestellung: *Gelten in allen Altersstufen und Schulformen die gleichen Dimensionen von Unterrichtsqualität im Fremdsprachenunterricht?*
- Analyse von Voraussetzungen für qualitativ vollen Unterricht (z.B. Hofrichter et al., in diesem Band)
Mögliche Fragestellung: *Wie hängt das Professionswissen von Lehrkräften (CK, PCK; Kirchhoff, 2017; Lindl & Kloiber, 2017) mit Unterrichtsqualität zusammen?*
- Zusammenschluss mit praxeologischen Ansätzen (z.B. Bonnet et al., 2022; Folkerts, in diesem Band)
Mögliche Fragestellung: *Welche Lehr- und Lernprozesse werden im Unterricht tatsächlich realisiert und auf welchen Vorannahmen von Unterrichtsqualität und Einstellungen beruhen diese?*
- (konzeptuelle) Klärung des Verhältnisses zu Kernpraktiken des Unterrichtens (z.B. Grossman & Fraefel, 2024; Fraefel, 2022)
Mögliche Fragestellung: *Wie stehen Kernpraktiken des Unterrichtens und Qualitätsmerkmale von Lehr-Lernprozessen zueinander?*
- Entwicklung von qualitativ hochwertigem Unterricht in den Fremdsprachen über die Phasen der Lehrkräftebildung hinweg
Mögliche Fragestellung: *Wie kann der Transfer von Forschungsergebnissen zur Unterrichtsqualität in den schulischen Fremdsprachenunterricht gelingen?*

Zudem kann eine fremdsprachendidaktisch ausgearbeitete Modellierung von Unterrichtsqualität in Kombination mit der Vermittlung von Kernpraktiken (z.B. Grossman & Fraefel, 2024; Fraefel, 2022) ein wertvolles konzeptuelles Hilfsmittel für die Lehrkräftebildung darstellen. Denn hierdurch wird es Dozentinnen und Dozenten in allen Phasen der Lehrkräftebildung ermöglicht, die Kohärenz zwischen bildungswissenschaftlichen und fremdsprachendidaktischen Wissensbeständen und Ausbildungsinhalten über die verschiedenen Phasen hinweg zu steigern.

5 Kurzvorstellung der Beiträge in diesem Band

Der Beitrag von *Armin Jentsch* und *Petra Kirchoff* mit dem Titel „Was bedeutet Fachspezifität? Überlegungen und Beispiele zu fächerübergreifender sowie fachdidaktisch konkretisierter Unterrichtsqualitätsforschung“ verdeutlicht, dass in der Forschung zum Teil sehr divergente Vorstellungen von Fachspezifität existieren und bei der Betrachtung verschiedene Ebenen unterschieden werden können. Die Heterogenität dieser Auffassungen wurde bislang kaum reflektiert, auch wenn sie wesentliche Implikationen und eventuell sogar Limitationen für die jeweiligen Studien zu Unterrichtsqualität mit sich bringen sowie für die Interpretation und Tragweite von deren Befunden haben. Der Beitrag stellt einige dieser Zugänge vor, vergleicht dabei insbesondere den Mathematik- und Englischunterricht und formuliert im Anschluss einige Forschungsdesiderate einer fachdidaktischen Konkretisierung von Unterrichtsqualität.

Joel Guttke und *Raphaela Porsch* fokussieren in ihrem Beitrag „Inhaltsvalidierung von Fragebogenitems zu kognitiver Aktivierung im Englischunterricht der Grundschule mittels Expert:innenbefragung“ die Rolle der Basisdimension kognitive Aktivierung, zu welcher die Befundlage bisher eher inkonsistent ist. Der Beitrag setzt sich daher genauer mit dieser Dimension auseinander und präsentiert Ergebnisse einer Onlinebefragung von selbst ausgewählten Expertinnen und Experten zur Inhaltsvalidität theoriebasiert konstruierter Items, um kognitive Aktivierung im Englischunterricht der vierten Jahrgangsstufe in einem nächsten Schritt mittels Fragebögen für Schüler:innen sowie Lehrkräfte zu erfassen. Daraus werden im Anschluss Implikationen hinsichtlich der Operationalisierung des Konstrukts und einer Revision von Items abgeleitet. Die beschriebene Vorgehensweise macht somit auch den Mehrwert von Expertinnen- und Expertenbefragungen bei der Konstruktion neuer Instrumente in der empirisch forschenden Fremdsprachendidaktik deutlich.

Manuela Franke setzt sich in ihrem Beitrag „Wie setzen Lehrkräfte das Lehrwerk im Französisch- und Spanischunterricht ein? – eine Studie zu Unterrichtsqualität in der Sekundarstufe I mit Fokus auf Sozialform, Binnendifferenzierung und Feedback“ mit Lehrwerken als einem wesentlichen Element der Materialkultur auseinander. Sie geht davon aus, dass nicht das Lehrwerk an sich die Unterrichtsqualität beeinflusst, sondern vor allem die Lehr- und Lernprozesse, die Lehrkräfte mit ihm initiieren. Der Beitrag stellt einen Teil der empirischen Studie zum Umgang von Französisch- und Spanischlehrkräften mit Lehrwerken der Sekundarstufe I vor, in dem es unter dem Ziel der Hypothesenbildung um die Inszenierungsformen von lehrwerkbasierter Unterrichtsphasen und deren Bedeutung hinsichtlich Unterrichtsqualität geht.

Der Beitrag „Was macht guten Fremdsprachenunterricht aus? Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern Englisch und Latein“ von *Maria Hofrichter*, *Alfred Lindl*, *Mara Rader* und *Annika Krämer* stellt zuerst den Status Quo der fachdidaktischen Forschung zu Unterrichtsqualität in den genannten Disziplinen dar. Darauf fußen die bei-

den in diesem Beitrag zentralen Fragestellungen „Welche Merkmale machen nach Ansicht von (angehenden) Lehrkräften guten Unterricht im Fach Englisch bzw. Latein aus?“ und „Wie bewerten (angehende) Lehrkräfte in der bildungswissenschaftlichen und fachdidaktischen Forschung prominent diskutierte Merkmale guten Unterrichts aus ihrer praxisorientierten Perspektive?“. Diese werden anhand einer Onlinebefragung mit 122 (angehenden) Lehrkräften (Lehramtsstudierenden und Lehrkräften im Beruf) aus beiden Fächern untersucht und die entsprechenden Ergebnisse detailreich vorgestellt sowie abschließend diskutiert.

In seinem Beitrag „Unterrichtsqualität in den romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen (Französisch, Spanisch, Italienisch, Portugiesisch)“ beschäftigt sich *Daniel Reimann* mit den vier romanischen Sprachen Französisch, Spanisch, Italienisch sowie Portugiesisch und fokussiert ihre Rolle als in der Regel spät beginnende Fremdsprachen. Er geht der Frage nach, wie Unterrichtsqualität in diesem Zusammenhang auf Grundlage des aktuellen Forschungsstands in diesen Fächern fachspezifisch modelliert werden kann. Der Autor leitet dafür zunächst 13 fachspezifische Merkmale der romanischen Sprachen ab und entwickelt im Anschluss daran 14 Spezifika in Bezug auf ihre Rolle als zweite, dritte und spät beginnende Fremdsprachen.

Sonja Brunsmeier beleuchtet in ihrem Beitrag „Zusammenspiel von generischen und fachspezifischen Aspekten: Ein Konzept für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn“ Teilergebnisse aus dem Forschungsprojekt ‚EMMa‘, welches ein Konzept für digital gestütztes aufgabenbasiertes Lehren und Lernen generiert. Die Autorin fokussiert den Fremdsprachenfrühbeginn im Fach Englisch und die Rolle digitaler Medien und Materialien zur Weiterentwicklung des Englischunterrichts in der Primarstufe. In Bezug auf Unterrichtsqualität setzt sich ihr Beitrag zum Ziel, aufzuzeigen, wie fachspezifische (d. h. Englischkompetenzen) und generische Kompetenzen (z. B. Medienkompetenzen) durch eine entsprechende Gestaltung des Unterrichtsangebots im Fremdsprachenfrühbeginn gefördert werden können.

In seinem Beitrag „Der schulische Fremdsprachenunterricht Russisch: Unterrichtsqualität und Fachlichkeit“ fokussiert *Dennis Tark* die besondere Rolle der russischen Sprache als Fremd- und Herkunftssprache sowie vor allem der enormen Leistungsheterogenität der Lernenden in diesem Schulfach. Aufgrund dieser speziellen Herausforderungen und des diesbezüglich defizitären Forschungsstands der Russischdidaktik ist die Frage nach Qualitätsmerkmalen von Russischunterricht noch weitgehend unbeantwortet. In diesem Beitrag geht es daher vorerst darum, vor dem Hintergrund dieser fachbezogenen Besonderheiten erste Überlegungen anzustellen, was guten Unterricht im Fach Russisch ausmacht. Als Ausgangs-, Referenz- und Kontrapunkt hierfür dient das Syntheseframework bzw. MAIN-TEACH-Modell.

Jens-Folkert Folkerts widmet sich in seinem Beitrag „Wer Ohren hat zu hören – Qualitätskriterien für guten Hörverstehensunterricht aus der Perspektive von Englischfachleitungen“ der Frage, was eigentlich guten Hörverstehensunterricht ausmacht. Dazu beschreibt er zunächst wichtige Prozesse, die beim Hören

ablaufen, und charakterisiert die Besonderheiten des trans- und interaktionalen Hörens. Den Kern seines Beitrags bildet eine qualitative Interviewstudie, in der er sieben Englischfachleitungen zu ihrem Verständnis hinsichtlich der Qualitätskriterien guten Hörverstehensunterrichts befragt hat. Die Studie und ihre Ergebnisse, die hier präsentiert werden, sind Teil eines übergeordneten designbasierten Forschungsprojekts, welches das Ziel verfolgt, ein umfassendes Kriterienraster für kompetenzorientierten Hörverstehensunterricht zu entwickeln.

Literatur

- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiß, M. (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich. <https://doi.org/10.1007/978-3-322-83412-6>
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Begrich, L., Fauth, B., Kunter, M. & Klieme, E. (2017). Wie informativ ist der erste Eindruck? Das Thin-Slices-Verfahren zur videobasierten Erfassung des Unterrichts. *Zeitschrift für Erziehungswissenschaft*, 20(Suppl. 1), 23–47. <https://doi.org/10.1007/s11618-017-0730-x>
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56, 205–213. <https://doi.org/10.1177/0022487105275904>
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies. Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Blömeke, S., Jentsch, A., Ross, N., Kaiser, G. & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, 79, 101600. <https://doi.org/10.1016/j.learninstruc.2022.101600>
- Blömeke, S. & Kaiser, G. (2017). Understanding the development of teachers' professional competencies as personally, situationally, and socially determined. In D. J. Clandinin & J. Husu (Hrsg.), *International handbook of research on teacher education* (S. 783–802). Thousand Oakes: Sage.
- Bonnet, A. & Hericks, U. (2022). Von „Messbarkeitsphobie“ und Durchprozessierungslogik. Kooperatives Lernen im Englischunterricht und die Professionalisierung von Lehrpersonen. In R. Bohnsack, A. Bonnet & U. Hericks (Hrsg.), *Praxeologisch-wissenschaftliche Professionsforschung. Perspektiven aus Früh- und Schulpädagogik, Fachdidaktik und Sozialer Arbeit* (S. 155–174). Bad Heilbrunn: Julius Klinkhardt. <https://doi.org/10.35468/5973-07>
- Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Bern: Huber.
- Brophy, J. (2000). *Teaching*. Brussels: IAE. http://www.ibe.unesco.org/fileadmin/user_upload/archive/Publications/educationalpracticesseriespdf/prac01e
- Brunsmeyer, S. (2016). *Die Entwicklung Interkultureller Kommunikativer Kompetenz im Englischunterricht der Grundschule. Grundlagen, Erfahrungen, Perspektiven*. Tübingen: Narr. <https://doi.org/10.1515/east-2017-0047>

- Charalambous, C. Y. & Praetorius, A.-K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM – Mathematics Education*, 50(3), 355–366. <https://doi.org/10.1007/s11858-018-0914-8>
- Charalambous, C. Y. & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, 67, 100894. <https://doi.org/10.1016/j.stueduc.2020.100894>
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Münster: Waxmann.
- Decristan, J., Hess, M., Holzberger, D. & Praetorius, A. K. (2020). Oberflächen- und Tiefenmerkmale: Eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. *Zeitschrift für Pädagogik*, 66. Beiheft, 102–116. <https://doi.org/10.3262/ZPB2001102>
- Desimone, L. M., Smith, T. M. & Frisvold, D. E. (2010). Survey measures of classroom instruction. Comparing student and teacher reports. *Educational Policy*, 24(2), 267–329. <https://doi.org/10.1177/0895904808330173>
- Dorfner, T., Förtsch, C. & Neuhaus, B. (2017). Die methodische und inhaltliche Ausrichtung quantitativer Videostudien zur Unterrichtsqualität im mathematisch-naturwissenschaftlichen Unterricht – Ein Review. *Zeitschrift für Didaktik der Naturwissenschaften*, 23, 261–285. <https://doi.org/10.1007/s40573-017-0058-3>
- Drechsel, B. & Schindler, A. K. (2019). Unterrichtsqualität. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 353–372). Berlin: Springer. https://doi.org/10.1007/978-3-662-55754-9_18
- Dreher, A. & Leuders, T. (2021). Fachspezifität von Unterrichtsqualität – aus der Perspektive der Mathematikdidaktik. *Unterrichtswissenschaft*, 49(2), 285–292. <https://doi.org/10.1007/s42010-021-00116-9>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Die Qualität von Grundschulunterricht aus der Sicht von Schülerinnen und Schüler, Lehrpersonen und externen Beobachtern: Zusammenhänge zwischen den Perspektiven und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28, 127–137. <https://doi.org/10.1024/1010-0652/a000129>
- Fauth, B., Göllner, R., Lenke, G., Praetorius, A.-K. & Wagner, W. (2020). Who Sees What? Conceptual Considerations on the Measurement of Teaching Quality from Different Perspectives. *Zeitschrift für Pädagogik*, 66. Beiheft, 138–155. <https://doi.org/10.3262/ZPB2001138>
- Fraefel, U. (2022). Mittels Kernpraktiken zu professionellem Unterrichten. *Journal für LehrerInnenbildung*, 22(3), 16–29. <https://doi.org/10.35468/jlb-03-2022-01>
- Freitag-Hild, B. (2010). *Theorie, Aufgabentypologie und Unterrichtspraxis inter- und transkultureller Literaturdidaktik*. *British Fictions of Migration im Fremdsprachenunterricht*. Trier: WVT.
- Gates Foundation (2013). *Ensuring Fair and Reliable Measures of Effective Teaching. Culminating Findings from MET Project's Three Year Study*. http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf
- Göllner, R., Fauth, B., Lenke, G., Praetorius, A.-K. & Wagner, W. (2020). Do Student Ratings of Classroom Management Tell us More About Teachers or About Classroom Composition? *Zeitschrift für Pädagogik*, 66. Beiheft, 156–172. <https://doi.org/10.3262/ZPB2001156>
- Göllner, R., Wagner, W., Klieme, E., Lüdtke, O., Nagengast, B. & Trautwein, U. (2016). Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen: Chancen, Grenzen und Forschungsperspektiven. In Bundesministerium für Bildung und Forschung (Hrsg.),

- Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments* (S. 63–82). Berlin: Bundesministerium für Bildung und Forschung.
- Grossman, P. & Fraefel, U. (Hrsg.) (2024). *Core Practices in Teacher Education. A Global Perspective*. Cambridge: Harvard Education Press.
- Guttknecht, J. (2023). Kognitive Aktivierung im Fremdsprachenunterricht: Ein systematisches Review von Forschungsarbeiten aus dem deutschsprachigen Raum. *Zeitschrift für Fremdsprachenforschung*, 34(2), 145–176.
- Hattie, J. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? *Pädagogik*, 2, 42–45.
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung. Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Seelze-Velber: Klett.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie* (S. 345–363). Weinheim: Beltz.
- Helsper, D. M. (in Vorbereitung). *Kognitive Aktivierung im Englischunterricht*. Trier.
- Herbert, B. & Schweig, J. (2021). Erfassung des Potenzials zur kognitiven Aktivierung über Unterrichtsmaterialien im Mathematikunterricht. *Zeitschrift für Erziehungswissenschaft*, 24(4), 955–983. <https://doi.org/10.1007/s11618-021-01020-9>
- Hochweber, J. & Vieluf, S. (2016). Gender differences in reading achievement and enjoyment of reading: The role of perceived teaching quality. *Journal of Educational Research*, 1–16. <https://doi.org/10.1080/00220671.2016.1253536>
- Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M., Löwen, K., Brunner, M. & Kunter, M. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Kersten, K. (2020). Der Zusammenhang von kognitiven Fähigkeiten und Zweisprachigkeit im frühen L2-Erwerb. Individuelle und externe Faktoren. In H. Böttger, J. Festman & T. Müller (Hrsg.), *Language Education and Acquisition Research. Focusing Early Language Learning* (S. 82–116). Bad Heilbrunn: Julius Klinkhardt.
- Kersten, K. (2021). L2 input and characteristics of instructional techniques in early foreign language classrooms: Underlying theory and pedagogical practice. *The European Journal of Applied Linguistics and TEFL*, 10(2), 27–59.
- Kersten, K., Bruhn, A.-C., Ponto, K., Böhnke, J. & Greve, W. (2018). Teacher Input Observation Scheme (TIOS). *Studies on Multilingualism in Language Education*, 4. Universität Hildesheim. <https://doi.org/10.13140/RG.2.2.22467.35368>
- Kirchhoff, P. (2017). FALCO-E: Fachspezifisches professionelles Wissen von Englischlehrkräften. Entwicklung und Validierung eines domänenspezifischen Testinstruments. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. Mulder (Hrsg.), *Fachspezifische Lehrkräftekompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 113–152). Münster: Waxmann.
- Kleickmann, T., Praetorius, A. K. & Riecke-Baulecke, T. (2019). Beurteilung von Unterrichtsqualität. In F. Zimmermann, J. Möller & T. Riecke-Baulecke (Hrsg.), *Schulische Diagnostik und Leistungsbeurteilung* (S. 207–223). Seelze: Kallmeyer.
- Kleickmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of Teaching in Science Education. More Than Three Basic Dimensions? *Zeitschrift für Pädagogik*, 66. Beiheft, 37–55. <https://doi.org/10.3262/ZPB2001037>

- Klieme, E. (2006). Empirische Unterrichtsforschung: Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52(6), 765–773.
- Klieme, E. (2007). *Zusammenfassung zentraler Ergebnisse der DESI-Studie*. https://www.bildungserver.de/onlineresource.html?onlineresourcen_id=33016
- Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff, H.-G., Schröder, K., Thomé, G. & Willenberg, H. (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS – Impulse für Schule und Unterricht, Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–58). Bonn: BMBF.
- Kostka, N. (2020). Produktives Sprechen im Englischunterricht der Grundschule – Eine empirische Studie zur Bedeutung formelhafter Sequenzen. *Giessener Fremdsprachendidaktik Online*, 15. http://geb.uni-giessen.de/geb/volltexte/2020/15205/pdf/GiFon_15.pdf
- Kranefeld, U. (2021). Der Diskurs um Unterrichtsqualität in der Musikdidaktik zwischen generischen und fachspezifischen Dimensionen. *Unterrichtswissenschaft*, 49(2), 221–233. <https://doi.org/10.1007/s42010-021-00113-y>
- Krauss, S. (2020). Expertise-Paradigma in der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 154–162). Bad Heilbrunn: Klinkhardt. <https://doi.org/10.35468/hblb2020-018>
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N. & Blum, W. (2020). Competence as a continuum in the COACTIV study: The “cascade model”. *ZDM – Mathematics Education*, 52, 311–327. <https://doi.org/10.1007/s11858-020-01151-z>
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B., Kirchhoff, P. & Mulder, R. H. (Hrsg.) (2017). *FALCO – Fachspezifische Lehrkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Münster: Waxmann.
- Kunter, M. & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9, 231–251. <https://doi.org/10.1007/s10984-006-9015-7>
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.) (2011). *Professionelle Kompetenz von Lehrkräften, Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann. <https://doi.org/10.31244/9783830974338>
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional Competence of Teachers: Effects on Instructional Quality and Student Development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Schöningh. <https://doi.org/10.36198/9783838538952>
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften, Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Münster: Waxmann. <https://doi.org/10.31244/9783830974338>

- Lindl, A., Ehrich, P., Gutmiedl, M., Rader, M., Simböck, L., Gürtner, M., Böhringer, S., Krämer, A., Kirchhoff, J. & Frei, M. (2024). Und wo bleibt die Ästhetik? – Betrachtungen zu einer weiteren Dimension von Unterrichtsqualität aus interdisziplinärer Perspektive. In M. Hemmer, C. Angele, C. Bertsch, S. Kapelari, G. Leitner & M. Rothgangel (Hrsg.), *Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung* (S. 371–388). Münster: Waxmann.
- Lindl, A., Gaier, L., Weich, M., Frei, M., Ehras, C., Gastl-Pischetsrieder, M., Elmer, M., Asen-Molz, K., Ruck, A.-M., Heinze, J., Murmann, R., Gunga, E. & Röhl, S. (2019). Eine ‚gute‘ Erklärung für alle?! Gruppenspezifische Unterschiede in der Beurteilung von Erklärqualität – erste Ergebnisse aus dem interdisziplinären Forschungsprojekt FALKE. In T. Ehmke, P. Kuhl & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten. Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 128–141). Weinheim/Basel: Beltz Juventa.
- Lindl, A. & Kloiber, H. (2017). *artifices docendi* – FALKE-L: Modellierung und Messung domänenspezifischer Kompetenzen von Lateinlehrkräften. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. H. Mulder (Hrsg.), *FALKE – Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 153–199). Münster: Waxmann.
- Lindl, A. & Krämer, A. (2022). Was macht guten Lateinunterricht aus? Eine Delphi-Studie zu Unterrichtsqualität im Fach Latein. *Die Alten Sprachen im Unterricht*, 69(1), 9–15.
- Lindl, A., Krauss, S. & Buchholtz, N. (2023). Professionswissen von Mathematiklehrkräften – eine einleitende Übersicht. In S. Krauss & A. Lindl (Hrsg.), *Professionswissen von Mathematiklehrkräften – Implikationen aus der Forschung für die Praxis* (S. 1–35). Berlin: Springer. https://doi.org/10.1007/978-3-662-64381-5_1
- Lindl, A., Simböck, L., Rader, M., Hofrichter, M., Gutmiedl, M., Böhringer, S. & Ehrich, P. (eingereicht). Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften – Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik. *Zeitschrift für Erziehungswissenschaft*.
- Lipowsky, F. & Bleck, V. (2019). Was wissen wir über guten Unterricht? – Ein Update. In U. Steffens & R. Messner (Hrsg.), *Unterrichtsqualität: Konzepte und Bilanzen gelingenden Lehrens und Lernens. Grundlagen der Qualität von Schule* (S. 219–249). Münster: Waxmann.
- Lohe, V., Lindl, A. & Kirchhoff, P. (2022). Tagungsbericht zum Symposium „Unterrichtsqualität in schulischen Fremdsprachen“. *Zeitschrift für Fremdsprachenforschung*, 33(2), 279–280.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Malden: Wiley-Blackwell.
- Maag Merki, K. (2011). The introduction of state-wide exit examinations: Empirical effects on math and English teaching in German academically oriented secondary schools. In M. A. Pereyra, H.-G. Kotthoff & R. Cowen (Hrsg.), *Pisa under examination. Changing knowledge, changing tests, and changing schools* (S. 125–142). Rotterdam/Boston/Taipei: Sense Publishers.
- Meyer, H. (2004). *Was ist guter Unterricht?* Berlin: Cornelsen.
- Nold, G. & Rossa, H. (2008). Hörverstehen Englisch. In E. Klieme, W. Eichler, A. Helmke, R. H. Lehmann, G. Nold, H.-G. Rolf, K. Schröder, G. Thomé & H. Willenberg, H. (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 178–196). Weinheim: Beltz.

- Porsch, R. & Wilden, E. (2021). Teaching English Out-of-Field in Primary School: Differences in Professional Characteristics and Effects on Instructional Quality. In L. Hobbs & R. Porsch (Hrsg.), *Out-of-field Teaching across the Teaching Disciplines and Contexts* (S. 117–134). Singapore: Springer. https://doi.org/10.1007/978-981-16-9328-1_6
- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings*. Münster: Waxmann.
- Praetorius, A. K. & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM – Mathematics Education*, 50(3), 535–553. <https://doi.org/10.1007/s11858-018-0946-0>
- Praetorius, A. K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>
- Praetorius, A. K., Grünkorn, J. & Klieme, E. (2020a). Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen. *Zeitschrift für Pädagogik*, 66. Beiheft, 9–14. <https://doi.org/10.3262/ZPB2001009>
- Praetorius, A. K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020b). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A. K. & Kleickmann, T. (2022). Nutzung von Lerngelegenheiten im Unterricht: Konzeptuelle und methodische Zugänge und Herausforderungen: Einführung in den Thementeil. *Unterrichtswissenschaft*, 50(2), 149–155. <https://doi.org/10.1007/s42010-022-00145-y>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality. The German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Praetorius, A.-K., Laueremann, F., Klassen, R. M., Dickhäuser, O., Janke, S. & Dresel, M. (2017). Longitudinal relations between teaching-related motivations and student-reported teaching quality. *Teaching and Teacher Education*, 65, 241–254. <https://doi.org/10.1016/j.tate.2017.03.023>
- Praetorius, A. K. & Nehring, A. (2020). Unterrichtsqualität zwischen Generik und Fachspezifik: Einführung in den Thementeil. *Unterrichtswissenschaft*, 48(3), 297–301. <https://doi.org/10.1007/s42010-020-00081-9>
- Praetorius, A. K., Rogh, W. & Kleickmann, T. (2020c). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft*, 48(3), 303–318. <https://doi.org/10.1007/s42010-020-00072-w>
- Reusser, K. (2020). Unterrichtsqualität zwischen empirisch-analytischer Forschung und pädagogisch-didaktischer Theorie – Ein Kommentar. *Zeitschrift für Pädagogik*, 66. Beiheft, 236–254. <https://doi.org/10.3262/ZPB2001236>
- Reusser, K. & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49(2), 189–202. <https://doi.org/10.1007/s42010-021-00117-8>
- Rothgangel, M. (2021). Unterrichtsqualität in der Religionsdidaktik – fachspezifische und fachübergreifende Aspekte. *Unterrichtswissenschaft*, 49(2), 253–260. <https://doi.org/10.1007/s42010-021-00107-w>
- Rothland, M. (2024). Auf Sand gebaut? Über den Versuch, die Basisdimensionen der Unterrichtsqualität in erziehungswissenschaftlicher „Theorie“ zu verankern. *Zeitschrift für Bildungsforschung*. <https://doi.org/10.1007/s35834-024-00411-2>

- Schilcher, A., Krauss, S., Kirchhoff, P., Lindl, A., Hilbert, S., Asen-Molz, K., Ehras, C., Elmer, M., Frei, M., Gaier, L., Gastl-Pischetsrieder, M., Gunga, E., Murmann, R., Röhrl, S., Ruck, A.-M., Weich, M., Dittmer, A., Fricke, M., Hofmann, B., ... Thimmabrey, C. (2021). FALKE: Experiences From Transdisciplinary Educational Research by Fourteen Disciplines. *Frontiers in Education*, 5, 579982. <https://doi.org/10.3389/fe-duc.2020.579982>
- Schlesinger, L., Jentsch, A., Kaiser, G., König, J. & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM – Mathematics Education*, 50(3), 475–490. <https://doi.org/10.1007/s11858-018-0917-5>
- Seidel, T., Renkl, A. & Rieß, W. (2021). Basisdimensionen für Unterrichtsqualität im Fachkontext konkretisieren: Gestaltungsmerkmale und Bestimmung von Standardsituationen. *Unterrichtswissenschaft*, 49(2), 293–301. <https://doi.org/10.1007/s42010-021-00108-9>
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.2307/1175860>
- Tian, L. & Macaro, E. (2012). Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students. A Lexical Focus-on-Form study. *Language Teaching Research*, 16(3), 361–385. <https://doi.org/10.1177/1362168812436909>
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Vieluf, S., Praetorius, A.-K., Rakoczy, K., Kleinknecht, M. & Pietsch, M. (2020). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts. Ein kritischer Vergleich verschiedener Modellvarianten. *Zeitschrift für Pädagogik*, 66. Beiheft, 63–80. <https://doi.org/10.3262/ZPB2001063>
- Waldis, M., Grob, U., Pauli, C. & Reusser, K. (2010). In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität – Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 171–208). Münster: Waxmann.
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>
- Wilden, E. & Porsch, R. (2019). The impact of teaching quality and learning time on primary EFL learners' receptive proficiency: Preliminary findings from the TEPS study. *AILA Review* 32(4), 160–177. <https://doi.org/10.1075/aila.00025.wil>
- Wilden, E., Porsch, R. & Schurig, M. (2020). An early start in primary EFL education and the role of teacher qualification and teaching quality. *Language Teaching for Young Learners*, 2(1), 28–51. <https://doi.org/10.1075/ltyl.19002.wil>
- Wiprächtiger-Geppert, M., Stahns, R. & Riegler, S. (2021). Fachspezifität von Unterrichtsqualität in der Deutschdidaktik. *Unterrichtswissenschaft*, 49(2), 203–209. <https://doi.org/10.1007/s42010-021-00109-8>
- Wisniewski, B., Zierer, K., Dresel, M. & Daumiller, M. (2020). Obtaining secondary students' perceptions of instructional quality: Two level structure and measurement invariance. *Learning and Instruction*, 66, 101303. <https://doi.org/10.1016/j.learninstruc.2020.101303>

Armin Jentsch und Petra Kirchhoff

Was bedeutet Fachspezifität?

Überlegungen und Beispiele zu fächerübergreifender sowie fachdidaktisch konkretisierter Unterrichtsqualitätsforschung

Abstract der Herausgeber:innen

The article by Armin Jentsch and Petra Kirchhoff is entitled “What is subject-specificity? Considerations on and examples of instructional quality from a subject-specific and an interdisciplinary research perspective”. The authors show that studies in the field can be based on very different ideas of subject-specificity and that studies often distinguish different levels of instructional quality in their analyses. Even though this has significant implications and possibly even limitations for studies of instructional quality, as well as for the interpretation and scope of their findings, this underlying diversity within one conceptual label has received little attention. In this paper, the authors present a critical analysis that compares Mathematics and English teaching. Research desiderata are formulated for the theoretical basis of instructional quality from the perspective of subject-matter didactics.

1 Einleitung

In aktuellen Arbeiten wird immer wieder gefordert, dass Unterrichtsqualitätsforschung auch die Spezifika der jeweiligen Fachdisziplinen berücksichtigen möge, über die sie Aussagen trifft. Insbesondere wird diskutiert, inwieweit fächerübergreifende (generische) Modelle und Messinstrumente von einer fachspezifischen Anreicherung relevanter Unterrichtsmerkmale profitieren könnten (u. a. Brunner, 2018; Heinitz et al., 2022). In diesen Überlegungen werden jedoch unterschiedliche Konzeptualisierungen von Fachspezifität deutlich, die bisher kaum reflektiert wurden, aber weitreichende Konsequenzen für konkrete interdisziplinäre Forschungsprogramme zu Unterrichtsqualität haben. Im vorliegenden Kapitel werden diese verschiedenen theoretischen Sichtweisen herausgearbeitet und ihre Konsequenzen für die Forschung zur Unterrichtsqualität beschrieben.

Die Bildungsforschung unterscheidet üblicherweise zwischen Oberflächen- und Tiefenstrukturmerkmalen von Unterricht (Decristan et al., 2020; Reusser, 2006; im Überblick Kunter & Ewald, 2016; vgl. auch Lindl et al., in diesem Band). Unter Oberflächenstrukturen werden unmittelbar beobachtbare Merkmale wie Unterrichtsettings und -methoden verstanden (z. B. Gruppenarbeit oder direkte Instruktion; Helmke, 2012). Die Tiefenstrukturen beziehen sich dagegen auf die

zugrunde liegenden Beziehungen zwischen Lehrpersonen, Schüler:innen und Unterrichtsfach (oder -inhalt; Cohen et al., 2003). Die empirische Befundlage macht deutlich, dass Lernzuwächse von Schüler:innen vor allem durch tiefenstrukturelle Merkmale von Unterricht erklärt werden können (z. B. Hattie, 2009). In ihrer Gesamtheit werden sie auch als *Unterrichtsqualität* bezeichnet. Ein entsprechendes und mittlerweile im deutschsprachigen Raum etabliertes fächerübergreifendes Modell mit drei Basisdimensionen (effiziente Klassenführung, konstruktive Unterstützung, Potenzial zur kognitiven Aktivierung) wurde im Anschluss an die TIMS-Videostudie entwickelt (u. a. Klieme et al., 2001; Kunter & Ewald, 2016; Praetorius et al., 2018) und hat sich bereits mehrfach empirisch bewährt (u. a. Baumert et al., 2010; Fauth et al., 2014; Lipowsky et al., 2009). Aktuelle theoretische Bestrebungen der Unterrichtsforschung beziehen sich vor allem darauf, dieses Modell durch weitere Dimensionen zu ergänzen, (fachspezifisch) auszudifferenzieren oder andere, insbesondere internationale Frameworks zu integrieren (Praetorius & Charalambous, 2018; Ufer & Praetorius, 2022; vgl. auch Lindl et al., in diesem Band).

Grundannahme dieses Beitrags ist, dass eine fachspezifische Modellierung von Unterrichtsqualität abhängig vom Forschungsinteresse (z. B. dem in Hinblick auf den Unterricht gewählten „Auflösungsgrad“; Reusser & Pauli, 2021; Lindl et al., in diesem Band) notwendig und hilfreich, in manchen Fällen aber auch hinderlich sein kann. Zunächst wird anhand einiger Beispiele aus der Literatur dargestellt, dass in der Bildungsforschung keine einheitliche Definition von Fachspezifität verwendet wird. Abschließend werden in Anlehnung an das einleitende Kapitel (Lindl et al., in diesem Band) sowie Reusser und Pauli (2021) einige offene Forschungsfragen formuliert, derer sich Forschungsprogramme zur Unterrichtsqualität künftig annehmen könnten.

2 Warum Unterrichtsqualitätsforschung (nicht) fachspezifisch konkretisieren?

Gemäß dem didaktischen Dreieck (Cohen et al., 2003; Reusser, 2006) ist der „Stoff“ oder Unterrichtsinhalt neben der Lehrkraft und den Schüler:innen ein Grundpfeiler von Unterricht. Unterrichtsinhalte außer Acht zu lassen hieße also, eine zentrale Komponente von Unterricht (und damit seiner Qualität) zu vernachlässigen. Mit der Bedeutsamkeit des Fachs für Unterricht gehen Unterrichtsziele einher, die zumindest für einige Fächer wie Deutsch, Mathematik oder der ersten Fremdsprache in nationalen Bildungsstandards formuliert sind. Dies hat insofern Konsequenzen für die Unterrichtswirklichkeit, als die Bildungsstandards ein intendiertes Curriculum darstellen, das auf Landesebene konkretisiert und im (Fach-)Unterricht implementiert werden sollte. Da Unterrichtsziele und – damit verbunden – die Unterrichtswirklichkeit also sowohl fachübergreifende, vor allem aber fachspezifische Komponenten haben, sollten Forschungsarbeiten zur Unterrichtsqualität die Fachspezifität berücksichtigen, wenn sie umfassende Aus-

sagen über Prozessmerkmale erfolgreichen Unterrichtens treffen wollen. Klieme und Rakoczy (2008) empfehlen etwa, dass das Modell der drei Basisdimensionen in Bezug auf die kognitive Aktivierung fachlich und fachdidaktisch konkretisiert werden sollte (vgl. Guttke et al. für Englisch in der Grundschule, in diesem Band), um das Wirkungsgeflecht von Lehrpersonen, Schüler:innen und Unterrichtsfach besser beschreiben zu können. Hier ist auf Anregung von Klieme und Rakoczy (2008) ein interdisziplinäres Forschungsfeld mit komplementären Aufgaben entstanden, das es vermag, Fach, Fachdidaktik und fächerübergreifende Lehr-Lernforschung zu verbinden.

In Anlehnung an das Kaskadenmodell (Lindl et al., in diesem Band) sowie das Modell von Blömeke et al. (2015), das die Kompetenz von Lehrkräften als ein Kontinuum beschreibt, könnte man des Weiteren vermuten, dass eine hohe Unterrichtsqualität ohne ein gewisses Maß an fachspezifischem Professionswissen der Lehrkräfte nicht realisierbar ist (vgl. auch Bromme, 1995; Brunner, 2018, Klieme et al., 2008). Insbesondere für den Mathematikunterricht liegen diesbezüglich auch empirische Nachweise vor (u. a. Baumert et al., 2010; Blömeke et al., 2022; Hill et al., 2008). Forschungsarbeiten zu diesen Zusammenhängen in fremdsprachlichen Fächern stehen noch aus.

Nicht unerwähnt bleiben sollte jedoch, dass der Transfer der empirisch gut abgesicherten, fächerübergreifenden Lehr-Lernforschung zu Unterrichtsqualität auch Vorteile für die Praxis hat. Soll die Qualität des Unterrichts an einer Schule beispielsweise im Rahmen der allgemeinen Schulentwicklung mit Phasen der Evaluation, des Feedbacks und der Setzung und Überprüfung von Entwicklungszielen verbessert werden, so können für alle Fächer gleichermaßen gültige Unterrichtsprinzipien eine belastbare Basis (Pietsch & Tosana, 2008; Taut & Rakoczy, 2016) mit einem in Teilen einheitlichen Bezugsrahmen über alle Fächer hinweg bieten. Eine umfassende Berücksichtigung von fachspezifischen Unterrichtsmerkmalen könnte sich derzeit noch als problematisch erweisen, da wichtige, empirisch abgesicherte Forschungserkenntnisse fehlen und die Vergleichbarkeit und Interpretation der Befunde aufgrund von (fachspezifischen) Kontexteffekten zumindest erschwert, wenn nicht sogar unmöglich sein könnte (Details in Abschnitt 3.3; Kane & Staiger, 2012; vgl. auch Cohen et al., 2018). Ein fachspezifisches Merkmal des Unterrichts moderner Fremdsprachen ist beispielsweise die Verwendung der Zielsprache. Wenngleich die DESI-Studie feststellt, dass unter anderem der Gebrauch von Englisch als Unterrichtssprache positiv mit Leistungen beim Hörverstehen zusammenhängt (Helmke et al., 2008), so gerät dieses Qualitätsmerkmal spätestens mit der Forderung nach einer Entwicklung plurilingueller Kompetenz (KMK, 2024) ins Wanken.

3 Zur Verwendung des Begriffs Fachspezifität in der Literatur

3.1 Fachspezifische Konzeptualisierung von Unterrichtsqualität

Fachspezifische Konzeptualisierungen von Unterrichtsqualität liegen immer dann vor, wenn in verschiedenen Fächern ein unterschiedliches theoretisches Verständnis von Unterrichtsqualität als übergeordnetem Konstrukt oder von einzelnen Unterrichtsqualitätsmerkmalen besteht. Dies kann auch bedeuten, dass ein Merkmal als essenziell angesehen wird, um Unterrichtsqualität in einem bestimmten Fach (oder in einer Fächergruppe) zu beschreiben, während dies für andere Fächer oder in der fächerübergreifenden Lehr-Lernforschung nicht der Fall ist. Kranefeld (2021) nennt beispielsweise die ästhetische Aktivierung von Schüler:innen als „Modus der Welterfahrung“ und damit als zu ergänzendes Merkmal von Unterrichtsqualität für den Musikunterricht (zusammenfassend Lindl et al., 2024). Schwarz (2019) schlägt als zusätzliches Qualitätskriterium für den Religionsunterricht vor, dass dieser einen besonderen Beitrag für sozial-emotionale Aspekte des Schullebens leisten solle. Mit Bezug auf das Schulfach Englisch verweist Wilden (2021, S. 5) unter anderem auf die „Doppelrolle der Zielsprache Englisch als Gegenstand und Medium des Unterrichts“, den positiv, konstruktiven Umgang mit Fehlern und die hohe Bedeutung der zielsprachlichen Kompetenz der Lehrkraft (vgl. dazu auch Gürtner et al., in diesem Band).

Die bisher genannten Beispiele können als mögliche Ergänzungen fächerübergreifender Modelle der Unterrichtsqualität durch fachspezifische Merkmale verstanden werden. Brunner (2018) stellt Überlegungen zu einer Kategorisierung von fachspezifischen Frameworks für den Mathematikunterricht an, die in ähnlicher Weise von Heinitz et al. (2022) für die Naturwissenschaften aufgegriffen wird. Hierbei diskutieren die Autor:innen, inwiefern fachspezifische Merkmale der Unterrichtsqualität in den untersuchten Frameworks zu einer gemeinsamen Dimension kondensiert oder zusammen mit fächerübergreifenden Merkmalen in integrierten Dimensionen zusammengeführt werden können. Beide Möglichkeiten wurden in Studien zur Qualität des Mathematikunterrichts realisiert. Exemplarisch zeigt Tabelle 1 einen Vergleich der Konzeptualisierungen des in der „Pythagoras-Studie“ eingesetzten Beobachtungsinstruments (Rakoczy & Pauli, 2006; Drollinger-Vetter, 2011) als eines mit einer eigenen fachspezifischen Dimension und des im Rahmen der Studie TEDS-Unterricht entwickelten Instruments (Jentsch et al., 2020; Schlesinger et al., 2018) mit integrierten Dimensionen, in denen fächerübergreifende und fachspezifische Qualitätsmerkmale gleichermaßen vorkommen (für die gewählten Kategorien in Tab. 1 vgl. Brunner, 2018).¹

Fraglich ist dagegen, inwieweit die vor allem im Kontext von Mathematikunterricht vielfach erwogene fachliche Korrektheit der Unterrichtsinhalte (Brunner, 2018; Schlesinger et al., 2018) als ein – im hier diskutierten Sinne – fachspezi-

1 Einschränkung ist anzumerken, dass der thematische Fokus der „Pythagoras-Studie“ möglicherweise bereits eine bestimmte Art der fachspezifischen Konkretisierung von Merkmalen und Instrumenten vorgibt oder wenigstens nahelegt.

fisches Merkmal angesehen werden kann. Wie fachliche Korrektheit in den einzelnen Fächern definiert wird, unterscheidet sich bedingt durch unterschiedliche Wege des Erkenntnisgewinns und der Begriffsbildung nach Drollinger-Vetter (2011) von einem Fach zum anderen. So hängt die Feststellung fachlicher Korrektheit von den Normen und Überzeugungen einer Fachkultur ab (Grossman & Stodolsky, 1995). Dies spräche dafür, dass die fachspezifische Konzeptualisierung fachlicher Korrektheit in einigen Fällen sinnvoll sein kann. In den fremdsprachlichen Fächern denkt man bei fachlicher Korrektheit zunächst an normativ-linguistische Setzungen beispielsweise in Bezug auf Grammatik, Lexik und Aussprache. Fachliche Korrektheit ist aber ebenso entscheidend, wenn historisches Hintergrundwissen als Teil der interkulturellen Kompetenz (Byram, 1997; Hu & Byram, 2009) entwickelt werden soll. Besonders deutlich wird dies in diesen Tagen anhand der historisch-kulturellen Auseinandersetzung mit dem israelisch-palästinensischen Konflikt, die Teil einer breiteren *Citizenship Education* ist. An diesen Beispielen lässt sich im fachinternen und im fachübergreifenden Vergleich bereits ablesen, dass das jeweils zugrunde liegende Wissen von unterschiedlicher epistemischer Natur und Gültigkeit ist.

Tabelle 1: Vergleich zweier Instrumente zur Einschätzung von Unterrichtsqualität im Mathematikunterricht

Kategorie	„Pythagoras-Studie“	TEDS-Unterricht
Quellen	Rakoczy und Pauli (2006), Drollinger-Vetter (2011)	Jentsch et al. (2020), Schlesinger et al. (2018)
Forschungsperspektiven	Bildungswissenschaften und Mathematikdidaktik	
Generische Qualitätsmerkmale	vorhanden	
Fachliche Qualitätsmerkmale (Korrektheit)	werden implizit erfasst	werden als Einzelmerkmale erfasst
Fachdidaktische Qualitätsmerkmale	vorhanden in den impliziten Ergänzungen aus mathematikdidaktischer Sicht	vorhanden in den punktuellen Operationalisierungen der Qualitätsmerkmale aus mathematikdidaktischer Sicht
Inhaltsspezifität	in einzelnen Qualitätsmerkmalen vorhanden	inhaltsunabhängig
Verhältnis generischer und fachspezifischer Qualitätsmerkmale	additiv	integrierend

Mit Blick auf den Mathematikunterricht kann sich die Bedeutung fachlicher Korrektheit auch in Abhängigkeit vom Unterrichtsziel unterscheiden. Beispielsweise könnten Lehrkräfte im Rahmen einer Unterrichtseinheit zur Arithmetik (z. B. Zahlenfolgen) auf besonders präzise Sprache fokussieren und damit deutlich machen, dass Mathematik einen Genauigkeitsgrad ermöglicht, der in der alltäglichen

Sprache kaum erreicht werden kann. In der Mathematik können in einer Aussage der Form „für fast alle Elemente der Menge X gilt“ etwa beliebig viele Ausnahmen akzeptiert werden, sofern diese Anzahl *endlich* ist (d.h. die Anzahl der Atome im Universum eingeschlossen). Der mathematische Gehalt dieser Aussage ist unbedingt zu unterscheiden von „für alle Elemente der Menge X gilt“, bei der selbst eine einzige Ausnahme ausgeschlossen wird, auch wenn im alltäglichen Sprachgebrauch zwischen der Anzahl der Atome im Universum und „unendlich“ kaum ein Unterschied besteht. Neben diesem Lernziel, das gerade wissenschaftspropädeutisch wichtig erscheint, strebt Mathematikunterricht aber auch andere Ziele an. Demnach könnte die Fokussierung auf fachlich präzise Sprache in stärker entdeckenden Unterrichtseinheiten oder im Rahmen von mathematischer Modellierung (z.B. wie viele Getränke müssen wir für die Klassenfeier einkaufen?) vielleicht sogar hinderlich sein.

Zusammenfassend können wir festhalten, dass die fachliche Korrektheit der behandelten Unterrichtsinhalte zwar fächerübergreifend als normatives Merkmal guten Unterrichts akzeptiert wird, inwieweit sie jedoch ein lernförderliches Unterrichtskriterium darstellt, hängt mitunter vom verfolgten Lernziel ab. Zudem unterscheiden sich die (wissenschaftlichen) Methoden zur Feststellung der fachlichen Korrektheit zwischen und innerhalb der Fachkulturen.

Die beiden in Tabelle 1 angeführten Studien zum Mathematikunterricht unterscheiden sich außerdem hinsichtlich ihres *Verhältnisses* von generischen und fachspezifischen Qualitätsmerkmalen. Während TEDS-Unterricht fachspezifische Qualitätsmerkmale integriert, konzeptualisiert die Pythagoras-Studie einen gesonderten fachspezifischen Merkmalsbereich. Wenngleich die Integration bzw. Separation fachlich informierter Qualitätsmerkmale in quantitativen Studien durch entsprechende Verfahren statistisch abgesichert werden kann, spielt die grundsätzliche Annahme zum Verhältnis der Bereiche in der Konzeptualisierung, welche die theoretische Grundlage für die Operationalisierung bildet, eine große Rolle. Wenn zum Beispiel Autor:innen der fremdsprachendidaktischen Fächer ihre Operationalisierung auf der Prämisse der kognitiven Aktivierung ansetzen, so ist der Rahmen für die Merkmale eindeutig lernpsychologisch-kognitiv gesetzt. Gehen Studien jedoch von der Annahme eines separaten Merkmalsbündels aus, werden sie vermutlich eher nach den Merkmalen suchen, die einerseits am besten diskriminieren und andererseits aus den bislang vorliegenden theoretischen und empirischen Vorarbeiten in den jeweiligen Didaktiken hervorgehen. Aus diesen Überlegungen wird deutlich, dass der Begriff Fachspezifität nicht eine feste Größe bezeichnet, sondern das zugrundeliegende Verständnis von einer Studie zur anderen sehr unterschiedlich sein kann. Ein Grund hierfür mag die von Rothland (2024) herausgearbeitete „Theoriearmut“ der drei Basisdimensionen sein. Charalambous und Praetorius (2018) gehen in ihren theoretischen Überlegungen zu diesem Aspekt noch einen Schritt weiter, wenn sie ein Kontinuum der Fachspezifität für die Konzeptualisierung von Unterrichtsqualität mit generischen und fachspezifischen Merkmalen vorschlagen, das es erlaubt, die betrachteten Unterrichtsqualitätsmerkmale stärker oder weniger stark zu konkretisieren.

3.2 Fachspezifische Operationalisierung von Unterrichtsqualität

In Forschungsarbeiten, die danach streben, fachspezifische Merkmale im Kontext des generischen Modells mit drei Basisdimensionen integrierend zu entwickeln, werden vor allem fachspezifische Operationalisierungen der Dimension kognitive Aktivierung diskutiert (u. a. Klieme & Rakoczy, 2008; Kunter & Ewald, 2016). Diese Arbeiten basieren häufig auf einem ähnlichen theoretischen Verständnis von kognitiver Aktivierung. Sie können sich allerdings hinsichtlich ihrer unterschiedlichen Perspektiven auf Unterricht (z. B. von Statusgruppen) oder in der Verwendung von Erhebungsmethoden und Messinstrumenten (z. B. direkte Messung von Kompetenzen, Befragung zur Selbsteinschätzung oder Analyse von Lehr- und Lernmaterialien) unterscheiden. Kunter und Ewald (2016) fassen dazu einige Operationalisierungen aus verschiedenen Fachdidaktiken zusammen, die vom Modell der drei Basisdimensionen ausgehen, kognitive Aktivierung aber im jeweiligen Unterrichtsfach in unterschiedlichen Kontexten und mit verschiedenen gelagerten Lernzielen sowie auf unterschiedliche Weise erheben. So haben Jordan et al. (2008) im Rahmen der COACTIV-Studie („Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz“, Kunter et al., 2011) das kognitive Aktivierungspotenzial von Mathematikaufgaben im Unterricht, in Klassenarbeiten oder in Hausaufgaben erfasst, indem unter anderem die Anzahl der eingeforderten Lösungswege, die Komplexität des notwendigen Verstehens mathematikhaltiger Texte oder die Anforderung an die Kombination von mathematischen Grundvorstellungen berücksichtigt wurden. Für den Deutschunterricht haben Lotz et al. (2011) die Vermittlung von Lesestrategien als potenziell kognitiv aktivierend herausgestellt. Stahns (2013) ergänzt, dass Fragen, die zu ausführlichen und inhaltlich reichhaltigen Schüler:innenbeiträgen führen, als kognitiv aktivierend angesehen werden können. In eine ähnliche Richtung zielen auch Fauth et al. (2014) für den Sachunterricht in der Grundschule („Aufgaben, die zum Nachdenken anregen“). Englert und Reese-Schnitker (2011) stellen schließlich für den Religionsunterricht die Wichtigkeit von Bezügen zu religiösen Traditionen und relevanten Medien heraus. Hier stellt sich allerdings die Frage, inwiefern wirklich substanzielle Unterschiede vorliegen oder jeweils nur neue Begrifflichkeiten zur Beschreibung der kognitiven Aktivierung gefunden wurden. Explizit befasst sich bislang keine Arbeit in der Fremdsprachendidaktik mit der hier diskutierten Möglichkeit, kognitive Aktivierung fachspezifisch zu fassen. Einen ersten Versuch unternehmen Guttke et al. (in diesem Band) für die Grundschule. Es ist zu erwarten, dass alle Versuche dieser Art sich stark an den Erkenntnissen der kognitiv-psychologisch geprägten L2/L3-Spracherwerbsforschung orientieren werden (vgl. DeKaiser, 2011).

3.3 Fachspezifische Messung von Unterrichtsqualität

Für die fachspezifische Erfassung von Unterrichtsqualität gilt es in einem ersten Schritt Beobachtungsinstrumente zu entwickeln, die in verschiedenen Fächern zur vergleichenden Erfassung von Unterrichtsqualität eingesetzt werden können (Maag Merki, 2011). Um die Ergebnisse eines solchen Vergleichs sinnvoll interpretieren zu können, muss sichergestellt sein, dass das Beobachtungsinstrument in allen untersuchten Fächern auf dieselbe Weise *funktioniert*. Ist das nicht der Fall, dann sind die zwischen Fächern beobachteten Unterschiede nicht zwangsläufig auf die Unterrichtsqualität zurückzuführen, sondern könnten ebenso in den Eigenschaften der Beobachtungsinstrumente und ihrer Anwendung begründet sein. In diesem Zusammenhang wird in der quantitativen Sozialforschung von (fehlender) Messinvarianz gesprochen (Van De Schoot et al., 2015). Messinvarianz bedeutet, dass eine Erhebungsmethode über verschiedene (Mess-)Situationen hinweg vergleichbare Ergebnisse liefert (z.B. Länder, Zeitpunkte, weitere Stichprobenmerkmale), ebenso wie eine Stoppuhr die Zeit sowohl in einer Turnhalle wie im Freien akkurat messen sollte. Im vorliegenden Fall begreifen wir also verschiedene Unterrichtsfächer als solche Messgelegenheiten, in denen Unterrichtsqualität vergleichbar erfasst werden soll.

Reusser und Pauli (2021) diskutieren etwa, inwiefern Klassenführung im Sportunterricht und in anderen Schulfächern vergleichbar erfasst werden kann. Etwas konkreter müsste man zunächst fragen: Weist erfolgreiche Klassenführung in unterschiedlichen Fächern dieselben Merkmale auf? Hinsichtlich der Hypothesenüberprüfung ist die folgende Frage relevant: Sind Unterrichtsstörungen im Sportunterricht ebenso bedeutsam für die Effektivität der Klassenführung wie in anderen Unterrichtsfächern? Muss letztere Frage für ein gegebenes Beobachtungsinstrument verneint werden, weil Unterrichtsstörungen im Sportunterricht beispielsweise durchweg eine geringere Rolle spielen als in anderen Fächern, so könnten beobachtete Unterschiede mit der (problematischen) Erfassung von Unterrichtsstörungen zusammenhängen. Aus einem solchen Befund könnte man schließen, dass mit dem eingesetzten Beobachtungsinstrument im Sportunterricht eine *andere Art* Klassenführung erfasst wurde (z.B. Klassenführung mit vs. ohne Berücksichtigung von Unterrichtsstörungen). Klassenführung hätte in diesem Sinne eine fachspezifische Bedeutung und wäre deshalb nicht fächerübergreifend vergleichbar.

Dreher und Leuders (2021) nennen die Kompetenzen der Beobachter:innen, die Unterrichtsbeobachtungen vornehmen, als zusätzliche Voraussetzung für die Erfassung von Unterrichtsqualität. Oftmals ist neben einem von den Beobachter:innen gemeinsam geteilten Verständnis von Unterrichtsqualität auch fachliches oder fachdidaktisches Wissen notwendig, um eine treffende Einschätzung im jeweiligen Unterrichtsfach vornehmen zu können (vgl. die Gedanken zur fachlichen Korrektheit in Abschnitt 2.1). Interessant ist in diesem Zusammenhang auch der Fall, in dem in einer fächerübergreifenden Studie Beobachter:innen mit *unterschiedlichen* Vorkenntnissen im einzuschätzenden Unterrichtsfach zum

Einsatz kommen (z. B. je zwei Studierende der Unterrichtsfächer Deutsch und Mathematik). Aus fachdidaktischer Sicht würde vermutlich stets Beobachter:innen der Vorzug gegeben, die über ein Mindestmaß an fachlichem und fachdidaktischem Wissen im entsprechenden Unterrichtsfach verfügen (d. h., es würden die Mathematikstudierenden im Mathematik- und die Deutschstudierenden im Deutschunterricht eingesetzt werden). Aus psychometrischer Sicht wäre dagegen ein Studiendesign vorzuziehen, in dem Beobachter:innen Unterricht so oft wie möglich gemeinsam (fächerübergreifend) einschätzen, um die Vergleichbarkeit der Beobachtungen überhaupt prüfen zu können. Das für die beiden dargestellten Sichtweisen optimale Szenario, Studierende der Fächer Deutsch *und* Mathematik als Beobachter:innen zu gewinnen, wird realistisch wohl nur selten umzusetzen sein.

3.4 Fachspezifische Effekte von Unterrichtsqualität

Sind die unter 3.3 diskutierten Voraussetzungen gegeben, so lassen sich beobachtete fachspezifische Unterschiede in der Unterrichtsqualität interpretieren, und zwar sowohl zwischen Klassen als auch „innerhalb“ von Klassen, wenn etwa dieselbe Lehrperson in mehreren Unterrichtsstunden oder in verschiedenen Unterrichtsfächern beobachtet wurde. Im Rahmen eines solchen Designs haben Praetorius et al. (2015) Klassenführung und konstruktive Unterstützung bei 25 Lehrkräften untersucht, die in derselben Klasse sowohl Deutsch als auch Englisch unterrichten haben. Für die Klassenführung fanden sich nur geringfügige Unterschiede, die konstruktive Unterstützung unterschied sich allerdings beträchtlich zwischen den Fächern. In ähnlicher Weise haben Cohen et al. (2018) die Unterrichtsqualität bei über 300 Grundschullehrkräften verglichen, die in denselben Klassen Englisch und Mathematik unterrichten haben. Auch in dieser Studie fanden sich stärkere Zusammenhänge für Aspekte der Klassenführung und etwas schwächere für Merkmale konstruktiver Unterstützung, die Unterschiede waren aber insgesamt gering (Korrelationen zwischen Fächern: Classroom Organization $r = .55$, Emotional Support $r = .52$, Instructional Support $r = .42$, auf Itemebene $r = .25$ bis $r = .55$).

Als abschließendes Beispiel bemerken wir, dass sich Fachspezifität auch auf die Zusammenhänge zwischen Unterrichtsqualität und anderen Variablen beziehen kann, wenn diese sich über verschiedene Unterrichtsfächer hinweg unterscheiden (z. B. Schüler:innenleistung). Ist etwa Klassenführung in mehreren Fächern erhoben worden und es ergeben sich beispielsweise unterschiedliche Effekte je Unterrichtsfach auf das Allgemeinwissen von Schüler:innen, so könnten wir auch hier von fachspezifischen Befunden sprechen. Diese Lesart von Fachspezifität wird zwar bisweilen in internationalen Vergleichsstudien angewendet, ist in Veröffentlichungen zur Unterrichtsqualität bisher aber nur selten aufgegriffen worden und stellt deshalb ein zentrales Themenfeld für entsprechende interdisziplinäre Projekte dar (z. B. Gürtner et al., in diesem Band; Schilcher et al., 2021).

4 Desiderate einer fachdidaktischen Konkretisierung von Unterrichtsqualität

Wie voranstehend dargelegt werden von der Lehr-Lernforschung und den Fachdidaktiken einerseits unterschiedlich starke Konkretisierungen vorgenommen, um Unterrichtsqualität über verschiedene Fächer hinweg zu untersuchen. Andererseits enthält die Deutung generischer Operationalisierungen zwangsläufig fachspezifische Anteile. Abhängig von der Forschungsfrage (und vermutlich auch von der jeweiligen Fachkultur, Grossman & Stodolsky, 1995) werden fächerübergreifende theoretische Modelle zugrunde gelegt (etwa die drei Basisdimensionen; Kunter & Ewald, 2016) und entsprechend sensitive Erhebungsinstrumente eingesetzt (z.B. Cohen et al., 2018). Um diese Ausführungen durch forschungspraktische Beispiele zu illustrieren, werden nachfolgend einige Desiderate einer „fachdidaktisch konkretisierten Unterrichtsforschung“ (Reusser & Pauli, 2021, S. 189; vgl. auch Lindl et al., in diesem Band) betrachtet. Diese erfordern aus Sicht der genannten Autor:innen eine Kooperation von Fachdidaktiken und (fachunabhängiger) Lehr-Lernforschung. Um dies zu verdeutlichen, wird im Folgenden auf drei der bei Reusser und Pauli (2021) formulierten Desiderate fachdidaktischer Theoriebildung zur kognitiven Aktivierung, zum adaptiven Feedback und zur Qualität von Unterrichtsgesprächen näher eingegangen.

4.1 Kognitive Aktivierung

Kognitive Aktivierung wird als vertieftes Nachdenken von Lernenden über Unterrichtsinhalte verstanden und gilt daher als besonders wichtig für deren Wissensaufbau (im Überblick Kunter & Ewald, 2016). Unklar ist jedoch beispielsweise, in welchen Unterrichtssituationen und durch welche Lehrkräftehandlungen kognitive Aktivierung im Fachunterricht begünstigt wird (Reusser & Pauli, 2021). Die Analyse dieser Aspekte könnte für den Erwerb eines geeigneten Handlungsrepertoires in der Aus- und Weiterbildung von Lehrkräften eine wichtige Rolle spielen. Die in Abschnitt 2.2 aufgeführten Operationalisierungsansätze sind erste Schritte in diese Richtung, dürften für passende Interventionen oder Handlungsempfehlungen für Lehrkräfte aber noch nicht ausreichend sein. Daher müssen die Beschreibung, die Erklärung, die Wirkung und die unterrichtliche Förderung von kognitiver Aktivierung auch fachdidaktisch konkretisiert und untersucht werden. Tabelle 2 liefert hierzu einige Beispiele (vgl. auch Kunter & Ewald, 2016).

Tabelle 2: Fachspezifische Konzeptualisierung und Operationalisierung der kognitiven Aktivierung

Fach	Konzeptualisierung	Datenerhebung	Quelle
Mathematik	Aufgaben mit multiplen Lösungen, Modellierungsaufgaben	Aufgabenanalyse	Jordan et al., 2008
Deutsch	Vermittlung von Lesestrategien	Beobachtung (Video)	Lotz et al., 2011
Sach- und Grammatikunterricht	Aufgaben, die zum Nachdenken bzw. langen Antworten anregen	Beobachtung (Video)	Fauth et al., 2014; Stahns, 2013
Geschichte, Religion, Mathematik	Metakognitive und diskursive Aktivitäten	Beobachtung (Video)	Nowinska, 2018
Englisch	Taxonomische Analyse im Englischunterricht	Beobachtung (Video)	Schadtle, 2023

Reusser und Pauli (2021) betonen zudem, dass sich Denk- und Lernprozesse im Unterricht vor allem auf die Klärung von *fachspezifischen* Sachverhalten beziehen (z. B. Verstehenselemente, Drollinger-Vetter, 2011), mit denen Lernende auf diese Weise *fachspezifische* Begriffe (re-)konstruieren. Damit geht einher, dass durch Unterrichtsfächer verschiedene *scientific communities* mit unterschiedlichen epistemischen Ansätzen und Methoden abgebildet werden (Grossman & Stodolsky, 1995), das Verständnis von Wissen und, wie dieses erworben wird, also fachspezifisch ist. Beispielsweise genügt eine Klärung von Sachverhalten z. B. in Form von Vokabelwissen im Fremdsprachenunterricht nicht, vielmehr müssen ebenso übende Elemente, die auf eine Habitualisierung abzielen, Bestandteil der Anforderungen an Unterrichtsqualität in einer Kombination zweier Dimensionen sein.

4.2 Adaptives Feedback

Damit möglichst viele Schüler:innen im Unterricht kognitiv aktiviert werden, sind nach ersten Ergebnissen adaptive Unterstützungsmaßnahmen (z. B. Feedback, Scaffolding) und Erklärungen durch die Lehrkräfte relevant (im Überblick Kunter & Ewald, 2016). Nach Reusser und Pauli (2021) benötigen globale Merkmale von Unterrichtsqualität gerade dann eine fachdidaktische Konkretisierung, wenn Lernaufgaben eine besondere Herausforderung für Schüler:innen darstellen und Verständnisschwierigkeiten das fachliche Lernen erschweren. Ein Beispiel für eine solche Konkretisierung stellen aktuelle fachdidaktische Projekte zur professionellen Unterrichtswahrnehmung von Lehrkräften dar (u. a. Grewe & Möller, 2020). Für diese besteht zumindest im Fach Mathematik ein direkter Zusammenhang zur Unterrichtsqualität (Blömeke et al., 2015; Kersting et al., 2012), dessen fächerübergreifende empirische Bestätigung allerdings noch aussteht.

4.3 Gegenstandsbezogene Unterrichtsgespräche

Wie in den Abschnitten 2.3 und 2.4 bereits an anderen Beispielen illustriert wurde, dürfte also auch die unterrichtliche Gesprächsführung ein Themenfeld für fächerübergreifende (z.B. Unterrichtsklima) und fachdidaktisch konkretisierte Forschungsfragen sein (z.B. Lehr-Lern-Diskursen in verschiedenen Unterrichtsfächern). Pauli und Reusser (2018) berichten Ergebnisse einer Studie, in der Lehrkräfte im Hinblick auf Qualitätsmerkmale für anspruchsvolle Unterrichtsgespräche in Mathematik und Geschichte fortgebildet wurden. Nach der Intervention erhöhte sich allerdings nicht nur die Beteiligung der Schüler:innen am Unterricht, sondern auch deren fachlicher Gehalt. Die Autor:innen erklären dieses Phänomen damit, dass die teilnehmenden Lehrkräfte im Anschluss an die Fortbildung über ein umfangreiches Strategierepertoire verfügten, mit dem sie die Schüler:innen dazu motivieren konnten, sich am Unterrichtsgeschehen zu beteiligen. Hierbei handelte es sich sowohl um fächerübergreifende als auch um fachdidaktische Strategien der unterrichtlichen Gesprächsführung und des Einsatzes von Unterrichtsmaterialien (z.B. Aufgaben, Jordan et al., 2008). Insgesamt liegt die Vermutung nahe, dass Aspekte der Unterrichtsgestaltung, die direkt den Aufbau von Kompetenzen adressieren, vor allem eine fachspezifische Operationalisierung (ggf. auch Konzeptualisierung) der untersuchten Qualitätsmerkmale erfordern. Demnach könnten beispielsweise auch die unterrichtliche Strukturierung (Phasierung) in einigen Fällen sinnvoll fachdidaktisch konkretisiert werden wie etwa die Phasen des Englischunterrichts an der Grundschule. Allerdings wird diese bei Helmke (2012) und anderen Autor:innen als fächerübergreifendes Qualitätsmerkmal aufgeführt.

5 Diskussion

Im Beitrag wurden verschiedene Konzeptualisierungen von Fachspezifität dargestellt, die in der Unterrichtsqualitätsforschung Verwendung finden, in der Literatur jedoch noch selten diskutiert werden. Manche Fragestellungen setzen implizit ein fächerübergreifendes theoretisches Modell oder Erhebungsinstrument voraus, um in einer empirischen Studie fachspezifische Befunde generieren und diese wiederum fächerübergreifend interpretieren zu können. Diese Überlegungen sind insbesondere im Rahmen von Effektivitätsforschung (im Überblick Seidel & Shavelson, 2007) bedenkenswert, wenn für ein fächerübergreifendes Forschungsinteresse zusätzliche Anforderungen an die Messinstrumente gestellt werden müssen, damit eine Vergleichbarkeit der Befunde über unterschiedliche Fächer hinweg gewährleistet ist. Dagegen kann eine Konkretisierung zentral sein, wenn die (fachspezifische) Professionalisierung von Lehrpersonen im Vordergrund des Forschungsinteresses steht (Bromme, 1995; Grewe & Möller, 2020), weil gewisse Unterrichtsqualitätsmerkmale beispielsweise im Rahmen einer Lehrkräftefortbildung fokussiert werden sollten.

Einschränkend ist anzumerken, dass dieser Beitrag beispielhaft bleiben muss und nur einige Anwendungsfälle von fachspezifischer Konkretisierung in der Unterrichtsforschung beleuchtet werden konnten. Diskutiert wurde etwa, inwieweit sich die Erhebung des kognitiven Aktivierungspotenzials im Unterricht für eine fachspezifische Konkretisierung eignet. Dieses Beispiel wurde gewählt, weil kognitive Aktivierung von Klieme und Rakoczy (2008) bereits im Anschluss an die „Pythagoras-Studie“ als diejenige Basisdimension identifiziert wurde, die eine solche Konkretisierung erfordert. In den letzten Jahren hat sich herausgestellt, dass viele Fachdidaktiker:innen diese Ansicht teilen (u. a. Englert & Reese-Schnicker, 2011; Kranefeld, 2020; Wilden, 2021). Eine fachdidaktische Konkretisierung bietet sich allerdings auch für andere Dimensionen der Unterrichtsqualität an. So beschreibt Kleickmann (2012) in Anlehnung an das *Classroom Assessment Scoring System* (CLASS, Pianta & Hamre, 2009) eine fachdidaktische Konkretisierung der Qualitätsdimension Strukturierung für den Sachunterricht. Diese Dimension schließt das adaptive Unterstützungsverhalten von Lehrpersonen im (Fach-) Unterricht ein, welches auch im vorliegenden Beitrag diskutiert wurde.

Schließlich soll noch darauf hingewiesen werden, dass eine Konkretisierung der Modellierung oder Erfassung von Unterrichtsqualität nicht nur für verschiedene Fächer, sondern auch „innerhalb“ eines Fachs vorgenommen werden kann, wenn beispielsweise ein themenspezifischer Fokus vorliegt (z. B. in der TALIS-Videostudie Deutschland, Grünkorn et al., 2020; vgl. auch Lipowsky et al., 2009 für die „Pythagoras-Studie“ bzw. allgemein Lindl et al., in diesem Band): Da das professionelle Wissen von Lehrpersonen inhaltsspezifisch organisiert ist, dürfte auch die (fachspezifische) Unterrichtsqualität themenabhängig variieren. Es ist beispielsweise naheliegend, dass sich das fremdsprachliche Lehren und Lernen von Lexik und Grammatik in vielerlei Hinsicht von der Vermittlung literaturästhetischer Merkmale oder von kulturellem Lernen unterscheidet. Sind diese Unterschiede für ein Forschungsprojekt relevant, so sollten sie durch fachdidaktische Konkretisierungen der Qualitätsmerkmale berücksichtigt werden. Seidel et al. (2021) schlagen deshalb die Definition *fachspezifischer* Standardsituationen vor, um fächerübergreifende, fach- und themenspezifische Perspektiven zusammenzubringen („core practices“, Grossman, 2018; Kernpraktiken, Fraefel, 2022). Solche Standardsituationen, die für alle Fächer relevant sind, sind beispielsweise die Anleitung von Gruppendiskussionen oder das Einholen und Interpretieren des Denkens der Schüler:innen (Teaching Works, 2023). Für die Forschung zur Unterrichtsqualität und für die Aus- sowie Weiterbildung von Lehrkräften dürfte dieses Konzept zukünftig eine wichtige Rolle spielen und sollte deshalb auch von den Fachdidaktiken nutzbar gemacht und vorangetrieben werden.

Literatur

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Bromme, R. (1995). Was ist „pedagogical content knowledge“? Kritische Anmerkungen zu einem fruchtbaren Forschungsprogramm. *Zeitschrift für Pädagogik, Beiheft* 33, 105–115.
- Brunner, E. (2018). Qualität von Mathematikunterricht: Eine Frage der Perspektive. *Journal für Mathematik-Didaktik*, 39(2), 257–284. <https://doi.org/10.1007/s13138-017-0122-z>
- Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Bristol: Multilingual Matters.
- Charalambous, C. & Praetorius, A.-K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM – Mathematics Education*, 50(3), 355–366. <https://doi.org/10.1007/s11858-018-0914-8>
- Cohen, D. K., Raudenbush, S. W. & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142. <https://doi.org/10.3102/01623737025002119>
- Cohen, J., Ruzek, E. & Sandilos, L. (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open*, 4(3), 1–16. <https://doi.org/10.1177/2332858418794492>
- DeKeyser, R. (2011). Cognitive-Psychological Processes in Second Language Learning. In M. H. Long & C. J. Doughty (Hrsg.), *The Handbook of Language Teaching* (S. 119–138). Chichester: Wiley Blackwell. <https://doi.org/10.1002/9781444315783.ch8>
- Dreher, A. & Leuders, T. (2021). Fachspezifität von Unterrichtsqualität – aus der Perspektive der Mathematikdidaktik. *Unterrichtswissenschaft*, 49(2), 285–292. <https://doi.org/10.1007/s42010-021-00116-9>
- Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit. Fachdidaktische Qualität der Anleitung von mathematischen Verstehensprozessen im Unterricht*. Münster: Waxmann.
- Englert, R. & Reese-Schnicker, A. (2011). Varianten korrelativer Didaktik im Religionsunterricht – Eine Essener Unterrichtsstudie. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, I. H. Schön, H. J. Vollmer & H.-G. Weigand (Hrsg.), *Empirische Fundierung in den Fachdidaktiken* (S. 59–73). Münster: Waxmann.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3), 127–137. <https://doi.org/10.1024/1010-0652/a000129>
- Fraefel, U. (2022). Mittels Kernpraktiken zu professionellem Unterrichten. *Journal für LehrerInnenbildung*, 22(3), 16–29. <https://doi.org/10.35468/jlb-03-2022-01>
- Grewe, O. & Möller, K. (2020). Die professionelle Unterrichtswahrnehmung von sprachsensiblen Maßnahmen im Sachunterricht der Grundschule fördern – ein video- und praxisbasiertes Seminar im Master of Education. *Herausforderung Lehrer*innenbildung – Zeitschrift zur Konzeption, Gestaltung und Diskussion*, 3(1), 323–359.

- Grossman, P. (2018). *Teaching core practices in teacher education*. Cambridge: Harvard Education Press.
- Grossman, P. & Stodolsky, S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(8), 5–23. <https://doi.org/10.3102/0013189X024008005>
- Grünkorn, J., Klieme, E., Praetorius, A.-K. & Schreyer, P. (2020). *Mathematikunterricht im internationalen Vergleich. Ergebnisse aus der TALIS-Videostudie Deutschland*. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation.
- Hattie, J. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement*. London: Routledge.
- Heinitz, B., Szogs, M., Förtsch, C., Korneck, F., Neuhaus, B. J. & Nehring, A. (2022). Unterrichtsqualität in den Naturwissenschaften. Eine vergleichende Gegenüberstellung von Ansätzen zwischen Fachspezifik und Generik. *Zeitschrift für Didaktik der Naturwissenschaften*, 28(1), 10. <https://doi.org/10.1007/s40573-022-00146-5>
- Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (4. Aufl.). Seelze-Velber: Klett-Kallmeyer.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In Eckhard Klieme (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 345–363). Weinheim: Beltz Pädagogik.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L. & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511. <https://doi.org/10.1080/07370000802177235>
- Hu, A. & Byram, M. (2009). *Interkulturelle Kompetenz und fremdsprachliches Lernen. Modelle, Empirie, Evaluation / Intercultural Competence and Foreign Language Learning. Models, Empiricism, Assessment*. Tübingen: Gunter Narr.
- Jentsch, A., Schlesinger, L., Heinrichs, H., Kaiser, G., König, J. & Blömeke, S. (2020). Erfassung der fachspezifischen Qualität von Mathematikunterricht: Faktorenstruktur und Zusammenhänge zur professionellen Kompetenz von Mathematiklehrpersonen. *Journal für Mathematik-Didaktik*, 42(1), 97–121. <https://doi.org/10.1007/s13138-020-00168-x>
- Jordan, A., Krauss, S., Löwen, K., Blum, W., Neubrand, M., Brunner, M., Kunter, M. & Baumert, J. (2008). Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven Aktivierungspotentials im deutschen Mathematikunterricht. *Journal für Mathematik-Didaktik*, 29(2) 83–107. <https://doi.org/10.1007/BF03339055>
- Kane, T. J. & Staiger, D. O. (2012). *Gathering feedback for teachers: Combining high-quality observations with student surveys and achievement gains*. Bill and Melinda Gates Foundation.
- Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R. & Stigler, J. W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, 49(3), 568–589. <https://doi.org/10.3102/0002831212437853>
- Kleickmann, T. (2012). *Kognitiv aktivieren und inhaltlich strukturieren im naturwissenschaftlichen Sachunterricht*. Kiel: IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik an der Universität Kiel. http://www.sinus-an-grundschulen.de/fileadmin/uploads/Material_aus_SGS/Handreichung_Kleickmann.pdf
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.

- Kranefeld, U. (2021). Der Diskurs um Unterrichtsqualität in der Musikdidaktik zwischen generischen und fachspezifischen Dimensionen. *Unterrichtswissenschaft*, 49(2), 221–233. <https://doi.org/10.1007/s42010-021-00113-y>
- Kunter, M. & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Hotappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Münster: Waxmann.
- Lindl, A., Ehrich, P., Gutmiedl, M., Rader, M., Simböck, L., Gürtner, M., Böhringer, S., Krämer, A., Kirchoff, J. & Frei, M. (2024). Und wo bleibt die Ästhetik? – Betrachtungen zu einer weiteren Dimension von Unterrichtsqualität aus interdisziplinärer Perspektive. In M. Hemmer, C. Angele, C. Bertsch, S. Kapelari, G. Leitner & M. Rothgangel (Hrsg.), *Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung* (S. 371–388). Münster: Waxmann.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E. & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Lotz, M., Lipowsky, F. & Faust, G. (2011). Kognitive Aktivierung im Leseunterricht der Grundschule. Konzeptionelle Überlegungen und erste empirische Ergebnisse zu ausgewählten Merkmalen kognitiv aktivierender Unterrichtsgespräche. *Osnabrücker Beiträge zur Sprachförderung*, 80, 146–165.
- Maag Merki, K. (2011). The introduction of state-wide exit examinations. Empirical effects on math and English teaching in German academically oriented secondary schools. In M. A. Pereyra, H.-G. Kotthoff & R. Cowen (Hrsg.), *PISA under Examination. Changing knowledge, changing tests, and changing schools* (S. 125–142). Rotterdam: Sense Publishers.
- Pianta, R. C. & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38(2), 109–119. <https://doi.org/10.3102/0013189X09332374>
- Pietsch, M. & Tosana, S. (2008). Beurteilereffekte bei der Messung von Unterrichtsqualität. Das Multifacetten-Rasch-Modell und die Generalisierbarkeitstheorie als Methoden der Qualitätssicherung in der externen Evaluation von Schulen. *Zeitschrift für Erziehungswissenschaft*, 11(3), 430–452. <https://doi.org/10.1007/s11618-008-0021-7>
- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings*. Münster: Waxmann.
- Praetorius, A. K. & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM – Mathematics Education*, 50(3), 535–553. <https://doi.org/10.1007/s11858-018-0946-0>
- Praetorius, A.-K., Vieluf, S., Saß, S. & Bernholt, A. (2015). The same in German as in English? Investigating the subject-specificity of teaching quality. *Zeitschrift für Erziehungswissenschaft*, 19(1), 191–209. <https://doi.org/10.1007/s11618-015-0660-4>
- Rakoczy, K. & Pauli, C. (2006). Hoch inferentes Rating: Beurteilung der Qualität unterrichtlicher Prozesse. In I. Hugener, C. Pauli & K. Reusser (Hrsg.), *Videoanalysen. Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“* (Bd. 3, S. 206–233). Frankfurt am Main: Gesellschaft zur Förderung Pädagogischer Forschung (GFPF)/Deutsches Institut für Internationale Pädagogische Forschung (DIPF).
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs, P. Füglistler, K. Reusser & H. Wyss

- (Hrsg.), *Didaktik auf psychologischer Grundlage. Von Hans Aebli kognitionspsychologischer Didaktik zur modernen Lehr- und Lernforschung* (S. 151–168). Bern: hep.
- Reusser, K. & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49(2), 189–202. <https://doi.org/10.1007/s42010-021-00117-8>
- Schilcher, A., Krauss, S., Kirchhoff, P., Lindl, A., Hilbert, S., Asen-Molz, K., Ehras, C., Elmer, M., Frei, M., Gaier, L., Gastl-Pischetsrieder, M., Gunga, E., Murmann, R., Röhrl, S., Ruck, A.-M., Weich, M., Dittmer, A., Fricke, M., Hofmann, B., ... Thim-Mabrey, C. (2021). FALKE: Experiences from transdisciplinary educational research by fourteen disciplines. *Frontiers in Education*, 5, 579982. <https://doi.org/10.3389/educ.2020.579982>
- Schlesinger, L., Jentsch, A., Kaiser, G., König, J. & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM – Mathematics Education*, 50(3), 475–490. <https://doi.org/10.1007/s11858-018-0917-5>
- Schwarz, S. (2019). *SchülerInnenperspektiven und Religionsunterricht. Empirische Einblicke – Theoretische Überlegungen*. Stuttgart: Kohlhammer. <https://doi.org/10.17433/978-3-17-037469-0>
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Stahns, R. (2013). *Kognitive Aktivierung im Grammatikunterricht. Videoanalysen zum Deutschunterricht*. Scheider Hohengehren.
- Taut, S. & Rakoczy, K. (2016). Observing instructional quality in the context of school evaluation. *Learning and Instruction*, 46, 45–60. <https://doi.org/10.1016/j.learninstruc.2016.08.003>
- Teaching Works (2023). *Core Teaching Practices*. Verfügbar unter <https://library.teaching-works.org/curriculum-resources/high-leverage-practices> [15.11.2023].
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Van de Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K. & Zondervan-Zijnenburg, M. (2015). Editorial: Measurement Invariance. *Frontiers in Psychology*, 6, 1064. <https://doi.org/10.3389/fpsyg.2015.01064>
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>

Inhaltsvalidierung von Fragebogenitems zu kognitiver Aktivierung im Englischunterricht der Grundschule mittels Expert:innenbefragung

Abstract der Herausgeber:innen

In their article “Content validation of questionnaire items on cognitive activation in primary English as a Foreign Language (EFL) lessons using an expert survey,” Joel Guttke and Raphaela Porsch focus on the role of the basic dimension of cognitive activation in EFL lessons. Up to now, the findings on this central dimension of instructional quality have been rather inconsistent. Therefore, this contribution takes a closer empirically founded look at this dimension. It presents the results of an online survey with selected experts on the content validity of theory-based items designed to collect data on cognitive activation in fourth-grade EFL classes. The results suggest that the majority of items tap into construct-relevant aspects. Implications for the operationalisation of the construct and the revision of items will then be derived. The approach described also highlights the added value of expert surveys in the process of constructing new instruments in research-based foreign language teaching.

1 Einleitung

Im deutschsprachigen Raum gelten die drei Basisdimensionen von Unterrichtsqualität als etabliertes Modell zur empirischen Erfassung von Unterrichtsmerkmalen und Erklärung der Lernwirksamkeit von Unterricht (Klieme, 2019). Im Gegensatz zu anderen Modellen der Unterrichtsqualität fußt das Modell hinsichtlich seiner faktoriellen Struktur auf einer vergleichsweise stabilen empirischen Basis. In Bezug auf die prädiktive Validität für Schüler:innenleistungen fällt die Befundlage jedoch – insbesondere für die Basisdimension der kognitiven Aktivierung – inkonsistent aus (z. B. Praetorius et al., 2018). Dennoch wird kognitiver Aktivierung auch in anderen Modellen der Unterrichtsqualität weiterhin eine zentrale Rolle zugeschrieben (Lindl et al., in diesem Band). Die Inkonsistenz, die sich in den empirischen Daten widerspiegelt, wird häufig auf einen Mangel in der theoretischen Fundierung des Konstrukts zurückgeführt (Praetorius & Gräsel, 2021). Um das Konstruktverständnis zu schärfen, wurden jüngst Überlegungen artikuliert, kognitive Aktivierung fachlich zu spezifizieren (Begrich et al., 2023) oder die bisherigen Basisdimensionen von Unterrichtsqualität weiterhin als generisch zu interpretieren und um weitere, fachspezifische Basisdimensionen zu ergänzen (z. B. Lipowsky & Bleck, 2019).

In der Fremdsprachenforschung wurde die Frage nach einer fachspezifischen Konstruktdefinition und -operationalisierung kognitiver Aktivierung bisher nicht systematisch untersucht. Dies ist jedoch eine notwendige Voraussetzung für die valide Erfassung des Konstrukts. Für das Fach Englisch existieren zwar erste Beobachtungsinstrumente (z.B. Kersten et al., 2018) und Fragebogenskalen (z.B. Porsch & Wilden, 2022), die kognitive Aktivierung im Fach Englisch zu erfassen vermögen (Lindl et al., in diesem Band). Der Entwicklungsprozess für diese Instrumente und die Prüfung von Konstruktvalidität als ein Schritt zur qualitativen Eignung wurden bislang jedoch nicht beschrieben. Aufgrund der konzeptionellen Unschärfe kognitiver Aktivierung ist insbesondere die Diskussion über Inhaltsvalidität von zentraler Bedeutung. Dies gilt umso mehr, wenn die mit diesen Instrumenten erhobenen Daten Einsatz in der evidenzbasierten Unterrichtsentwicklung finden sollen (Grünkorn et al., 2018).

Der vorliegende Beitrag adressiert das skizzierte Forschungsdesiderat in dreifacher Hinsicht. Erstens werden Ergebnisse einer Online-Expert:innenbefragung zur Inhaltsvalidität neukonstruierter Items zur Erfassung kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 mittels Schüler:innen- und Lehrkräftefragebögen dargelegt. Zweitens werden anhand der Ergebnisse Implikationen hinsichtlich der Konstruktoperationalisierung und Itemrevision abgeleitet. Schließlich leistet die Studie einen Beitrag zur Fragebogenentwicklung in der Fremdsprachenforschung, indem sie die Online-Expert:innenbefragung als eine Methode zur Entwicklung von Items aufzeigt, die bislang selten zur Anwendung gekommen ist.

2 Forschungsstand

2.1 Kognitive Aktivierung und die Notwendigkeit einer fachlichen Konstruktspezifikation

Dieser Abschnitt fasst jüngste Entwicklungen in der Erforschung kognitiver Aktivierung als Basisdimension von Unterrichtsqualität zusammen. Dazu werden zunächst verschiedene Facetten kognitiver Aktivierung begrifflich voneinander getrennt, die im bisherigen Forschungsdiskurs nur implizit unterschieden werden. Daraufhin werden Potenziale einer fachlichen Spezifikation kognitiver Aktivierung genannt, die in einer Konstruktdefinition für den Englischunterricht der Primarstufe münden.

Ursprünglich stellte kognitive Aktivierung im Kontext der TIMS-Studie einen faktorenanalytisch identifizierten Oberbegriff dar, der Unterrichtsmerkmale zur „Komplexität von Aufgabenstellungen und Argumentationen und die Intensität des fachlichen Lernens“ (Klieme et al., 2001, S. 51) umfasste. Seit dieser Veröffentlichung wurden die drei Basisdimensionen von Unterrichtsqualität über den Mathematikunterricht hinaus intensiv rezipiert und empirisch untersucht. Dies führte gleichzeitig zu einer Diffusion des Konstrukts kognitiver Aktivierung aufgrund

variierender theoretischer Verankerungen und Operationalisierungen (Schreyer, 2024). Derzeit erscheint ein Verständnis von kognitiver Aktivierung als Anregung „zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand“ (Lipowsky, 2020, S. 92) konsensfähig.

Die Unbestimmtheit der Konstruktdefinition stellt eine Herausforderung in der Erforschung des Gegenstands dar, da die Grenzen zu weiteren Dimensionen der Unterrichtsqualität verschwimmen. Dies hat zur Folge, so stellen Lipowsky und Hess (2019) fest, dass unter kognitiver Aktivierung „auch inhaltlich teilweise recht unterschiedliche Maßnahmen der Lehrperson und/oder Aktivitäten der Lernenden verstanden werden“ (S. 81). In einer Zusammenschau ausgewählter empirischer Studien identifizieren sie zehn Facetten kognitiver Aktivierung, die von der ko-konstruktiven Weiterentwicklung von Schüler:innenvorstellungen im Unterrichtsgespräch bis hin zur Adaptivität des Unterrichts reichen. Daraus leiten sie zwölf fachunabhängige Maßnahmen zur kognitiven Aktivierung der Lernenden für die Unterrichtspraxis ab. Zu diesen Maßnahmen zählen unter anderem die „[s]ystematische Variation von Aufgaben(teilen) und Anregung zur Entdeckung von Regelmäßigkeiten“ sowie der Einbezug von Aufgaben, „die über die Anwendung von Routinen hinausgehen“ (Lipowsky & Hess, 2019, S. 92).

Das Syntheseframework zu Unterrichtsqualität (Praetorius et al., 2020) ergänzt die drei Basisdimensionen auf Grundlage einer Synthese von zwölf internationalen *frameworks* zur Erfassung von Unterrichtsqualität um vier weitere Dimensionen. Darin wird – trotz häufiger Verweise auf die Abhängigkeit von kognitiver Aktivierung und Lerngegenständen – eine Trennung zwischen kognitiver Aktivierung und Unterrichtsinhalten vorgenommen, indem sie als zwei separate Unterrichtsqualitätsdimensionen formuliert werden. Kognitiver Aktivierung werden vier Subdimensionen zugeschrieben: (1) Auswahl und (2) Einsatz fachlich gehaltvoller und auf das kognitive Niveau der Schüler:innen abgestimmter Aufgaben sowie (3) Unterstützung der kognitiven Aktivität und (4) des metakognitiven Lernens der Schüler:innen anhand kognitiv aktivierender Aufgaben.

Bei genauerer Betrachtung dieser Konstruktbeschreibungen lassen sich aus den unterschiedlichen Indikatoren kognitiver Aktivierung drei separate Facetten herausarbeiten (Kleickmann et al., 2020; Rieser & Decristan, 2023): (1) das Potenzial zu kognitiver Aktivierung, (2) die kognitive Unterstützung und (3) die kognitive Aktiviertheit. Als Potenziale zu kognitiver Aktivierung gelten grundsätzlich alle Unterrichtsmerkmale, von denen begründet angenommen wird, dass sie der kognitiven Aktivierung von Schüler:innen dienlich sind. Inwiefern diese Potenziale aufseiten der Schüler:innen tatsächlich in kognitiver Aktiviertheit resultieren, ist von interindividuellen Faktoren und der Wahrnehmung des Unterrichtsangebots abhängig. Maßnahmen der kognitiven Unterstützung zielen darauf ab, die Ansprüche komplexer Lernumgebungen – die gewissermaßen notwendig für kognitive Aktivierung sind – zu reduzieren, sodass eine Bearbeitung durch unterschiedlich leistungsstarke Schüler:innen möglich wird. Ob diese drei Facetten Teil eines gemeinsamen Konstrukts kognitiver Aktivierung sind oder in Teilen separate Basisdimensionen der Unterrichtsqualität bilden, ist noch unklar.

Unstrittig ist hingegen, dass die drei Facetten kognitiver Aktivierung konform mit dem Angebots-Nutzungs-Modell schulischer Lehr-Lernprozesse sind (Helmke, 2014).

Die stark divergierenden Definitionen und Operationalisierungen kognitiver Aktivierung lassen sich zum Teil auf die drei beschriebenen Facetten zurückführen. Während die von Lipowsky und Hess (2019) identifizierte Offenheit von Problemstellungen ein Potenzial zu kognitiver Aktivierung darstellt, bildet die Individualisierung von Unterricht einen Indikator kognitiver Unterstützung. Die Unterteilung in drei Facetten kognitiver Aktivierung birgt Implikationen nicht nur auf inhaltlicher, sondern auch auf methodischer Ebene. Rieser und Decristan (2023) zeigen beispielsweise, dass sich Skalen, die konzeptionell zwischen dem Potenzial zu kognitiver Aktivierung und kognitiver Aktiviertheit unterscheiden, auch empirisch auf Individual- und Klassenebene voneinander trennen lassen und auf Klassenebene nicht signifikant miteinander korrelieren. Vor diesem Hintergrund ist es ratsam, bei der Neuentwicklung von Instrumenten zur Erfassung von kognitiver Aktivierung bereits während der Itemkonstruktion die drei Facetten des Konstrukts zu berücksichtigen.

Das Potenzial zu kognitiver Aktivierung wurde bisher größtenteils in Unterrichtsmaterialien wie Lehrwerken (für das Fach Englisch: Wilden et al., im Erscheinen) sowie Lernaufgaben verortet und generisch untersucht, jedoch nur selten als Teil des Lehrkräftehandelns begriffen. Maier et al. (2010) schlagen beispielsweise ein allgemeindidaktisches Kategoriensystem zur Analyse des kognitiven Potenzials von Aufgaben vor. Dazu beschreiben sie sieben Dimensionen zur Kategorisierung von Aufgaben: Art des Wissens, die zur Bewältigung der Aufgabe notwendigen kognitiven Prozesse, Anzahl der bei der Aufgabebearbeitung aktivierten Wissenseinheiten, Offenheit der Aufgabenstellung, Lebensweltbezug, sprachlogische Komplexität und Repräsentationsformen des Wissens. Wird das Kategoriensystem zur Analyse von Aufgabenmaterial (*task as workplan*) eingesetzt, lässt sich daran eine Aussage über das Potenzial zur kognitiven Aktivierung treffen, das wiederum von der Implementation des Aufgabenmaterials in das Unterrichtsgeschehen (*task in process*) und den dadurch initiierten Lernprozessen zu unterscheiden ist.

Die beispielhaft aufgeführten – und bei Weitem nicht vollständigen – Facetten und Indikatoren kognitiver Aktivierung tragen dazu bei, das Konstrukt im Sinne einer extensionalen Nominaldefinition einzugrenzen (Döring & Bortz, 2016). In der Anwendung und Operationalisierung erfordern diese Indikatoren in Abhängigkeit von dem jeweils betrachteten Schulfach dennoch einen Transferschritt, zu dessen Durchführung fachdidaktische Forschung unerlässlich ist. So stellt sich beispielsweise die Frage, welche Charakteristika im Englischunterricht fachlich gehaltvolle Aufgaben kennzeichnen oder welche Bestandteile einer Aufgabe auf ihre sprachlogische Komplexität hin zu prüfen sind (lediglich die Aufgabeninstruktion oder auch der Ausgangstext bei einer Aufgabe zum Leseverstehen; Wilden et al., im Erscheinen). Des Weiteren werden einige dieser Fragen

erst durch einen komparativen Ansatz aus Perspektive unterschiedlicher Fachdidaktiken beantwortbar (z. B. Lindl et al., in diesem Band; Praetorius et al., 2020).

Im Forschungsdiskurs zu Unterrichtsqualität bildet sich die Tendenz zu einem fachspezifischen Verständnis kognitiver Aktivierung in zweierlei Hinsicht ab. Einige Forscher:innen sprechen sich für eine „Erweiterung kognitiver Aktivierung über Fächer hinweg oder aber lediglich für einzelne Fächer hin zu einer kognitiv-motorischen, kognitiv-ästhetischen oder kommunikativ-kognitiven Aktivierung“ (Praetorius & Gräsel, 2021, S. 178) aus. Andere Ansätze schlagen hingegen vor, „diese ergänzenden Aspekte nicht als Teil kognitiver Aktivierung, sondern als eigenständige Dimensionen zu behandeln“ (Praetorius & Gräsel, 2021, S. 178; Lindl et al., 2024). Ungeachtet dieser Entscheidung birgt die fachliche Spezifikation kognitiver Aktivierung das Potenzial, die Lernwirksamkeit des Fachunterrichts differenzierter und damit den Gegenstand angemessener zu untersuchen.

Für die fachspezifische Operationalisierung kognitiver Aktivierung sind die Bildungsziele eines Fachs von zentraler Bedeutung. In den Bildungsstandards für den ersten und mittleren Schulabschluss wird die Entwicklung funktionaler kommunikativer Kompetenz als übergeordnetes Ziel des Fremdsprachenunterrichts postuliert (KMK, 2023). Ergänzend werden Kompetenzbereiche wie interkulturelle Kompetenz, Sprachbewusstheit oder Text- und Medienkompetenz als nebengeordnete Ziele benannt. Aufgrund ihrer nebengeordneten Rolle und um den Umfang der zu entwickelnden Fragebögen zu begrenzen, werden diese Kompetenzbereiche in diesem Forschungsprojekt nicht explizit adressiert, obgleich sie vereinzelt in Items abgebildet sind. Dies soll nicht bedeuten, dass den fünf nebengeordneten Kompetenzbereichen keine Relevanz hinsichtlich der Qualität des Englischunterrichts zugeschrieben wird; stattdessen steht die Untersuchung kognitiver Aktivierung während des Zweitspracherwerbs im Zentrum des Forschungsinteresses.

Zum Fremdsprachenunterricht liegen aus dem deutschsprachigen Raum bisher nur vereinzelt konzeptionelle Überlegungen und empirische Befunde vor, die sich im weitesten Sinne der Untersuchung von kognitiver Aktivierung zuordnen lassen. Guttke (2023) identifiziert in einem systematischen Forschungsreview sechs Aspekte, die kognitiver Aktivierung im Fremdsprachenunterricht dienlich seien: die Auseinandersetzung mit einer kommunikativen Problemstellung, der zielgerichtete Einsatz der Zielsprache in einem bedeutsamen Kontext, die kritische Reflexion des eigenen Sprachhandelns, vielfältiger und verständlicher *language input*, Gelegenheit zur Produktion von *language output* während der Interaktion in der Zielsprache sowie begleitendes *corrective feedback*. Anhand dieser Überlegungen lässt sich eine vorläufige, fachspezifische Konstruktdefinition formulieren: Kognitive Aktivierung im Fremdsprachenunterricht entspricht einer möglichst intensiven Verarbeitung der Zielsprache. Unter Berücksichtigung der Forschung zu (*Instructed*) *Second Language Acquisition* (Gass et al., 2020, S. 575–589; Loewen, 2020) und der sechs Aspekte kognitiver Aktivierung im Fremdsprachenunterricht (Guttke, 2023) drückt sich eine solche möglichst intensive Verarbeitung der Zielsprache potenziell in drei Prozessen aus: (1) in der Transformation von *language*

input zu *intake*, (2) in Schüler:innenreaktionen auf *corrective feedback (uptake)* sowie (3) während der Produktion zielsprachlicher Äußerungen (*comprehensible/pushed output*).

2.2 Erfassung kognitiver Aktivierung mittels Fragebögen

Im Folgenden werden schriftliche Befragungen von Schüler:innen und Lehrkräften mittels Fragebögen als Methode zur Erfassung von Unterrichtsqualität einschließlich kognitiver Aktivierung sowie empirische Befunde zu ihrer psychometrischen Güte präsentiert. Anhand dieser Befunde wird die Relevanz von Itemperspektive und -referenz in der Itemformulierung für die Entwicklung valider Fragebögen diskutiert, woraus Implikationen für die Itemkonstruktion im nachfolgend beschriebenen Promotionsprojekt abgeleitet werden.

Neben der Analyse von Aufgaben hinsichtlich ihres Potenzials zu kognitiver Aktivierung mittels Kodierschemata stellen Erhebungen mit Fragebögen, die eine Anzahl inhaltlich zusammengehöriger Items bzw. Aussagen zur Bewertung eines Sachverhalts umfassen, ein üblicherweise genutztes und ökonomisches Instrumentarium dar, um Unterrichtsqualität einschließlich kognitiver Aktivierung quantitativ zu messen. Die aus den Fragebogenerhebungen gewonnenen Kennwerte stellen Indikatoren für eine hohe oder geringe Qualität dar und können wiederum mit weiteren Indikatoren von Unterrichtsqualität, Prozessmerkmalen des Unterrichts, Merkmalen der Lehrkraft oder der Schüler:innen sowie *outcomes* des Unterrichts wie Schüler:innenleistungen in einen Zusammenhang gebracht werden (vgl. Lindl et al., in diesem Band). Theoretische Zusammenhänge dieser verschiedenen Merkmalsgruppen lassen sich mithilfe des Angebots-Nutzungs-Modells des Unterrichts von Helmke (2014) erklären. Zentral sind in dem Modell neben der Annahme von Zusammenhängen die Unterscheidung von Merkmalen der Lehrperson, des Unterrichts als ein Angebot, den Lernaktivitäten, d. h. die Nutzung des Angebots seitens der Schüler:innen, und den Wirkungen bzw. Ertrag wie den Leistungen oder Veränderungen non-kognitiver Merkmale (z. B. Selbstwirksamkeitsüberzeugungen). Dahingehend stellt sich für die Entwicklung von Skalen zur Erfassung von kognitiver Aktivierung im Unterricht die Frage, ob (a) das *Potenzial* zu kognitiver Aktivierung oder (b) die *Nutzung* des unterrichtlichen Angebots erfasst werden sollen (s. Abschnitt 2.1). Nachteil ist für letztgenannten Ansatz, dass damit Aspekte kognitiver Aktiviertheit abgebildet werden, die vorrangig für Schüler:innen zugänglich sind und nur eingeschränkt durch die Berücksichtigung weiterer Perspektiven auf ihre Gültigkeit geprüft werden können. Vor diesem Hintergrund fokussieren die im vorliegenden Beitrag präsentierten Items das Potenzial zu kognitiver Aktivierung.

Weiterführend lassen sich zur Einschätzung der Qualität des Unterrichts in Bezug auf das Potenzial zur kognitiven Aktivierung der Schüler:innen drei Perspektiven unterscheiden: die der Lehrkraft, die der Schüler:innen (individuell und in Form von aggregierten Klassenmerkmalen) und die von externen Beob-

achter:innen (einschließlich Videografie). Alle Perspektiven besitzen bestimmte Potenziale als auch Nachteile (z. B. Gruehn, 2000) und empirische Arbeiten zeigen, dass sich die Wahrnehmung der verschiedenen Gruppen unterscheidet (z. B. Fauth et al., 2014). Göllner et al. (2016) diskutieren insbesondere die Erfassung von Unterrichtsqualität aus Schüler:innensicht. Die Perspektive der Schüler:innen hat unter anderem den Vorteil, dass eine hohe Anzahl an Beurteiler:innen zur Verfügung steht, um das Unterrichtsgeschehen angemessen abzubilden. Laut den Autor:innen liegen relativ wenige Befunde zur psychometrischen Güte von Schüler:innenurteilen vor; die verfügbaren Befunde deuten jedoch darauf hin, dass sich Schüler:innenurteile als Datenquelle mit hinreichender Validität zur Erfassung von Unterrichtsqualität eignen. Insgesamt seien jedoch alle Perspektiven relevant und keine zu bevorzugen, da sich jede Perspektive für bestimmte Konfigurationen von Beurteilungsgegenstand und Itemreferenz besonders gut eignet (S. 70).

Schließlich können in Fragebögen neben den verschiedenen Personen Bewertungen von Merkmalen des Unterrichts oder eine Bewertung der Unterrichtsqualität in Bezug auf (a) eine spezifische Unterrichtsstunde oder (b) retrospektiv summierend Unterricht (z. B. bezogen auf das zurückliegende Schuljahr) in einem Fach eingeschätzt werden. In Bezug auf (a) kann es sich um spezifische, einfach beobachtbare Geschehnisse im Unterricht handeln. Alternativ sind hoch-inferente Beobachtungen (im Sinne von Ratings) möglich, die von der bzw. dem Beurteiler:in verlangen, „über das konkret beobachtete Verhalten hinaus auf abstrakte Sachverhalte oder allgemeine Verhaltenstendenzen zu schließen“ (Göllner et al., 2016, S. 64), was zum Erreichen reliabler Ergebnisse bestenfalls mit einer vorangegangenen Schulung verbunden ist. Grundsätzlich verlangt die Bewertung einer Aussage bzw. die Beantwortung eines Items unterschiedlich komplexe kognitive Prozesse.

Neben grundsätzlichen Anforderungen an die Itemkonstruktion (z. B. Eindeutigkeit, sprachliche Verständlichkeit), um die Beantwortung zu ermöglichen, können laut Göllner et al. (2016, S. 70) weitere Anforderungsmerkmale unterschieden werden: (a) Adressat:innenbezug des Items (Wer wird von den Items adressiert bzw. wer soll bewertet werden? Gesamte Gruppe/Klasse oder einzelne:r Schüler:in), (b) die Wahrnehmungsperspektive (Aus welcher Sicht erfolgt die Beurteilung? Gesamte Gruppe/Klasse oder einzelne:r Schüler:in) und (c) der Zeitbezug (Auf welchen Zeitraum bezieht sich die Beurteilung eines Items?), ein Aspekt, der bereits oben thematisiert wurde. Ergänzt um die Interaktion zwischen Lehrer:innen und Schüler:innen entwickelten Fauth et al. (2020) eine Matrix zur Item-Referenz (s. Abb. 1).

		Item-Reference											
		Teachers		Teachers/ Students		Students							
Rater Perspective	Teachers	Self Information	+ -	Ego-involve. +	-	Combi. Information	+ -	Ego-involve. +	-	Other Information	+ -	Ego-involve. +	-
	Students	Other Information	+ -	Ego-involve. +	-	Combi. Information	+ -	Ego-involve. +	-	Self Information	+ -	Ego-involve. +	-
	External Observers	Other Information	+ -	Ego-involve. +	-	Other Information	+ -	Ego-involve. +	-	Other Information	+ -	Ego-involve. +	-

Abbildung 1: Matrix zur Klassifikation von Items nach ihrer Referenz und Perspektive (Fauth et al., 2020, S. 144)

Die Matrix zeigt einerseits die drei Gruppen, die eine Beantwortung vornehmen können (Lehrkräfte, Schüler:innen, externe Beobachter:innen). Sie unterscheidet andererseits darin, dass sich der Inhalt eines Items auf das Handeln der Lehrkraft, die Schüler:innen oder die Interaktion zwischen beiden Gruppen beziehen kann. Schließlich wird klassifiziert, ob sie ihr eigenes Verhalten oder das anderer bewerten sollen. Die kritische Frage stellt sich, inwieweit jeweils für letzteren Punkt stets ausreichend Informationen zur Verfügung stehen (bspw. um als Schüler:in eine Aussage für die gesamte Klasse treffen zu können). Für die Bewertung des eigenen Verhaltens ist anzunehmen, dass diese bestimmten Verzerrungen unterworfen sein kann (u. a. Schutz des Selbstwerts oder im Sinne sozialer Erwünschtheit). Göllner et al. (2016) und Fauth et al. (2020) zeigen auf, dass Inventare zur Messung von Unterrichtsqualität die verschiedenen Anforderungen abbilden und mitunter sogar unterschiedliche Perspektiven innerhalb einer Skala eingenommen werden sollen. Auch der Zeitbezug ist nicht immer eindeutig für die Beurteiler:innen erkennbar.

Um kognitive Aktivierung angemessen zu operationalisieren, bedarf es daher neben der Festlegung auf eine fachspezifische Konstruktdefinition eines systematischen Vorgehens bei der Formulierung der Items und die Sicherstellung, dass die Befragten die Items mit ihnen zur Verfügung stehenden Informationen eindeutig für einen festgelegten Zeitraum beantworten können. In der nachfolgend vorgestellten Studie wurden diese Voraussetzungen in der Itemkonstruktion wie folgt umgesetzt: Bei den Lehrkräftemitems wurde darauf geachtet, dass die Items in der Ich-Perspektive formuliert sind und sich die Iteminhalte (sofern möglich) ausschließlich auf das Lehrkräfteverhalten beziehen. Die Schüler:innenitems wur-

den so formuliert, dass sie Beurteilungen des Lehrkräfteverhaltens auf Individual- statt auf Klassenebene elizitieren. In den Items beider Zielgruppen besteht eine klare Referenz hinsichtlich des zeitlichen (zurückliegende Englischstunde) und inhaltlichen Bezugs zu einem möglichst beobachtbaren Verhalten.

2.3 Expert:innenbefragungen als Instrument zur Inhaltsvalidierung von Fragebögen

Da der vorliegende Beitrag eine Validierungsstudie darstellt, fasst dieser Abschnitt das Konzept der Inhaltsvalidität zusammen. Darüber hinaus werden anhand empirischer Studien aus der Fachdidaktik und den Bildungswissenschaften methodische Zugänge zum Erreichen von Inhaltsvalidität aufgezeigt, welche die Grundlage für das in diesem Beitrag gewählte methodische Vorgehen bilden.

Im Zuge der Test- und Fragebogenentwicklung bildet die Überprüfung der Güte eines Instruments einen zentralen Bestandteil des Forschungsprozesses. Ursprünglich wurde Validität als ein einem Fragebogen inhärentes Merkmal begriffen. Inzwischen wird das Ausmaß des Gütekriteriums hingegen anhand der Interpretation und weiteren Verwendung von *scores*, die aus der Anwendung eines Fragebogens resultieren, bewertet (Messick, 1995, S. 742). Damit wird die Validierung eines Fragebogens zu einem fortdauernden und argumentativen Prozess, für den es je nach Anwendungskontext geeignete Evidenz anzuführen gilt.

Trotz dieses zunehmend integrativen Validitätsbegriffs kann es zur Identifikation geeigneter Evidenz nützlich sein, folgende Validitätsfacetten voneinander zu unterscheiden: Kriteriumsvalidität, Konstruktvalidität sowie schließlich Inhaltsvalidität, die Gegenstand des vorliegenden Beitrags ist (Moosbrugger & Kelava, 2020). Anhand der Inhaltsvalidität wird untersucht, inwiefern „die Items eines Tests/Fragebogens eine repräsentative Stichprobe an Verhaltens- und Erlebensweisen aus jenem Itemuniversum (d. h. allen merkmalsrelevanten Verhaltens- und Erlebensweisen) darstellen, mit dem das interessierende Merkmal vollständig erfasst werden könnte“ (Moosbrugger & Kelava, 2020, S. 32). Aus dieser Definition lassen sich mit der Repräsentanz und inhaltlichen Relevanz von Items zwei Kriterien zur Überprüfung von Inhaltsvalidität formulieren (Klauer, 1984; Messick, 1995). Enthält ein Fragebogen beispielsweise zu wenige Items, um ein komplexes Konstrukt zu erfassen, resultiert dies in „*construct underrepresentation*“ (Messick, 1995, S. 742). Andererseits kann ein Fragebogen zwar ausreichend Items enthalten, von denen einige jedoch nicht inhaltlich relevant für das zu erfassende Konstrukt sind und im ungünstigsten Fall sogar ein anderes Konstrukt erfassen („*construct-irrelevant variance*“; Messick, 1995, S. 742). Somit lässt sich Inhaltsvalidität sowohl auf Itemebene als auch holistisch auf Fragebogenebene betrachten.

Trotz der aufgezeigten Relevanz findet eine explizite Untersuchung von Inhaltsvalidität in fremdsprachendidaktischen Forschungsarbeiten bislang selten statt. Grundlage für die Überprüfung von Inhaltsvalidität bildet die Formulierung einer Konstruktspezifikation, anhand derer alle konstruierten Items unter Verweis

auf Bezugstheorien des Konstrukts legitimiert werden können (Döring & Bortz, 2016). Davon ausgehend lassen sich Annahmen zur systematischen Konstruktion eines Itempools innerhalb eines nomologischen Netzes generieren, um ein ausreichendes Maß an Repräsentanz sicherzustellen (Klauer, 1984). Die inhaltliche Relevanz des Itempools lässt sich weitaus weniger systematisch prüfen, da die Inhaltsvalidierung „theoretisch-argumentativ und gestützt durch Urteile von Fachexperten“ (Döring & Bortz, 2016, S. 446) erfolgt, statt sie anhand eines numerischen Kennwertes zu quantifizieren. Wie eine solche Expert:innenbeurteilung methodisch umgesetzt wird, soll nachfolgend anhand ausgewählter Studien illustriert werden. Aufgrund der seltenen Diskussion von Inhaltsvalidität in fremdsprachendidaktischer Forschung wird dazu auch auf Studien anderer Fachdidaktiken sowie der empirischen Bildungsforschung zurückgegriffen.

Eine in der fachdidaktischen Forschung relativ etablierte Methode zur Erfassung von Expert:innenurteilen bilden Delphi-Befragungen. Dabei beantworten Expert:innen schriftlich über mehrere Runden hinweg Items, deren Ergebnisse im Anschluss an jede Runde anonymisiert und kumuliert präsentiert werden. Dadurch wird eine schrittweise, moderierte Konsensbildung unter den beteiligten Expert:innen möglich. Weber et al. (2020) führten beispielsweise eine drei Runden umfassende und über ein Jahr andauernde Delphi-Befragung zur Konstruktion von Leitlinien für die Konzeption einer Lehrkräftefortbildung zum Thema Quantenphysik durch. Da es wünschenswert ist, dass zumindest ein Teil der Expert:innen über mehrere Runden an Delphi-Befragungen teilnehmen, verdeutlicht diese Studie bereits, welche hohen Ansprüche die Methode an die zeitlichen Ressourcen ihrer Teilnehmer:innen stellt.

Im Rahmen des interdisziplinären Projekts FALKO beschreibt Kirchhoff (2016, 2017) die Entwicklung eines Testinstruments zur Erfassung fachspezifischen professionellen Wissens von Englischlehrkräften. Zur Prüfung der Inhalts- und Augenscheinvalidität wurden 15 Expert:inneninterviews – darunter elf Englischlehrkräfte unterschiedlicher Schulformen der Sekundarstufe und vier Fachdidaktiker:innen – im Modus des lauten Denkens durchgeführt. Zu den Inhalten oder der Auswertung dieser Interviews werden in den gesichteten Publikationen keine näheren Angaben gemacht. Zudem wurden in der Vorstudie 125 Lehramtsstudierende mit dem Fach Englisch sowie Englischlehrkräfte schriftlich zur Relevanz der Testitems für die Unterrichtspraxis befragt („Ich halte diesen Inhalt für die Ausbildung von Lehrkräften für nicht relevant/kaum relevant/ziemlich relevant/sehr relevant“; Kirchhoff, 2016, S. 88).

Schließlich werden in einer Vielzahl an Forschungsarbeiten (z. B. Hofrichter et al., in diesem Band; Jenßen et al., 2015; Senkbeil et al., 2013; Wibowo & Heemsoth, 2019) Varianten schriftlicher Expert:innenbefragungen zur Prüfung der Inhaltsvalidität eingesetzt. Dabei werden anhand zuvor explizit formulierter Selektionskriterien identifizierte Expert:innen mittels Ratingskalen zur Beurteilung von Items aufgefordert. Von Interesse sind dabei beispielsweise die Relevanz und Repräsentanz der Items für das zu messende Konstrukt oder die Zugehörigkeit der Items zu unterschiedlichen Wissensfacetten in Professionstests. Vereinzelt werden

die Ratingskalen um Freitextantworten ergänzt, um schriftliche Kommentare der Expert:innen zu ermöglichen. Die Stichprobengröße variiert in den vier gesichteten Publikationen zwischen sechs und 55 Teilnehmer:innen. In gewisser Hinsicht stellt die schriftliche Expert:innenbefragung damit eine Mischform der beiden zuvor beschriebenen Zugänge dar, indem sie durch ihren quantitativen Schwerpunkt einerseits die wenig zeitintensive Befragung möglichst vieler Expert:innen erlaubt und andererseits Begründungen der Expert:innen zu ihren Urteilen einfordert.

3 Forschungsfragen

Dieser Beitrag ist Teil eines Promotionsprojekts, in dessen Rahmen Schüler:innen- und Lehrkräfteskalen zur Erfassung des Potenzials zu kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 entwickelt, erprobt und validiert werden. Damit verfolgt der Beitrag eine der einleitend formulierten Forschungsaufgaben in Bezug auf fremdsprachliche Unterrichtsqualität (Lindl et al., in diesem Band). Dazu wurde in einem ersten Schritt der Forschungsstand zu kognitiver Aktivierung im Fremdsprachenunterricht im deutschsprachigen Raum systematisch aufgearbeitet (Guttke, 2023). Auf Grundlage dieser Ergebnisse und Befunde der *Instructed Second Language Acquisition* (Loewen, 2020) wurde eine vorläufige Konstruktdefinition formuliert (s. Abschnitt 2.1). Diese diente der Konstruktion von Items, die daraufhin in mehreren Teilstudien validiert wurden. Neben der Durchführung kognitiver Interviews zur Prüfung der Iteminterpretation und einer Pilotierung zur deskriptiv-statistischen Itemanalyse wurde die nachfolgend vorgestellte Online-Expert:innenbefragung zur Inhaltsvalidierung der Items durchgeführt. Mithilfe der Daten dieser Befragung sollen die folgenden Forschungsfragen beantwortet werden:

1. In welchem Ausmaß bilden die theoriebasiert konstruierten Fragebogenitems zur Erfassung kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 laut Expert:innenurteil relevante Aspekte hinsichtlich des Konstrukts ab?
2. Welche Itemmerkmale bewerten die Expert:innen positiv/negativ hinsichtlich der Itemkonstruktion (z. B. Itemlänge oder Kongruenz von Schüler:innen- und Lehrkräfteversion eines Items)?

Das weiterführende Ziel der Beantwortung von Forschungsfrage 1 soll sein, ambige oder unpassende Items zu identifizieren, die laut Expert:innenmeinung entweder keinen Aspekt kognitiver Aktivierung abbilden oder möglicherweise solche Aspekte erfassen, die einer anderen Dimension von Unterrichtsqualität zugehörig sind. Gleichzeitig können damit Leerstellen aufgezeigt und neue Items konstruiert werden, um Aspekte kognitiver Aktivierung zu erfassen, die im Itempool bislang nicht berücksichtigt wurden. Forschungsfrage 2 bietet das Potenzial, die Item-

konstruktion durch die Expertise der Teilnehmer:innen an der Online-Befragung evaluieren zu lassen und daraus Hinweise für die Revision der Items abzuleiten.

4 Methodisches Vorgehen

4.1 Rekrutierung von Expert:innen und Stichprobe

Die Eignung zur Inhaltsvalidierung des Itempools setzt seitens der Teilnehmer:innen Expertise hinsichtlich kognitiver Aktivierung voraus. Diese Expertise kann, so wurde im Rahmen der Studie angenommen, entweder aus der wissenschaftlichen Auseinandersetzung mit dem Forschungsgegenstand oder der Reflexion unterrichtspraktischer Erfahrung erwachsen. Folglich ergaben sich drei Personengruppen für die Teilnehmer:innenakquise: Hochschullehrer:innen, Grundschullehrkräfte und Fachleitungen der Zentren für schulpraktische Lehrerausbildung (ZfsL) in Nordrhein-Westfalen (NRW). Zur Berücksichtigung bei der Akquise sollten Hochschullehrer:innen forschend in der Fremdsprachendidaktik oder der empirischen Bildungsforschung tätig sein und in den zurückliegenden fünf Jahren mindestens eine Publikation zu Unterrichtsqualität oder kognitiver Aktivierung veröffentlicht haben. Bei den Grundschullehrkräften wurde ein Abschluss des Vorbereitungsdienstes vorausgesetzt. An den ZfsL wurden ausschließlich Fachleitungen für das Fach Englisch in der Grundschule kontaktiert.

Die Akquise der Teilnehmer:innen für die Online-Expert:innenbefragung erfolgte per E-Mail. In einer ersten E-Mail wurde die Teilnahmebereitschaft der gemäß den Auswahlkriterien identifizierten Expert:innen erfragt. Diese Einladung enthielt nähere Informationen zum Forschungsprojekt sowie zum Umfang und Ablauf der Online-Befragung. Alle Expert:innen, die dieser Einladung folgten, erhielten mit einer zweiten E-Mail einen personalisierten Zugangslink zur Online-Befragung und eine schriftliche Konstruktdefinition kognitiver Aktivierung im Englischunterricht der Primarstufe, um eine gemeinsame Ausgangsbasis für das Rating der Items unter den Expert:innen zu gewährleisten.

Letztendlich nahmen 28 Expert:innen an der Online-Befragung teil, von denen 24 die Online-Befragung vollständig absolvierten. Die vier unvollständigen Datensätze wurden von der weiteren Datenauswertung ausgeschlossen. Die Expert:innen setzen sich aus zwölf Hochschullehrer:innen und zwölf Grundschullehrkräften zusammen, von denen acht zusätzlich als Fachleitungen an ZfsL in NRW tätig sind. Durchschnittlich wiesen die Hochschullehrer:innen eine Berufserfahrung von elf Jahren ($M = 10,92$; $SD = 3,57$) und die Grundschullehrkräfte von zehn Jahren ($M = 10,17$; $SD = 5,89$) auf. Vier der Hochschullehrer:innen ordneten sich der empirischen Bildungsforschung, acht der Fremdsprachendidaktik zu. Hinsichtlich ihrer Qualifikation gaben sieben Grundschullehrkräfte Englisch als Fach in Lehramtsstudium und Vorbereitungsdienst an, wohingegen bei zwei Grundschullehrkräften das Fach Englisch Teil des Lehramtsstudiums, nicht jedoch des Vorbereitungsdienstes war. Drei Grundschullehrkräfte studierten das

Fach Englisch nicht, sondern erwarben ihre Lehrqualifikation durch eine Nachqualifizierung.

4.2 Erhebungsinstrument

Die Online-Befragung fand im Zeitraum vom 04.05.–19.06.2022 statt und wurde über LimeSurvey durchgeführt. Als Erhebungsinstrument diente der gesamte theoriebasiert konstruierte Itempool, der zum damaligen Zeitpunkt 41 Itempaare verteilt auf vier Skalen (*language input*, *information processing*, *language output*, *corrective feedback*; vgl. Tab. 1) umfasste. Jedes Itempaar bestand aus einem Schüler:innen- und Lehrkräftemitem, welche dieselben Inhalte aus unterschiedlichen Perspektiven beschrieben. Die Lehrkräftemitems wurden in der Ich-Perspektive (Selbstbewertung) formuliert, wohingegen die Schüler:innenitems größtenteils Aussagen über die Englischlehrkraft enthielten (Fremdbewertung: „Mein:e Englischlehrer:in ...“).

Tabelle 1: Beispielimitem zu den vier in der Online-Expert:innenbefragung bewerteten Skalen

Itemcode	Lehrkräftemitem	Schüler:innenitem
<i>Language input</i>		
<i>modifikation</i>	Ich verändere meinen Sprachgebrauch im Englischen je nach Lernstand der Schüler:innen, mit denen ich spreche.	Wenn mein:e Englischlehrer:in mit mir auf Englisch spricht, verstehe ich das meiste, was sie/er sagt.
<i>Information processing</i>		
<i>erstsprache</i>	Ich ermögliche den Schüler:innen, die englische Sprache mit ihrer Erstsprache zu vergleichen.	Im Englischunterricht vergleiche ich Englisch mit anderen Sprachen. Ich denke darüber nach, was bei den beiden Sprachen gleich und was verschieden ist.
<i>Language output</i>		
<i>nonverbal- aeussern</i>	Wenn Schüler:innen ein englisches Wort nicht kennen, um einen Satz zu formulieren, dürfen sie sich non-verbal äußern (z. B. Gestik, Mimik, Symbolkarten).	Wenn ich ein Wort auf Englisch nicht weiß, darf ich Hilfsmittel benutzen (z. B. meine Hände, Karten mit Bildern).
<i>Corrective feedback</i>		
<i>inhaltsfehler</i>	Ich korrigiere inhaltliche Fehler der Schüler:innen (z. B. „My mother is 9 years old.“).	Mein:e Englischlehrer:in berichtigt mich, wenn ich etwas sage, das nicht stimmt (z. B. „My mother is 9 years old.“).

Nach Abfrage einiger demografischer Merkmale wurde den Expert:innen nacheinander jedes Itempaar präsentiert (vgl. Abb. 2). Die Abfolge der Items innerhalb der Skalen wurde randomisiert. Die Repräsentanz und Relevanz eines Items konnte entlang einer vierstufigen Likert-Skala eingeschätzt werden. Zusätzlich standen den Expert:innen drei Freitextfelder für die Begründung ihrer Einschätzung, eines alternativen Formulierungsvorschlags sowie weitere Anmerkungen zur Verfügung.

**Bilden die Items Ihrer Meinung nach einen Aspekt
kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 ab?**

	[L] Ich wiederhole Wörter und Redewendungen, die zum Lernziel der Unterrichtsstunde gehören.	[S] Mein:e Englischlehrer:in sagt englische Wörter und Sätze, die ich lernen soll, extra oft.
Stimme voll zu	<input type="checkbox"/>	<input type="checkbox"/>
Stimme eher zu	<input type="checkbox"/>	<input type="checkbox"/>
Stimme weniger zu	<input type="checkbox"/>	<input type="checkbox"/>
Stimme nicht zu	<input type="checkbox"/>	<input type="checkbox"/>

Optional können Sie Ihre Einschätzung nachfolgend ausführen.

Bitte begründen Sie Ihre Einschätzung kurz (insbesondere dann, wenn Sie weniger oder nicht zustimmen):

Falls Sie mit der Itemformulierung unzufrieden sind, wie würden Sie das Item stattdessen formulieren:

Gegebenenfalls weitere Anmerkungen:

Abbildung 2: Illustration des Frageformats aus der Online-Expert:innenbefragung anhand eines Beispielitems der Skala *language input*

4.3 Datenauswertung

Aufgrund der Verwendung eines geschlossenen und eines offenen Antwortformats resultierten aus der Online-Befragung sowohl quantitative als auch qualitative Daten, die auf unterschiedliche Weise ausgewertet wurden. Die Daten der vierstufigen Likert-Skalen zur Einschätzung der Items wurden zunächst dichotomisiert. Dazu wurden jeweils zwei Ausprägungen der Likert-Skala (*stimme nicht zu/stimme weniger zu* sowie *stimme voll zu/stimme eher zu*) zu einer neuen Ausprägung (*Ablehnung* sowie *Zustimmung*) gebündelt. Damit ließ sich pro Item der Anteil an Zustimmung bzw. Ablehnung an der Gesamtzahl der Expert:innenurteile berechnen. Da die vier Ausprägungen der in der Online-Befragung verwendeten Likert-Skala nicht äquidistant sind, werden die Daten als ordinalskaliert aufgefasst (Brück & Toth, 2022; Rosebrock et al., 2019).

Die Freitextantworten der Expert:innen wurden hingegen qualitativ in Form einer inhaltlichen Kategorisierung ausgewertet. Dazu wurden alle Freitextantworten itemweise von den Autor:innen gesichtet und im Rahmen eines konsensuellen Vorgehens (Kuckartz, 2022) einer von vier Kategorien zugeordnet. Die Kategorien wurden nach einer ersten Sichtung aus den Daten heraus gebildet: Positive Rückmeldungen (Begründung des Expert:innenurteils im Falle von Zustimmung), Kritik (Begründung des Expert:innenurteils im Falle von Ablehnung), Vorschläge zur Änderung der Itemformulierung sowie Vorschläge zur Ergänzung des Itempools.

5 Ergebnisse

5.1 Quantitative Daten aus dem Expert:innenrating

Zunächst werden die quantitativen Daten der Expert:innenurteile über die Repräsentanz und Relevanz der Items präsentiert, die in Abbildungen 3 und 4 visualisiert sind. Die Abbildungen zeigen die Anzahl zustimmender und ablehnender Expert:innenurteile pro Item in aufsteigender Reihenfolge. Die unterschiedliche Färbung der Balken ermöglicht es, die absoluten Urteilsanteile den jeweiligen Expert:innengruppen zuzuordnen (hell: Grundschullehrkräfte und Fachleitungen, dunkel: Hochschullehrer:innen). Zustimmung bzw. Ablehnung bedeutet in diesem Kontext, dass die Expert:innen (nicht) der Auffassung waren, das jeweilige Item bilde einen Aspekt kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 ab. Die in diesem Abschnitt präsentierten Ergebnisse erlauben folglich Rückschlüsse über die Inhaltsvalidität der Items.

Von den 41 Items des Lehrkräftefragebogens erhielten sechs Items die vollständige Zustimmung der Expert:innen und insgesamt 31 Items 75 % (= 18 Urteile) oder mehr Zustimmung. Von den 41 Schüler:innenitems erhielten fünf Items die vollständige Zustimmung der Expert:innen und insgesamt 30 Items 75 % (= 18 Urteile) oder mehr Zustimmung. Aus beiden Itempools erhielt kein einziges Item weniger als 50 % (= 12 Urteile) Zustimmung.

Aufgrund des Anteils an Zustimmung durch die Expert:innen können die Items in Ränge eingeteilt werden. So bilden beispielsweise alle Items mit 24 zustimmenden Urteilen Rang eins, alle Items mit 23 zustimmenden Urteilen Rang zwei usw. Beim Rangvergleich der Items zwischen Lehrkräfte- und Schüler:innenfragebogen zeigt sich, dass mit 34 Items ein Großteil der Items entweder denselben Rang belegt oder einen Rangunterschied von 1 aufweist. Bei den übrigen sieben Items fällt dieser Rangunterschied höher aus. Die fünf Items *einsprachigkeit*, *koerpersprache*, *modifikation*, *spass* und *sprachfehler* unterscheiden sich zwischen den beiden Fragebögen um zwei Ränge. Bei den beiden Items *focusonforms* und *simulation* beträgt der Rangunterschied sogar 4.

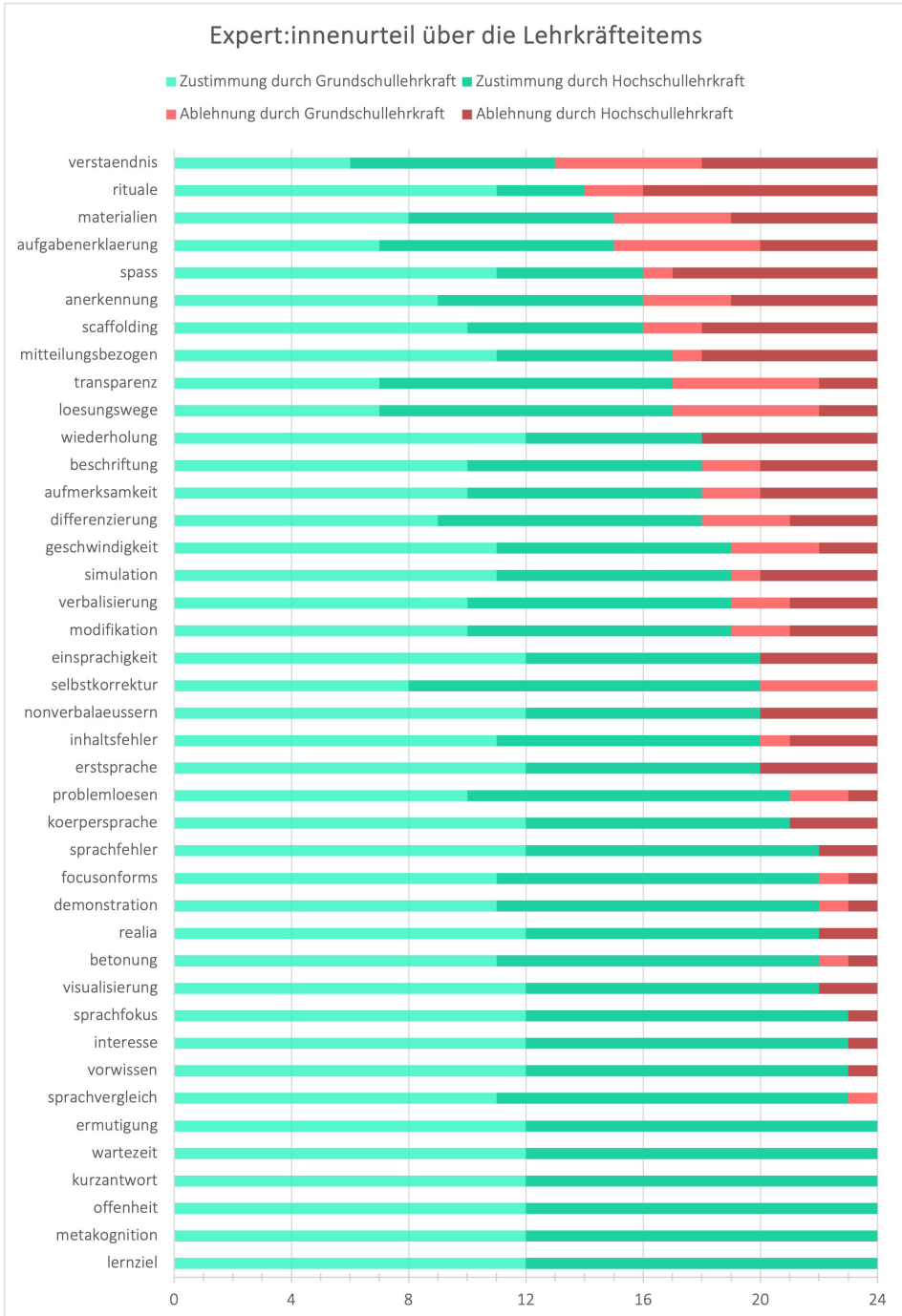


Abbildung 3: Visualisierung der Expert:innenurteile hinsichtlich der Lehrkräfteitems zu der Frage: „Bilden die Items Ihrer Meinung nach einen Aspekt kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 ab?“

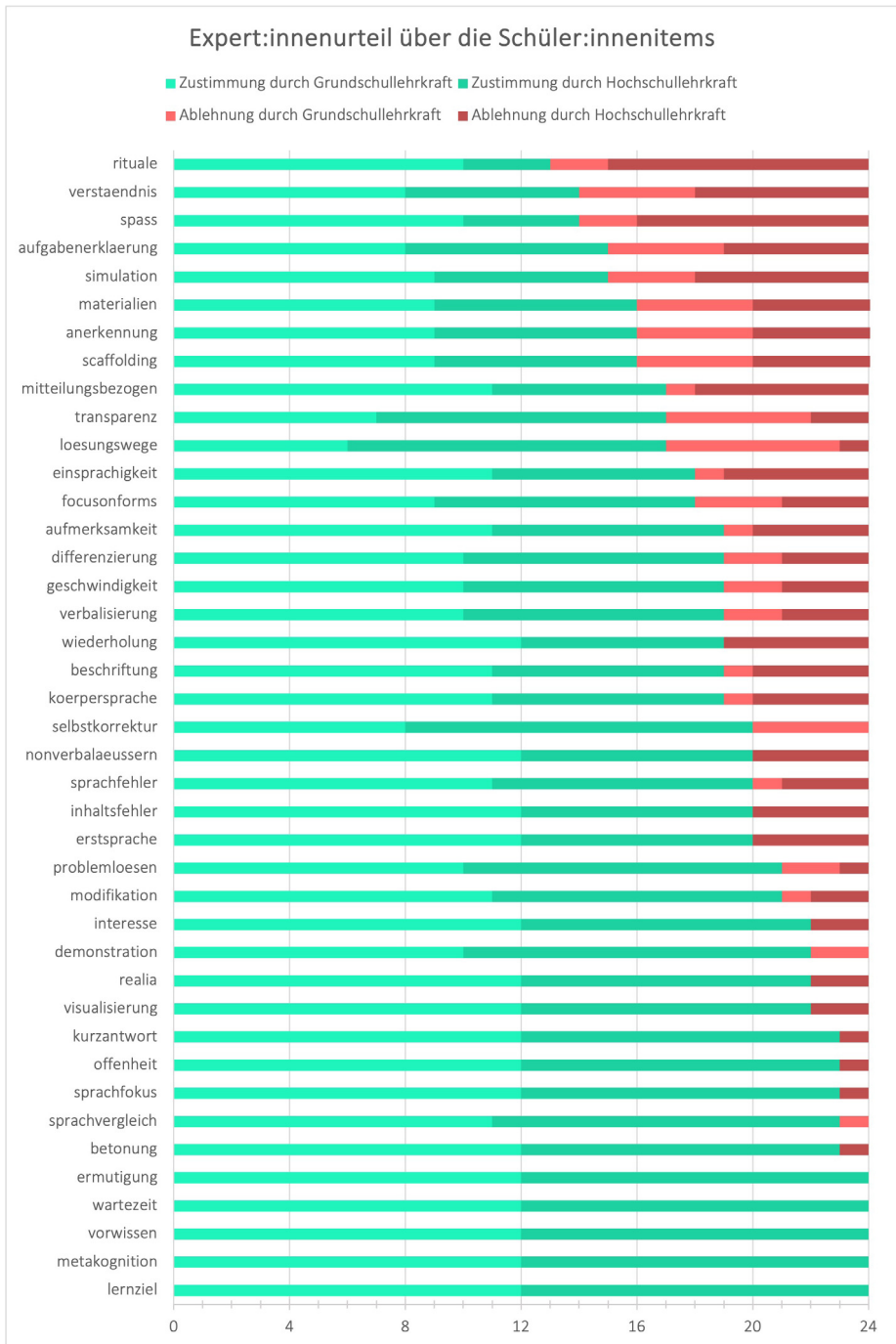


Abbildung 4: Visualisierung der Expert:innenurteile hinsichtlich der Schüler:innenitems zu der Frage: „Bilden die Items Ihrer Meinung nach einen Aspekt kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4 ab?“

5.2 Qualitative Daten aus den Expert:innenkommentaren

Nachfolgend werden die Expert:innenkommentare aus den Freitextantworten entlang der vier in Abschnitt 4.3 beschriebenen Kategorien zusammenfassend wiedergegeben. Kategorie 1 (positive Rückmeldungen) beinhaltet alle Kommentare, in denen ein zustimmendes Expert:innenurteil begründet wurde. Die Begründung des durch das Item intendierten Potenzials zu kognitiver Aktivierung erfolgte unter Bezug auf Zweitspracherwerbstheorien, pädagogisch-psychologische Lerntheorien oder die eigene Unterrichtserfahrung.

In Kategorie 2 (Kritik) wurden alle Kommentare zusammengefasst, mit denen die Expert:innen ihr ablehnendes Urteil begründeten. Diese Begründungen thematisierten folgende Aspekte:

- Gefahr konstruktirrelevanter Varianz: Vereinzelt wurde darauf hingewiesen, die Ausprägung eines Items hänge nicht ausschließlich mit dem Grad kognitiver Aktivierung zusammen. Ein Beispiel dafür ist das Item *beschriftung* ([L] „Ich beschrifte Gegenstände im Klassenraum mit englischen Wörtern oder Redewendungen.“, [S] „In meinem Klassenraum gibt es Sachen, die mit englischen Wörtern oder Sätzen beschriftet sind.“). Grundschullehrkräfte wiesen darauf hin, die Gestaltung des Klassenraums obliege in manchen Fällen der Klassenlehrkraft oder der Klassenraum sei schlicht so voll, dass eine Beschriftung einzelner Gegenstände nicht möglich sei.
- Iteminhalt differenziert nicht stark genug zwischen dem Potenzial zu kognitiver Aktivierung und kognitiver Aktivität: Bei einigen Items wurde darauf hingewiesen, es handele sich bei dem Iteminhalt lediglich um ein Oberflächenmerkmal von Unterricht, das Gelegenheitsstrukturen für das Potenzial zu kognitiver Aktivierung beschreibe. Hinsichtlich des Ausmaßes an kognitiver Aktivität sei es notwendig, den Iteminhalt in Bezug zu einer konkreten Unterrichtsaktivität zu setzen. Beispielfhaft sei hier auf das Item *materialien* ([L] „Ich stattete den Klassenraum mit englischsprachigen Materialien (z. B. Bilderbücher, Spiele, CDs) für die Schüler:innen aus.“, [S] „In meinem Klassenraum gibt es Bücher, Spiele oder CDs auf Englisch, die ich benutzen darf.“) verwiesen.
- Iteminhalt weist einen Mangel an unterrichtspraktischer Relevanz auf: Obwohl einzelne Teilnehmer:innen eine Passung zwischen Iteminhalt und Konstruktdefinition feststellten, beurteilten sie das Item mit Ablehnung, da der Iteminhalt selten bis gar nicht in ihrer Unterrichtspraxis vorkomme. Diese Form der Begründung wurde beispielsweise beim Item *loesungswege* ([L] „Wenn Schüler:innen unterschiedliche Formulierungen vorschlagen, um dasselbe zu sagen, diskutiere ich mit ihnen, welcher Vorschlag am angemessensten ist.“, [S] „Wenn es im Englischunterricht mehrere Antworten auf eine Frage gibt, überlegen wir als Klasse, welche Antwort die Beste ist.“) geäußert.
- Keine Passung zwischen Item und Unterrichtsqualitätsdimension kognitiver Aktivierung: Bei der Beurteilung des Items *spass* ([L] „Ich Sorge dafür, dass sich die Schüler:innen wohlfühlen und Spaß beim Lernen haben.“, [S] „Ich habe Spaß am Englischunterricht und spreche gerne auf Englisch.“) wurde

beispielsweise mehrfach auf eine höhere Passung des Items mit der Unterrichtsqualitätsdimension sozio-emotionale Unterstützung hingewiesen. Vorschläge für die Zuordnung eines Items zu einer anderen Unterrichtsqualitätsdimension waren nicht Teil der Online-Befragung, sondern wurden selbstständig von den Expert:innen angeführt, basierten also auf ihrem jeweils individuellen Vorwissen.

In Kategorie 3 fallen all jene Kommentare, in denen die Expert:innen Aspekte der Itemformulierung diskutierten oder Vorschläge zur Überarbeitung der Itemformulierung verfassten. In dieser Kategorie lassen sich vier wiederkehrende Themen identifizieren. Erstens wurde dazu geraten, die Formulierungen in den Lehrkräfte- und Schüler:innenitems stärker aneinander anzupassen, um die Konsistenz der beiden jeweils analog konstruierten Items zu erhöhen. So wurde beispielsweise kritisiert, dass die Umschreibung des Wortes „Anschauungsmaterialien“ aus dem Lehrkräfteitem *realia* mit dem Ausdruck „Sachen, die ich von zuhause kenne“ im Item der Schüler:innen inhaltlich zu ungenau sei.

Zweitens wiesen die Expert:innen darauf hin, einige der Items seien sprachlich zu komplex. Um diesem Problem zu begegnen, wurde zunächst dazu geraten, die Schüler:innenitems so kurz und simpel wie möglich zu formulieren. Außerdem schlugen die Expert:innen vor, im Grundschulalter seltener genutzte Wörter durch bekanntere Wörter zu ersetzen (z. B. „zeigen“ statt „verdeutlichen“ oder „erklären“). Schließlich identifizierten die Expert:innen in wenigen Items mehrdeutige Wörter, die ambige Iteminhalte verursachten. Zur Veranschaulichung sei das Lehrkräfteitem *wiederholung* („Ich wiederhole Wörter und Redewendungen, die zum Lernziel der Unterrichtsstunde gehören.“) angeführt. In diesem Item verstanden einige Expert:innen das Verb „wiederholen“ im intendierten Sinne der wiederholten Aussprache. In einer Englischstunde zum Unterrichtsvorhaben *Me and my family* mit dem Lernziel Wortschatzarbeit gelte es beispielsweise als lernförderlich, wenn die Englischlehrkraft Wörter und Redewendungen aus dem Wortfeld *relatives* im Unterrichtsgespräch wiederholt verwendet. Andere Expert:innen hingegen begriffen das Verb „wiederholen“ im Sinne des wiederholten Übens eines *language chunks*. Sie assoziierten damit den gehäuftten Einsatz von Übungsphasen (z. B. *gap filling exercises*, Dialoge), in denen die Schüler:innen den Einsatz neu eingeführter *chunks* erproben.

Drittens kommentierten die Expert:innen den Grad der Genauigkeit der Items hinsichtlich der Sicherung des Itemverständnisses. Sie rieten einerseits dazu, weitere Items um konkrete Beispiele zu ergänzen, da die Iteminhalte sonst zu abstrakt für Schüler:innen der Jahrgangsstufe 4 seien. Gleichzeitig äußerten einige der Expert:innen Bedenken darüber, Beispiele könnten aufgrund der Vielfalt möglicher Implementationen von Iteminhalten im Englischunterricht einschränkend wirken. Beispielsweise wurde vorgeschlagen, das Schüler:innenitem *rituale* („Im Englischunterricht gibt es Sachen, die mein:e Englischlehrer:in immer gleich macht.“) durch die Beschreibung einer konkreten Unterrichtssituation zu konkretisieren.

Viertens benannten die Expert:innen Probleme bezüglich der Itemreferenz. Im Falle der Lehrkräfteteims sei vereinzelt unklar, ob sich die Iteminhalte auf die Interaktion zwischen Lehrkraft und Schüler:in oder Lehrkraft und Lerngruppe bezögen. Aus den Schüler:innenitems ginge wiederum vereinzelt nicht hervor, ob die Iteminhalte mit Blick auf individuelle Schüler:innen oder die gesamte Lerngruppe zu bewerten seien. Zuletzt merkten die Expert:innen an, einige Items könnten hinsichtlich ihrer Formulierung geschärft werden, um den Unterschied zwischen lehrkräfteseitigem Unterrichtsangebot und schüler:innenseitiger Unterrichtsnutzung stärker herauszuarbeiten.

Kategorie 4 beinhaltet die Vorschläge der Expert:innen zur Erweiterung des Itempools um weitere Aspekte des Englischunterrichts in der Primarstufe, die sich wie folgt zusammenfassen lassen:

- Differenzierung der Items zu *corrective feedback* hinsichtlich konkreter Feedback-Strategien,
- Einbezug von für die Primarstufe typischen Unterrichtsgegenständen und Aufgabenstellungen,
- Einführung des Schriftbildes,
- Einsatz digitaler Medien,
- formatives Assessment,
- Formen der Ausspracheschulung (z.B. durch Chorsprechen, Lieder oder Sprachspiele),
- Hinzufügen weiterer Items zur Interaktion der Schüler:innen in der Zielsprache, um diesen Teilaspekt der Konstruktdefinition nicht zu vernachlässigen.

Schließlich gab es vereinzelt Expert:innen, die der Auffassung waren, der Itempool beinhalte zu viele Items zu allgemeinen Strategien des Fremdsprachenunterrichts, die nicht unmittelbar relevant für kognitive Aktivierung seien. Sie forderten die Formulierung zusätzlicher Items zum Problemlösen sowie zum Einsatz herausfordernder Aufgaben im Englischunterricht der Primarstufe. Als Beispiele führten sie offene Unterrichtssettings und Aufgabenformate wie das *task-based language teaching* an sowie Gelegenheiten zum selbstständigen Lernen.

6 Ergebnisdiskussion

Nachfolgend werden die präsentierten Ergebnisse der Online-Expert:innenbefragung diskutiert und auf dieser Grundlage Implikationen für die Weiterentwicklung des Fragebogens zur Erfassung kognitiver Aktivierung im Englischunterricht der Primarstufe herausgearbeitet. Eine eindeutige Interpretation der Expert:innenurteile zu den Schüler:innen- und Lehrkräfteteims ist insofern herausfordernd, als dass für die Wertung des Vorliegens von Inhaltsvalidität – anders als beispielsweise für Reliabilitätskoeffizienten – keine Normwerte existieren. In Studien, die ebenfalls einen quantitativen Zugang zur Annäherung an Inhaltsvalidität wählen (s. Abschnitt 2.3), formulieren die Autor:innen eigene Kriterien für die

Itemselektion, die vor dem Hintergrund des jeweiligen Forschungsprojekts plausibel erscheinen. Diesem Vorgehen folgend wurden für die vorliegende Teilstudie vor Durchführung der Online-Befragung folgende Kriterien formuliert:

- Kriterium K1: Items, die von mindestens der Hälfte der Expert:innen (= 12 Urteile) ein ablehnendes Urteil erhalten, gelten nicht als ausreichend inhaltsvalide und werden eliminiert.
- Kriterium K2: Items, die von maximal 15 % der Expert:innen (= 4 Urteile) ein ablehnendes Urteil erhalten, gelten als ausreichend inhaltsvalide und werden ohne inhaltliche Änderungen beibehalten.
- Kriterium K3: Items, die von mindestens 15 % bis maximal 50 % der Expert:innen ein ablehnendes Urteil erhalten, werden nach einer erneuten Prüfung entweder eliminiert oder auf Grundlage der Expert:innenkommentare überarbeitet.

Die Revisionsgrenze von 15 % wurde im Vergleich zu Studien mit einem ähnlichen methodischen Vorgehen aufgrund der in Abschnitt 2.1 geschilderten Unschärfe des Konstrukts kognitiver Aktivierung bewusst strenger gewählt. Anhand dieser Kriterien lässt sich vorsichtig schlussfolgern, dass die theoriebasierte Itemkonstruktion in einem inhaltsvaliden Itempool für beide Fragebögen resultierte. Keines der Schüler:innen- oder Lehrkräfteitems wurde durch K1 eliminiert. 23 Lehrkräfteitems und 21 Schüler:innenitems wurden gemäß K2 inhaltlich unverändert beibehalten. Dieser Anteil entspricht in beiden Fällen mehr als der Hälfte des Itempools. Damit unterlagen im Anschluss an die Online-Expert:innenbefragung 18 Lehrkräfteitems und 20 Schüler:innenitems einer erneuten Prüfung und Überarbeitung.

Die aus den Expert:innenurteilen hervorgegangenen Rangunterschiede zwischen den jeweils analogen Schüler:innen- und Lehrkräfteitems könnten ein Indiz für perspektivenspezifische Aspekte kognitiver Aktivierung liefern (Fauth et al., 2020; Lindl et al., in diesem Band). Diese Rangunterschiede können so interpretiert werden, dass die Expert:innen das jeweilige Item hinsichtlich kognitiver Aktivierung für Lehrkräfte relevanter halten als für Schüler:innen oder umgekehrt. So merkten einige Expert:innen zu Lehrkräfteitems mit einem starken Fokus auf das Lehrkräfteverhalten oder didaktische Aspekte an, der Iteminhalt sei konstruktirrelevant für den Schüler:innenfragebogen, da Schüler:innen nicht über genügend Informationen und Expertise verfügten, um das Item zu beantworten. Auskunft über die Gültigkeit dieser Vermutung werden Ergebnisse weiterer Validierungs- und Pilotierungsstudien liefern.

Zudem wurde bei der Itemkonstruktion darauf geachtet, Unterrichtsmerkmale zu fokussieren, die *Potenziale* zu kognitiver Aktivierung repräsentieren, und diese Unterrichtsmerkmale – sofern möglich – anhand konkreter Szenarien zu beschreiben. Dies bedeutet, dass die Lehrkräfteitems zu kognitiver Aktivierung mehrheitlich selbstreferent und -evaluativ sind, wohingegen die Schüler:innenitems mehrheitlich fremdreferent und -evaluativ sind. Vergleichsweise hohe Rangunterschiede treten primär bei solchen Items auf, die davon eine Ausnahme

bilden. Ein Beispiel dafür ist das Schüler:innenitem *focusonforms* („Wenn ich etwas auf Englisch falsch sage, erklärt mir mein:e Englischlehrer:in den Fehler. Ich denke dann ganz genau über den Fehler nach.“), das aufgrund der Expert:innenurteile vier Ränge über dem zugehörigen Lehrkräfteam („Wenn Schüler:innen im Gespräch einen sprachlichen Fehler machen, nutze ich das als Anlass, ihnen diesen Fehler zeitnah bewusst zu machen und zu erklären.“) liegt. Dieses Item fordert sowohl von Lehrkräften als auch von den Schüler:innen, ein kombiniertes Urteil über beide Personengruppen abzugeben.

Ferner könnten die Rangunterschiede in der Expertise der befragten Expert:innen begründet sein, die auf unterschiedlichen Wissensformen (Vogel, 2019) basiert. Die Hochschullehrkräfte verfügen allein aufgrund ihrer Disziplinzugehörigkeit über variierendes Fachwissen zu kognitiver Aktivierung, wohingegen sich die Grundschullehrkräfte in ihrem unterrichtspraktischen Erfahrungswissen unterscheiden. Zudem ist davon auszugehen, dass sich die Hochschullehrkräfte vorwiegend an wissenschaftlichem Wissen orientieren, das seitens der Grundschullehrkräfte während der Berufspraxis durch pädagogisches Alltags- und Professionswissen angereichert wurde. Betrachtet man die Expert:innenurteile separat nach Personengruppen, wird erkennbar, dass die Grundschullehrkräfte deutlich mehr Items (z. B. *wiederholung*, *nonverbalaeussern* oder *realia*) ausschließlich zustimmend beurteilten als die Hochschullehrkräfte. Umgekehrt ist dies nur bei drei Items der Fall (*sprachvergleich*, *demonstration* und *selbstkorrektur*). Eine Verzerrung aufgrund Zustimmung durch soziale Erwünschtheit ist seitens der Grundschullehrkräfte eher unwahrscheinlich, da sie im Rahmen der Online-Befragung nicht dazu aufgefordert waren, ihren eigenen Unterricht zu beurteilen. Auch wenn es nicht der Regelfall ist, scheint bei einigen Items durchaus Uneinigkeit zwischen beiden Personengruppen zu herrschen. Auch in den Expert:innenkommentaren aus den Freitextantworten dokumentieren sich personengruppenspezifische Merkmale. Die Hochschullehrkräfte – insbesondere solche, die sich der empirischen Bildungsforschung zugehörig fühlen – wählen mehrheitlich einen komparativen Ansatz und diskutieren die Items im Kontext der Unterrichtsqualitätsforschung. Sie verweisen beispielsweise auf andere Unterrichtsqualitätsdimensionen des Syntheseframeworks (Praetorius et al., 2020) oder vergleichen die neu konstruierten Items mit bereits veröffentlichten Items zur Erfassung kognitiver Aktivierung. Die Grundschullehrkräfte schilderten zur Begründung ihrer Itemurteile häufig Episoden des Englischunterrichts und leiteten daraus die Relevanz des jeweiligen Items ab. Obwohl diese unterrichtspraktischen Erfahrungen höchst subjektiv sind und nicht unmittelbar verallgemeinert werden sollten, liefern sie doch wichtige Informationen über die ökologische Validität der Items.

Die Vorschläge der Expert:innen zur Ergänzung des Itempools aus den Freitextantworten wurden einzeln geprüft. Die Prüfung erfolgte vor dem Hintergrund der Konstruktspezifikation kognitiver Aktivierung, der zur Konstruktion des Itempools gesichteten Literatur sowie dem Zweck der zu entwickelnden Fragebögen. Der Vorschlag, die Items zu *corrective feedback* zu differenzieren, wurde

beispielsweise bei der Revision des Itempools berücksichtigt. Im Entwurf des Itempools waren drei vergleichsweise allgemein formulierte Items zum Umgang der Lehrkraft mit Schüler:innenfehlern enthalten, die sich auf die Korrektur von Sach- und Sprachfehlern sowie die Anregung der Schüler:innen zur Selbstkorrektur sprachlicher Fehler bezogen. Die Expert:innen äußerten diesbezüglich mehrheitlich, die existierenden Items seien zu undifferenziert, da nicht alle Formen von *corrective feedback* gleichermaßen kognitiv aktivierend wirken. Diese Annahme unterschiedlicher Effekte verschiedener Feedback-Strategien wird durch die empirische Befundlage der Zweitspracherwerbsforschung gestützt (z. B. *immediate vs. delayed feedback, implicit vs. explicit feedback*; Li, 2010; Lyster et al., 2013). Im Zuge der Revision wurden sechs neue Items (*explicit correction, recast, request, metalinguistic cue, repetition, elicitation*; Oliver & Adams, 2021) formuliert, die empirisch identifizierte Feedback-Strategien des Fremdsprachenunterrichts abbilden und so dazu dienen, den Umgang der Lehrkraft mit Schüler:innenfehlern differenzierter zu erfassen.

Andere Ergänzungsvorschläge, wie beispielsweise der Einbezug von für die Primarstufe typischen Unterrichtsgegenständen und Aufgabenstellungen, wurden nach einer kritischen Würdigung und Prüfung verworfen. In diesem Fall war die Entscheidung darin begründet, dass die zu entwickelnden Skalen bzw. Inventare im weiteren Verlauf des Promotionsprojekts zur Untersuchung von Zusammenhängen mit weiteren Konstrukten möglichst breit einsetzbar sein sollen. Dieses Ziel wäre durch eine Ausrichtung der Items auf konkrete Unterrichtsgegenstände, die möglicherweise nur zu einem bestimmten Zeitpunkt im Schuljahr unterrichtet werden, eingeschränkt worden. Dennoch ist im Forschungsdiskurs zu Unterrichtsqualität eine Auseinandersetzung mit der Beziehung zwischen den Inhalten und der Qualität von Unterricht unter dem Begriff der Lerngegenstandsorientierung – auch bezüglich der Erfassung von Unterrichtsqualität – erkennbar: „In diesem Zusammenhang erscheint es zudem lohnend zu diskutieren, welche Dimensionen von Unterrichtsqualität für *alle* zentralen Ziele, Inhalte und Methoden hohe Relevanz haben und welche lediglich oder vornehmlich für *bestimmte* Ziel-Inhalts-Methoden-Konfigurationen bedeutsam sind“ (Praetorius et al., 2020, S. 429; Hervorhebung im Original). Es erscheint folglich plausibel, die pilotierten Fragebogenitems in zukünftigen Forschungsprojekten entlang von Lerngegenständen des Englischunterrichts in der Primarstufe zu konkretisieren und auf ihre Gültigkeit zu prüfen.

Abschließend werden Einschränkungen der vorliegenden Studie und Handlungsalternativen dargelegt. Für einen quantitativen Zugang, der zur Beurteilung der Inhaltsvalidität gewählt wurde, ist die Anzahl der Expert:innen relativ gering. Zudem stellen die Teilnehmer:innen aufgrund der Freiwilligkeit der Teilnahme an der Online-Befragung eine Positivauswahl dar. Eine umfangreichere Stichprobe hätte die Zuverlässigkeit des Itemratings sicherlich gesteigert; gleichzeitig ist die Anzahl der Expert:innen vergleichbar mit anderen Studien, in denen ein ähnliches methodisches Vorgehen gewählt wurde (Jenßen et al., 2015; Wibowo & Heemsoth, 2019). Des Weiteren wäre es denkbar gewesen, die Studie in einem

anderen Format wie einer Delphi-Befragung oder qualitativen Expert:inneninterviews durchzuführen. Während diese beiden Formate vermutlich eine intensivere inhaltliche Auseinandersetzung der Expert:innen mit dem Itempool ermöglicht hätten, führte die Entscheidung für eine Online-Befragung zu einer Abbildung mehrerer Perspektiven. Expert:inneninterviews in ähnlichem Umfang (z. B. Prusse-Hess & Prusse, 2018) hätten zwar auch unterschiedliche Perspektiven berücksichtigen können, wären jedoch sowohl für den Forschenden als auch die teilnehmenden Expert:innen deutlich ressourcenintensiver. Darüber hinaus bestand mit den Freitextantworten in der Online-Befragung die Gelegenheit zur Angabe von Begründungen, die von den Expert:innen in erfreulichem Ausmaß wahrgenommen wurde. Zuletzt ist anzumerken, dass die eigens formulierte Konstruktspezifikation kognitiver Aktivierung den Expert:innen ausschließlich schriftlich mitgeteilt wurde. Unklarheiten oder Missverständnisse seitens der Expert:innen, die möglicherweise zu invaliden Itembeurteilungen führen können, sind deshalb nicht auszuschließen.

7 Fazit

Auf Grundlage der Ergebnisse der Online-Expert:innenbefragung erwies sich der theoriebasiert konstruierte Itempool als solide Ausgangsbasis für die Konstruktion des Schüler:innen- und Lehrkräftefragebogens zu Erfassung kognitiver Aktivierung im Englischunterricht der Jahrgangsstufe 4. Dennoch enthielt die Expert:innenkritik konstruktive Vorschläge zur Ergänzung und Überarbeitung einzelner Items. Sie diente somit als Korrektiv für die Autor:innen, die den initialen Itempool eigens formulierten. Dieser wird im Rahmen der nächsten Teilstudie einer Überprüfung des Itemverständnisses mittels kognitiver Interviews mit Grundschullehrkräften und -schüler:innen unterzogen, um den Validierungsprozess fortzuführen (Guttke & Porsch, 2024).

Es lässt sich schlussfolgern, dass die Inhaltsvalidierung durch eine Online-Expert:innenbefragung einen produktiven, effizienten und qualitätssichernden Arbeitsschritt in der Instrumententwicklung bildet. Aus diesen Gründen ist es wünschenswert, dass die Sicherung der Inhaltsvalidität neu entwickelter Instrumente in zukünftigen Arbeiten innerhalb der Fremdsprachenforschung und darüber hinaus stärkere Berücksichtigung findet. Schließlich leistet die Studie einen Beitrag zu der gemeinsamen Aufgabe von Bildungswissenschaft und Fachdidaktik, Befunde zur Fachlichkeit von Unterrichtsqualität zu integrieren (Lindl et al., in diesem Band). So ist es für das Fach Englisch primär die Zweitspracherwerbsforschung, weniger die fachdidaktische Forschung, die Belege für potenziell kognitiv aktivierende Merkmale des Fremdsprachenunterrichts hervorgebracht hat. Es ist an der Zeit, die vielfältigen konzeptionellen Überlegungen der Fremdsprachendidaktik, welche die Unterrichtsgestaltung zweifelsohne bereichern, ebenfalls empirisch auf ihre Qualität zu prüfen. Zudem wird deutlich, dass sich die seitens der bildungswissenschaftlichen Forschung eingesetzten Instrumente zur Messung der

Unterrichtsqualitätsdimension kognitive Aktivierung lediglich eingeschränkt für fremdsprachendidaktische Untersuchungen eignen. Notwendig sind fachspezifische Anpassungen bzw. Entwicklungen von Items, die den Gegenstand angemessen zu erfassen vermögen. Die für die Fremdsprachenvermittlung bestehende Besonderheit, dass im Unterricht die Sprache Lerngegenstand und Medium zugleich darstellt, sollte dazu angemessen berücksichtigt werden. Die vorliegende Studie beschreibt einen Schritt zur Entwicklung solcher Items und lässt die Empfehlung zu, dass eine Expert:innenbefragung für die Entwicklung neuer Items notwendig ist, um inhaltsvalide Messinstrumente zu erstellen.

Literatur

- Begrich, L., Praetorius, A.-K., Decristan, J., Fauth, B., Göllner, R., Herrmann, C., Klein-knecht M., Taut, S. & Kunter, M. (2023). Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft. *Unterrichtswissenschaft*, 51, 63–97. <https://doi.org/10.1007/s42010-023-00163-4>
- Brück, N. & Toth, C. (2022). *Studienbuch Operationalisierungen*. Berlin: Springer. <https://doi.org/10.1007/978-3-658-30239-9>
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive. Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3), 127–137. <https://doi.org/10.1024/1010-0652/a000129>
- Fauth, B., Göllner, R., Lenske, G., Praetorius, A.-K. & Wagner, W. (2020). Who sees what? Conceptual considerations on the measurement of teaching quality from different perspectives. *Zeitschrift für Pädagogik*, 66 (Beiheft), 138–155. <https://doi.org/10.3262/ZPB2001138>
- Gass, S. M., Behney, J. & Plonsky, L. (2020). *Second language acquisition: An introductory course*. New York: Routledge. <https://doi.org/10.4324/9781315181752>
- Göllner, R., Wagner, W., Klieme, E., Lüdtke, O., Nagengast, B. & Trautwein, U. (2016). Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen: Chancen, Grenzen und Forschungsperspektiven. In Bundesministerium für Bildung und Forschung (Hrsg.), *Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments* (S. 63–82). Berlin: Bundesministerium für Bildung und Forschung.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Grünkorn, J., Klieme, E. & Stanat, P. (2018). Bildungsmonitoring und Qualitätssicherung. In O. Köller, M. Hasselhorn, F. W. Hesse, K. Maaz, J. Schrader, C. K. Spieß, H. Solga & K. Zimmer (Hrsg.), *Das Bildungswesen in Deutschland: Bestand und Potentiale* (S. 263–298). Bad Heilbrunn: UTB/Klinkhardt.
- Gutke, J. (2023). Kognitive Aktivierung im Fremdsprachenunterricht: Ein systematisches Review von Forschungsarbeiten aus dem deutschsprachigen Raum. *Zeitschrift für Fremdsprachenforschung*, 34(2), 145–175.

- Guttke, J. & Porsch, R. (2024). Kognitive Interviews als Methode der Inhaltsvalidierung von Fragebogenitems zur Erfassung kognitiver Aktivierung im Englischunterricht der Grundschule. *Empirische Pädagogik*, 38(1), 147–167. <https://doi.org/10.62350/WYBS7986>
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (5. Aufl.). Seelze-Velber: Klett/Kallmeyer.
- Jenßen, L., Dunekacke, S. & Blömeke, S. (2015). Qualitätssicherung in der Kompetenzforschung. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (S. 11–31). Weinheim/Basel: Beltz Juventa.
- Kersten, K., Bruhn, A.-C., Ponto, K., Böhnke, J. & Greve, W. (2018). Teacher input observation scheme (TIOS). *Studies on Multilingualism in Language Education*, 4. Universität Hildesheim.
- Kirchhoff, P. (2016). Was sollte eine gute Englischlehrkraft wissen? Über die Auswahl von Items im FALKO-E Test zum fachspezifischen Professionswissen. In M. Legutke & M. Scharf (Hrsg.), *Fremdsprachendidaktische Professionsforschung: Brennpunkt Lehrerbildung* (S. 75–98). Tübingen: Narr.
- Kirchhoff, P. (2017). FALKO-E: Fachspezifisches professionelles Wissen von Englischlehrkräften. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. Mulder (Hrsg.), *FALKO Fachspezifische Lehrerkompetenzen* (S. 113–152). Münster: Waxmann.
- Klauer, K. J. (1984). Kontentvalidität. *Diagnostica*, 30(1), 1–23.
- Kleickmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of Teaching in Science Education. In A.-K. Praetorius, J. Grünkorn & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität: Theoretische Grundfragen und quantitative Modellierungen* (S. 37–55). Weinheim: Beltz Juventa.
- Klieme, E. (2022). Unterrichtsqualität. In M. Harring, C. Rohlf's & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 411–426). Münster: Waxmann.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In E. Klieme & J. Baumert (Hrsg.), *TIMSS – Impulse für Schule und Unterricht* (S. 43–58). Bonn: Bundesministerium für Bildung und Forschung.
- Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*. Weinheim/Basel: Beltz/Juventa.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–235. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Lindl, A., Ehrich, P., Gutmiedl, M., Rader, M., Simböck, L., Gürtner, M., Böhringer, S., Krämer, A., Kirchhoff, J. & Frei, M. (2024). Und wo bleibt die Ästhetik? – Betrachtungen zu einer weiteren Dimension von Unterrichtsqualität aus interdisziplinärer Perspektive. In M. Hemmer, C. Angele, C. Bertsch, S. Kapelari, G. Leitner & M. Rothgangel (Hrsg.), *Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung* (S. 371–388). Münster: Waxmann.
- Lipowsky, F. (2020). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–118). Berlin: Springer. https://doi.org/10.1007/978-3-662-61403-7_4
- Lipowsky, F. & Bleck, V. (2019). Was wissen wir über guten Unterricht? – Ein Update. In U. Steffens & R. Messner (Hrsg.), *Unterrichtsqualität: Konzepte und Bilanzen gelingenden Lehrens und Lernens* (S. 219–249). Münster: Waxmann.
- Lipowsky, F. & Hess, M. (2019). Warum es manchmal hilfreich sein kann, das Lernen schwerer zu machen. In K. Schöppe & F. Schulz (Hrsg.), *Kreativität & Bildung – Nachhaltiges Lernen* (S. 77–132). München: kopaed.

- Loewen, S. (2020). *Introduction to instructed second language acquisition*. New York: Routledge. <https://doi.org/10.4324/9781315616797>
- Lyster, R., Saito, K. & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40. <https://doi.org/10.1017/S0261444812000365>
- Maier, U., Kleinknecht, M., Metz, K. & Bohl, T. (2010). Ein allgemeindidaktisches Kategoriensystem zur Analyse des kognitiven Potenzials von Aufgaben. *Beiträge zur Lehrerbildung*, 28(1), 84–96. <https://doi.org/10.36950/bzl.28.1.2010.9798>
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Moosbrugger, H. & Kelava, A. (2020). *Testtheorie und Fragebogenkonstruktion*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-61532-4>
- Oliver, R. & Adams, R. (2021). Oral corrective feedback. In H. Nassaji & E. Kartchava (Hrsg.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (S. 187–206). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108589789.010>
- Porsch, R. & Wilden, E. (2022). Teaching English out-of-field in primary school: Differences in professional characteristics and effects on instructional quality. In L. Hobbs & R. Porsch (Hrsg.), *Out-of-field teaching across teaching disciplines and contexts* (S. 117–134). Singapur: Springer. https://doi.org/10.1007/978-981-16-9328-1_6
- Praetorius, A.-K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49, 167–188. <https://doi.org/10.1007/s42010-021-00119-6>
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48, 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM – Mathematics Education*, 50, 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Prusse-Hess, B. & Prusse, M. (2018). *Wirksamer Englischunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Rieser, S. & Decristan, J. (2023). Kognitive Aktivierung in Befragungen von Schülerinnen und Schülern: Unterscheidung zwischen dem Potential zur kognitiven Aktivierung und der individuellen kognitiven Aktivierung. *Zeitschrift für Pädagogische Psychologie*, 1–15. <https://doi.org/10.1024/1010-0652/a000359>
- Rosebrock, A., Schlosser, S., Höhne J. & Kühnel, S. M. (2019). Einflüsse unterschiedlicher Formen der Verbalisierung von Antwortskalen auf das Antwortverhalten von Befragungspersonen. In N. Menold & T. Wolbring (Hrsg.), *Qualitätssicherung sozialwissenschaftlicher Erhebungsinstrumente* (S. 65–102). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-24517-7_3
- Schreyer, P. (2024). *Kognitive Aktivierung in der Unterrichtsinteraktion*. Münster: Waxmann. <https://doi.org/10.31244/9783830998358>
- Senkbeil, M., Ihme, J. & Wittwer, J. (2013). Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 16, 671–791. <https://doi.org/10.1007/s11618-013-0446-5>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] (2023). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Ersten Schulabschluss und den Mittleren Schulabschluss*. Verfügbar unter <https://>

- www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2023/2023_06_22-Bista-ESA-MSA-ErsteFremdsprache.pdf
- Weber, K.-A., Friege, G. & Scholz, R. (2020). Quantenphysik in der Schule – Was benötigen Lehrkräfte? Ergebnisse einer Delphi-Studie. *Zeitschrift für Didaktik der Naturwissenschaften*, 26, 173–190. <https://doi.org/10.1007/s40573-020-00119-6>
- Wibowo, J. & Heemsoth, T. (2019). Fachdidaktisches Wissen von Sportlehrer*innen testen: Überlegungen zur Inhaltsvalidität. *Zeitschrift für sportpädagogische Forschung*, 2, 88–108. <https://doi.org/10.5771/2196-5218-2019-2-88>
- Wilden, E., Porsch, R., Guttke, J. & Wellmans, L. (im Erscheinen). Das Potenzial zur kommunikativ-kognitiven Aktivierung im Englischunterricht – Befunde einer Analyse von Lernaufgaben aus Lehrwerken der Grundschule. *Fremdsprachen Lehren und Lernen*.

Wie setzen Lehrkräfte das Lehrwerk im Französisch- und Spanischunterricht ein?

Eine Studie zu Unterrichtsqualität in der Sekundarstufe I mit Fokus auf Sozialform, Binnendifferenzierung und Feedback

Abstract (der Herausgeber:innen)

In her article “How do teachers use textbooks in French and Spanish classes? A study on instructional quality in lower secondary school with a focus on interactional formats, internal differentiation and feedback,” Manuela Franke considers textbooks as an essential element of the material culture of language teaching. She argues that it is not only the textbook itself that influences the quality of teaching, but rather the teaching and learning processes that teachers initiate with it. The article presents part of her empirical study on how French and Spanish teachers work with textbooks at lower secondary level, focussing on the ways in which textbook-based phases are implemented and their significance for instructional quality. At the end, she develops hypotheses based on the research results of an ongoing empirical study.

1 Einleitung

Der Einsatz von Lehrwerken im Französisch- und Spanischunterricht muss basierend auf ihrer maßgeblichen Rolle im Unterrichtsalltag – vor allem der Sekundarstufe I – als wesentliches Element der Materialkultur schulischen Lernens und Lehrens verstanden werden. Im Durchschnitt ist ca. die Hälfte der Unterrichtszeit lehrwerkbasiert (Franke, in Vorbereitung). Das Lehrwerk stellt als konstituierender Faktor eines zeitgemäßen Unterrichts einen entscheidenden Aspekt hinsichtlich der Unterrichtsqualität dar (Fuchs et al., 2014, S. 23; Gräsel, 2010, S. 137). Die Qualität von Unterricht hängt allerdings nicht ausschließlich von der Qualität der in ihm verwendeten Materialien ab, sondern insbesondere davon, wie diese von Lehrkräften eingesetzt werden, um Lernprozesse zu initiieren. Auch wenn durch Lehrwerke bestimmt wird, was Schüler:innen lernen und den Lehrpersonen eine einfache Möglichkeit gegeben wird, sich an die Vorgaben der Curricula zu halten (Sandfuchs, 2010, S. 23), beeinflusst die didaktische Expertise von Lehrkräften die Unterrichtsqualität also maßgeblich (Korell, 2024, S. 467). Je nachdem, welche Inhalte und Aufgaben Lehrkräfte aus Lehrwerken auswählen und wie sie diese in ihren Unterricht integrieren, ergeben sich (mehr oder weniger) qualitativ hoch-

wertige Lerngelegenheiten für Schüler:innen (siehe Kaskadenmodell nach Krauss et al., 2020, Säulen 3 und 4; ausführlich Lindl et al., in diesem Band).

Aktuelle Französisch- und Spanischlehrwerke bieten basierend auf Prinzipien eines modernen Fremdsprachenunterrichts ein vielseitiges Angebot an Materialien, Aufgaben und Übungen (differenzierendes Zusatzmaterial, konkrete Hinweise zur Gestaltung des Unterrichts usw.), sodass ein qualitativ hochwertiger Unterricht mit dem Lehrwerk möglich ist. Dennoch vernachlässigt ein Unterricht, der statt auf didaktischen Entscheidungen beruhend auf einem dogmatischen und strikt chronologischen Einsatz des Lehrwerks erfolgt, wie beispielsweise von Nieweler (2000, S. 17) und Kurtz (2010) kritisiert, die Kriterien eines qualitätvollen Fremdsprachenunterrichts. Hierzu zählen z. B. die Berücksichtigung individueller Bedürfnisse der Lernenden (Differenzierung und Adaptivität) oder der Einbezug aktueller Inhalte und Themen sowie wissenschaftlicher Aspekte in den Unterricht (Auswahl und Thematisierung von Inhalten und Fachmethoden; siehe Multi-layered and integrated in conceptualizing the quality of teaching (MAIN-TEACH)-Modell; Charalambous & Praetorius, 2020; Praetorius & Gräsel, 2021, Ufer & Praetorius, 2022, Lindl et al., in diesem Band).

In den letzten Jahrzehnten wurden neben theoretischen Arbeiten zur Verwendung des Lehrwerks im Fremdsprachenunterricht (z. B. Caspari, 2010; Martinez, 2005) auch einige empirische Studien publiziert. In der Regel handelt es sich bei diesen um Interviews mit Lehrkräften (z. B. Allen, 2008; Harlan, 2000), schriftliche Umfragen mit Schüler:innen und Lehrkräften (z. B. Bohnensteffen, 2011) und/oder um Unterrichtsbeobachtung weniger Unterrichtseinheiten (z. B. Sikorová & Červenková, 2014; Nausica Marcos, 2015), die bestätigen, dass das Lehrwerk in der Sekundarstufe I regelmäßig verwendet wird. Die Fragebogenstudie von Bohnensteffen (2011) verdeutlicht zudem, dass Lehrkräfte sich nur schwer vorstellen können, ohne ein Lehrwerk zu unterrichten. Dennoch variiert der Grad des Lehrwerkeinsatzes und dessen Anpassung unter den Lehrpersonen stark (Nausica Marcos, 2015; Fuchs & Bock, 2018). Die bis dato vorliegenden Studienergebnisse verweisen auf ein Spannungsfeld zwischen der Verbreitung des Mediums bzw. dessen intensiver Nutzung in der Sekundarstufe I und den individuellen Anpassungs- und enthaltenen Variationsmöglichkeiten zur Gestaltung von Unterrichtsettings.

Eben dieses Spannungsfeld nimmt auch die Studie von Korell (2024) in den Blick und betrachtet die Lehrwerknutzung im Spanischunterricht aus drei Perspektiven: (1) die subjektive Sichtweise der Lehrkräfte auf Lehrwerke und deren Einsatz, (2) die Rezeption und Verwendung von Lehrwerken durch Spanischlehrkräfte im Rahmen der Unterrichtsvorbereitung, -nachbereitung und im Unterricht sowie (3) strukturelle Faktoren, die die Rahmenbedingungen für den Umgang mit Lehrwerken bestimmen. Zur Untersuchung dieser Aspekte wurde ein dreigliedriges Mixed-Methods-Design angewendet, das sowohl eine Fragebogenstudie als auch Interviews und videografische Unterrichtsbeobachtungen umfasst. Die Studie liefert wertvolle Erkenntnisse zu den verschiedenen Dimensionen der Lehrwerkverwendung. Hinsichtlich des Einsatzes des Lehrwerks verdeutlicht sie

unter Berücksichtigung der verschiedenen Bereiche und ihrer Wechselwirkung u. a., dass Lehrkräfte ein lehrwerkbezogenes Aufgabenverständnis entwickeln, zum kritischen und kreativen Umgang mit dem im Lehrwerk vorgeschlagenen Material bereit sind und Interesse am Ausprobieren neuer Ansätze in der Lehrwerkarbeit aufbringen. Dabei werden sie durch fachdidaktische Kenntnisse, Einstellungen und situative Parameter beeinflusst. Gleichzeitig zeigt Korell (2024, S. 365 f.) auf, dass die Aussagen der Lehrkräfte hinsichtlich der Lehrwerknutzung nicht immer mit dem tatsächlichen Unterrichtsgeschehen übereinstimmen, was die Wichtigkeit von großflächig angelegten Beobachtungsstudien unterstreicht.

Während eine ausgiebige Nutzung der Lehrwerke auf ihre Eignung als Ressource hinweisen kann, um bestimmte Lehrinhalte zu vermitteln und Unterricht transparent zu strukturieren, ist die Variation im Einsatz und die Anpassung an spezifische Kontexte und Lerngruppen von der Lehrkompetenz der Lehrkraft abhängig. Von der Prämisse ausgehend, dass Lehrwerke das Leitmedium im Französisch- und Spanischunterricht der Sekundarstufe I darstellen (u. a. Fuchs et al., 2014, S. 9; Wiater, 2005, S. 43; Kurtz, 2011, S. 3; Stöber, 2010, S. 6; Thaler, 2011, S. 18), ist die Qualität des Unterrichts eng mit dem Einsatz des Lehrwerks und damit mit der Fähigkeit der Lehrpersonen verbunden, das im Lehrwerk offerierte Material effektiv und unter Berücksichtigung der Prinzipien eines modernen Fremdsprachenunterrichts an die Bedürfnisse der Lerngruppe und das jeweilige Kompetenzziel anzupassen und gegebenenfalls zu modifizieren.

Auch wenn die Nutzungskontexte von Lehrwerken in den letzten Jahren vermehrt fokussiert wurden (Fuchs & Bock, 2018, S. 1–2), liegen bislang keine empirisch abgesicherten Informationen vor, die den Zusammenhang zwischen Lehrwerken und der Qualität von Unterricht mit einem unterrichtspraktischen Blick auf den tatsächlichen Einsatz von Lehrwerken durch Lehrkräfte betrachten (Säule 3 des Kaskadenmodells nach Krauss et al., 2020: „Bereitstellung eines möglichst qualitätsvollen Unterrichtsangebots durch die Lehrkraft“; ausführlich Lindl et al., in diesem Band).

Die hier präsentierten Ergebnisse sind Teil einer empirischen Studie, die basierend auf 133 beobachteten Unterrichtseinheiten (45 oder 90 Minuten) Einblicke in die Französisch- und Spanischunterrichtsgestaltung mit dem Lehrwerk in der Sekundarstufe I bieten. Die erhobenen Daten versprechen Hinweise zu ausgewählten Aspekten verschiedener Dimensionen von Unterrichtsqualität nach Praetorius et al. (2020, S. 417–421, siehe Abschnitt 2) und sollen Routinen des Lehrwerkeinsatzes sichtbar machen: Wie wird Fremdsprachenunterricht mit dem Lehrwerk praktiziert? Welche Unterrichtsmuster und Zugänge zum Lehrwerk sind typisch? Etc.

In diesem Beitrag wird – mit dem Ziel der Hypothesenbildung – eine erste Auswertung hinsichtlich der Inszenierungsformen von lehrwerkbasierter Unterrichtsphasen und deren Bedeutung hinsichtlich der Unterrichtsqualität vorgenommen, die zusammen mit der später erfolgenden qualitativen Analyse ausgewählter Aspekte (z. B. die Arbeit mit Lehrwerktexten oder die Art der Differenzierung lehrwerkbasierter Aufgaben) die Grundlage für weitere Forschungsarbeiten

darstellt und eine kritische Reflexion der aktuellen Unterrichtspraxis mit dem Lehrwerk erlaubt.

2 Zum Zusammenhang von fremdsprachlicher Unterrichtsqualität und Lehrwerkeinsatz

Untersuchungen der Unterrichtsqualität basieren häufig auf dem von Klieme et al. (2006) vorgeschlagenen Modell, das drei Basisdimensionen (Klassenführung, konstruktive Unterstützung und kognitive Aktivierung) beinhaltet, wobei Unterrichtsqualität als „nicht substantiell definiert“ wird. Stattdessen umfasse sie „jegliches Merkmal der Lernumgebung, das dazu beitragen könnte, investierte Lernzeit möglichst effizient zu nutzen“ (Klieme et al., 2006, S. 127). Die drei Basisdimensionen werden von Praetorius et al. (2020, S. 417–421) zu sieben generischen Dimensionen erweitert, wobei diese konkret naturwissenschaftliche Fächer, Sport und Geschichte in den Blick nehmen: I Auswahl und Thematisierung von Inhalten und Fachmethoden, II kognitive Aktivierung, III Unterstützung des Übens, IV formatives Assessment, V Unterstützung des Lernens aller Schüler:innen, VI sozio-emotionale Unterstützung und VII Klassenführung. Diese sieben Dimensionen werden im MAIN-TEACH-Modell (Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) in Form einer integrativen Systematisierung zusammengeführt (ausführlich zum MAIN-TEACH-Modell siehe Lindl et al., in diesem Band). Einen Vorschlag zur Berücksichtigung der Schulfremdsprache Englisch, der auch auf die zweiten und dritten Fremdsprachen übertragbar ist, beinhaltet

„die Doppelrolle der Zielsprache [...] als Gegenstand und Medium des Unterrichts, die Bedeutung der Sprach*nutzung* durch Schüler*innen in kommunikativen und interaktiven Unterrichtssituationen, die auf den hohen Stellenwert der Lerner*innenorientierung hinweist, [den] positive[n] Umgang mit Fehlern als integralem Bestandteil des Sprachlernens, welcher im Zusammenhang mit der Funktion des Übens gesehen werden muss [...] [Hervorhebungen im Original]“ (Wilden, 2021, S. 217).

Durch das Lehrwerk bzw. dessen Einsatz kann jede der sieben Dimensionen von Unterrichtsqualität nach Praetorius et al. (2020; auch Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) beeinflusst werden. Denn durch die Vorauswahl von Inhalten bedingt es die im Unterricht tatsächlich bearbeiteten Themen und Aufgabenformate. Dabei sollten Lehrwerke auf das kognitive Niveau der Lernenden abgestimmt sein und die Leistungsüberprüfung z. B. durch den Einsatz von Zusatzmaterialien wie Klassenarbeitstrainer unterstützen. Zur Förderung des globalen Lernens und zum Erwerb von Lernstrategien tragen beispielsweise Grammatikbeilagen und Vokabeltrainer bei. Das Lehrwerk kann die Klassenführung durch einen ritualisierten Aufbau von Unterrichtsstunden und den Einsatz von Hausaufgaben strukturieren. Es ermöglicht zudem die Schaffung von Transparenz –

Lernende können bei lehrwerkbasierem Unterricht im Falle von Abwesenheit verlässlich nacharbeiten, sie können kommende Stunden vorbereiten, zum Teil Lösungen oder Transkripte von Hörtexten einsehen etc. Zudem kann durch den Einsatz des Lehrwerks die Interaktion zwischen Lernenden und Lehrkräften im Sinne einer sozio-emotionalen Unterstützung gestärkt werden (z. B. durch Strategieteile, Tipps, motivierende Begleiter:innen wie wiederkehrende Lehrwerkfiguren oder Tiere wie der Hund in *Découvertes*, Klett, direkte Ansprache der Lernenden).

In diesem Beitrag werden auf Unterrichtsbeobachtungen basierende Daten (Näheres hierzu in Abschnitt 3) hinsichtlich der folgenden Aspekte der Dimensionen III, IV und V von Unterrichtsqualität (Praetorius et al., 2020, S. 417–421), die auch in den Ebenen des MAIN-TEACH-Modells (Charalambous & Praetorius, 2020, ausführlich Lindl et al., in diesem Band) verankert sind, sowie der von Wilden (2021, S. 217) für die Schulfremdsprache Englisch vorgeschlagenen Ergänzung ausgewertet:

III. Unterstützung des Übens

III.2 Konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schüler:innen beim Üben

IV. Formatives Assessment

IV.3 Qualitativ hochwertiges *Feedback* an die Schüler:innen (z. B. beinhaltet feedback, feed-up und feed-forward Komponenten)

IV.4 Nutzung des *Feedbacks* als Grundlage für die Ausrichtung des weiteren Unterrichts

V. Unterstützung des Lernens aller Schüler:innen

V.1 Bereitstellung eines Lernumfelds, das *produktives Verhalten* fördert

V.2 *Differenzierung* und Adaptivität

V.3 Förderung der *aktiven Mitwirkung von allen Schüler:innen*

Ergänzung: fremdsprachenspezifisches Merkmal

Bedeutung der *Sprachnutzung* durch Schüler:innen in *kommunikativen und interaktiven Unterrichtssituationen*

In den einzelnen Komponenten eines Lehrwerks sind verschiedene Elemente integriert, die gezieltes Üben unterstützen sollen. Besonders im Arbeitsheft finden sich Aufgabenformate, die darauf abzielen, Schüler:innen tiefer in die Lehrinhalte einzuführen und das Verständnis bzw. fremdsprachliche Fertigkeiten zu festigen. Jedoch sind Lerngruppen äußerst vielfältig und heterogen in ihren individuellen Voraussetzungen und Bedürfnissen. Zudem ist Lernen ein äußerst individueller Prozess (Hoffmann, 2010, S. 161) und somit erfolgt auch der (Fremd-)Spracherwerb auf sehr unterschiedliche Art und Weise, wobei Lernende auf eine Vielfalt an sprachlichem und lerntheoretischem Wissen zurückgreifen können (Caspari & Holzbrecher, 2016, S. 9; Hoffmann, 2010, S. 162). Die folgenden Differenzlinien können hierbei Berücksichtigung finden: a) Leistungsfähigkeit, b) Motivation, In-

teresse, Lern- und Leistungsbereitschaft, c) Vorwissen, Sprachkenntnisse, Lernvoraussetzungen, d) Lernstil, e) Lerntempo, f) Lebenswelt, soziokulturelle Herkunft, g) Zuwanderungsgeschichte: Sprachlernerfahrungen, Sprachregister (Caspari & Holzbrecher, 2016, S. 10–13).

Die Heterogenität der Lernenden bringt primär die Aufgabe mit sich, Lerngelegenheiten zu gestalten, in denen alle Schüler:innen ihrer eigenen Vorgehensweise im Lernprozess folgen können und individuelle Änderungen derselben möglich sind bzw. zu solchen angeregt wird (Hoffmann, 2010, S. 162). Entsprechend eignen sich „didaktische Formen [...], die unterschiedliche Lernniveaus und Leistungsstände und damit diverse Aktivierungsstufen und Bewusstseins Ebenen zulassen und entsprechend fördern“ (Hoffmann, 2010, S. 163). Aktuelle Lehrwerke bieten insbesondere hinsichtlich des Anforderungsniveaus verschiedene Differenzierungsmöglichkeiten an, die zur gezielten und individuellen Unterstützung nicht nur, aber auch des Übens genutzt werden können. Zusätzlich ergibt sich aus der Heterogenität von Lerngruppen die Notwendigkeit der Adaptation von Lehrwerkelementen durch Lehrkräfte (unterste Ebene des MAIN-TEACH-Modells „Differenzierung und Adaptation“, Charalambous & Praetorius, 2020; Praetorius & Gräsel, 2021, Ufer & Praetorius, 2022, Lindl et al., in diesem Band). Für die in diesem Beitrag durchgeführte Analyse wird Differenzierung als die „Gesamtheit der unterrichtsorganisatorischen und didaktisch-methodischen Maßnahmen [verstanden], die innerhalb einer Lerngruppe mit der Intention ergriffen werden, die unterschiedlichen Voraussetzungen und individuell angemessene Lernprozesse zu ermöglichen“ (Tönshoff, 2004, S. 277), sodass alle Lernenden bestmöglich gefördert werden können.

Neben der Differenzierung gehört auch ein konstruktiver Umgang mit Fehlern bzw. Feedback zur Unterstützung des Lernens und stellt eine weitere wichtige Säule für die Förderung der Unterrichtsqualität dar (hinsichtlich naturwissenschaftlicher Fächer, Sport und Geschichte siehe: Dimension III.2 konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schüler:innen beim Üben, IV.3 Qualitativ hochwertiges Feedback an die Schüler:innen und IV.4 Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts, Praetorius et al., 2020, S. 415–416 und S. 419; für die Fremdsprachen siehe Wilden, 2021, S. 217; Charalambous & Praetorius, 2020, ausführlich Lindl et al., in diesem Band). Lehrkräfte können eine lernförderliche Umgebung schaffen, in der Schüler:innen Fehler als Gelegenheiten zur Reflexion und Verbesserung betrachten. Dabei liegt die Verantwortung für die Anleitung zum Umgang mit Fehlern ausschließlich bei den Lehrkräften, denn das Lehrwerk bietet Feedback in erster Linie durch die Bereitstellung von vereinzelt Selbstbewertungsbögen oder Lösungen zu Aktivitäten vor allem aus dem Arbeitsheft. Um die Schüler:innen zu ermutigen, aus Fehlern zu lernen und ihre Herangehensweise kontinuierlich zu verbessern, sind einfühlsame pädagogische und didaktische Ansätze im Umgang mit dem Lehrwerk erforderlich. Eine solche Feedbackkultur ermöglicht den Lernenden, ihre Fortschritte zu bewerten und an ihren Schwächen zu arbeiten. Dies kann beispielsweise durch den Austausch über Lösungsstrategien oder die Dis-

kussion über Fehler (Steffens & Haenisch, 2019, S. 299) geschehen. Feedback kann zudem dafür sorgen, dass Lernende (mehr) Verantwortung für ihren Lernprozess übernehmen (können), ihr Selbstbewusstsein stärken und entsprechend als zentral für die Förderung der Autonomie der Lernenden verstanden werden. Dabei wird Feedback hier nach Rösler (2004, S. 177) als Hinweis definiert, den Lernende hinsichtlich ihrer schriftlichen wie mündlichen Aktivitäten im Fremdsprachenunterricht erhalten und der der Einschätzung in Bezug auf die Angemessenheit bzw. Korrektheit dient (siehe auch Hattie & Timperley, 2007, S. 82).

Um die im Kontext des Fremdsprachenunterrichts zentralen sprachlichen Fertigkeiten optimal zu entwickeln, bedarf es entsprechender, an den Lernenden orientierter Voraussetzungen, Schauplätze und Anlässe (Wapenhaus, 2018, S. 256). Hierzu zählt auch die Gestaltung von Lernsettings, in denen möglichst viele Schüler:innen über einen möglichst langen Zeitraum aktiviert sind (vgl. in Bezug auf die Relevanz der Aktivierung der Lernenden im Kontext der Förderung von Mündlichkeit u. a. Blume & Nieweler, 2018, S. 5) und in denen den Lernenden eine Fremdsprachennutzung in „kommunikativen und interaktiven Unterrichtssituationen“ (Wilden, 2021, S. 217) ermöglicht wird. Die Wahl von Sozialformen kann Aussagen hinsichtlich der kommunikativen Gestaltung des Unterrichts ermöglichen und trägt daher wesentlich zur Erfassung qualitativer Aspekte des Unterrichts bei (Wilden, 2021, S. 217), da die Förderung einer aktiven Beteiligung aller Lernenden als Teil der Unterrichtsdimensionen V Unterstützung des Lernens aller Schüler:innen sowie VII Klassenführung und somit als grundlegender Bestandteil eines interaktiven Lehr- und Lernprozesses betrachtet werden kann (Praetorius et al., 2020, S. 420–421; Charalambous & Praetorius, 2020; Lindl et al., in diesem Band).

Unter Berücksichtigung der genannten Aspekte ergeben sich folgende Fragestellungen für die anschließende Datenanalyse:

- Welche Sozialform wählen Lehrkräfte für die Arbeit mit dem Lehrwerk aus? (Abschnitt 4.1)
- Wie wird bei der Arbeit mit dem Lehrwerk differenziert? (Abschnitt 4.2)
- In welcher Form geben Lehrkräfte in lehrwerkbasierten Unterrichtsphasen Feedback? (Abschnitt 4.3)

3 Studiendesign

Die hier vorgestellte qualitative Studie gründet sich auf Beobachtungen von Französisch- und Spanischunterricht der Sekundarstufe I, in denen eine möglichst präzise Protokollierung der von den Lehrkräften formulierten Aussagen (z. B. Arbeitsaufträge, Impulse usw.) zur Anleitung von Unterrichtsphasen mit dem Lehrwerk vorgenommen wurde (siehe Abschnitt 3.1 zur Darstellung des Erhebungsinstruments). Die Datenerhebung fand im Rahmen des Praxissemesters im Master Lehramt Französisch und Spanisch an der Universität Potsdam statt und wurde von den Studierenden durchgeführt (jeweils 16 Personen pro Fach). Diese

Form der Datenerhebung wurde gewählt, da die Beobachtung durch Studierende im Praxissemester für die Lehrkräfte ein vertrautes und alltägliches Setting darstellt. Ihr Unterricht wurde vier Monate lang regelmäßig von den jeweiligen Studierenden hospitiert. Dadurch ergab sich – im Gegensatz zu z. B. gefilmten Stunden – ein dem tatsächlichen alltäglichen Unterrichtsgeschehen sehr nahekommendes Setting und blieb die Reaktivität durch die Beobachtungssituation gering (Döring & Bortz, 2016, S. 330). Zusätzlich konnten sich Beobachtungsfehler aufgrund der Anzahl an Beobachter:innen (32 Personen) ausmitteln, sodass Verzerrungen in den dokumentierten Unterrichtseinheiten minimiert werden (zu den Vor- und Nachteilen verschiedener methodischer Zugänge zur Erfassung von Unterrichtsqualität siehe Lindl et al., in diesem Band).

Es wurden in sich abgeschlossene Unterrichtseinheiten mit einer Dauer von 45 oder 90 Minuten beobachtet. Ganze Unterrichtsreihen bzw. -sequenzen wurden nicht dokumentiert. Das Beobachtungsobjekt (Döring & Bortz, 2016, S. 326) war die durch die Lehrkraft gestaltete Arbeit mit dem jeweiligen Lehrwerk. Die beobachteten Unterrichtseinheiten dienten gleichzeitig im das Praxissemester begleitenden Seminar als Ausgangspunkt für die Reflexion des beobachteten und eigenen Lehrwerkseinsatzes, die in Implikationen für zukünftige Unterrichtsplanungen mit dem Lehrwerk überführt wurde.

3.1 Stichprobe

Die Stichprobe umfasst 133 Französisch- und Spanischunterrichtseinheiten. Sie wurden von 44 Lehrkräften an 31 Schulen (17 an Gymnasien, 7 an Gesamtschulen und 4 an Integrierten Sekundar- bzw. Oberschulen) in Berlin und Brandenburg vor Einsetzen der pandemiebedingten Einschränkungen (Oktober 2019 bis Februar 2020) durchgeführt. Die relativ große Anzahl an beobachteten Unterrichtseinheiten minimiert den möglichen Einfluss des Unterrichtsstils einer bestimmten Lehrkraft auf die Gesamtstichprobe. Eine Lehrkraft wurde mindestens einmal und höchstens neunmal beobachtet. Eine informierte schriftliche Einwilligungserklärung der Lehrkräfte wurde eingeholt, die zusätzliche Informationen zum Forschungsprojekt bereitstellte, sodass es sich um eine offene Beobachtung handelte (Döring & Bortz, 2016, 329).

Es wurden Unterrichtseinheiten mit 31 verschiedenen Lehrwerken beobachtet, wobei das älteste Lehrwerk aus dem Jahr 2008 (*¡Apúntate!* 1, 1 Unterrichtsstunde¹) stammt und die beiden jüngsten Lehrwerke aus dem Jahr 2019 (*¡Apúntate!* 4 und *¿Qué pasa?* 4). Die genutzten Französischlehrwerke sind zwischen 2012 (*Découvertes* 1) und 2017 (*Horizons, À toi!* 3) erschienen. Alle in der Stichprobe verwendeten Lehrwerke verfügten über Differenzierungsangebote (in der Regel wurde durch ein Symbol auf entsprechende leichtere oder schwerere Aktivitäten im hinteren Teil des Lehrwerks verwiesen). Zu den Sozialformen, in denen Akti-

1 Die neueste Ausgabe von *¡Apúntate!* 1 (2016) wurde in 22 Einheiten genutzt.

vitäten stattfinden sollten, waren in allen Lehrwerken vereinzelt entweder durch entsprechende Symbole oder aufgrund der Aufgabenstellung selbst Vorschläge zu finden. Alle Lehrwerke verfügten über Lösungen zu einzelnen Aktivitäten vor allem aus den Arbeitsheften, die den Lernenden entweder zur Verfügung stehen (z. B. in Form eines integrierten Lösungshefts oder -teils) oder von der Lehrkraft bereitgestellt werden konnten.

3.2 Datenerhebung

3.2.1 Aufbau des Beobachtungsinstruments

Die von den Studierenden durchgeführte teilstrukturierte Beobachtung (Martin & Hanington, 2019, S. 158; Döring & Bortz, 2016, S. 328) erfolgte anhand eines zur inhaltlichen und methodischen Absicherung im *peer-group*-Verfahren entwickelten Beobachtungsbogens. Die Studierenden verzeichneten in diesem unter anderem die eingesetzte Lehrwerkkomponente (z. B. Arbeitsheft oder Schulbuch), Seitenzahl und Aufgabennummer, Beginn und Ende der Arbeit mit dem jeweiligen Lehrwerkelement (z. B. Aufgabenstellungen, Lehrwerktexte, Hörtexte, Zeichnungen etc.) sowie möglichst wortwörtlich den von der Lehrkraft formulierten Arbeitsauftrag und die Sozialform.

3.2.2 Prüfung des Beobachtungsinstruments

Zur Prüfung des Beobachtungsinstruments hinsichtlich seiner Eignung zur Dokumentation des Unterrichtsgeschehens durch mehrere Beobachter:innen wurde dieses von neun Lehramtsstudierenden des Faches Französisch und/oder Spanisch (nachfolgend Test-Beobachter:innen genannt) sowie der Autorin im Rahmen eines Pre-Tests im Vorfeld der Studie eingesetzt, um das lehrwerkbasierte Unterrichtsgeschehen in einer videografierten Unterrichtseinheit Englisch² zu dokumentieren. Anschließend wurde eine qualitative Prüfung des Instruments hinsichtlich der Anfälligkeit für Wahrnehmungsfehler (z. B. Döring & Bortz, 2016, S. 331) durchgeführt. Hierzu wurde ein Abgleich der zu der videografierten Un-

2 Die Videografie und die in ihr verwendeten Lehrwerkseiten sind öffentlich unter folgendem Link zugänglich: <https://tube.switch.ch/videos/eld59906>. Es handelt sich um eine Stunde von 45 Minuten, die aufgrund des Zuschnitts ein 36 Minuten langes Video ergibt. Die Stunde wurde 2010 an der Kantonschule Freudenberg (Zürich, Schweiz) in einer achten Klasse mit dem Lehrwerk *New English File Elementary*, o. J. durchgeführt. Im Video sind sowohl die Lehrkraft als auch Schüler:innen, Tafelbild und zum Teil die Materialien der Lernenden zu sehen. Die Videografie einer Englischstunde wurde gewählt, um die Prüfung des Beobachtungsinstruments sowohl mit Spanisch- als auch Französischstudierenden an einem gemeinsamen Termin durchzuführen. Alle Personen gaben in einer mündlichen Abfrage an, dass sie Englisch mindestens auf B2-Niveau beherrschen.

terrichtsstunde entstandenen Beobachtungsprotokolle hinsichtlich der Vollständigkeit der dokumentierten Lehrwerkelemente und der nachvollziehbaren Erfassung des Unterrichtsgeschehens mit dem Lehrwerk (z. B. (un-)veränderte Durchführung der Lehrwerkaktivitäten³, gewählte Sozialform oder Binnendifferenzierung usw.) vorgenommen. Die Sichtung der zur Testung des Instruments ausgefüllten Beobachtungsprotokolle erfolgte zur Qualitätssicherung durch eine dritte, nicht in das Projekt involvierte Person aus der *Peer-Group* der Autorin.

In der dokumentierten Videostunde wurden insgesamt sechs Aktivitäten exklusive der Hausaufgabe beobachtet. Die Lehrkraft gestaltete die lehrwerkbasier-ten Unterrichtsaktivitäten unterschiedlich komplex. Zum Teil wurden die Aktivitäten aus dem Lehrwerk (*New English File Elementary*) ergänzt: So besteht Aufgabe 1b auf Seite 20 darin, Begriffe aus einer Liste den passenden Beschreibungen zuzuordnen. Die Lehrkraft ließ die Lernenden zusätzlich die Beschreibungen laut im Plenum vorlesen. Die Dokumentation einer Aktivität durch die Test-Beobachter:innen galt dann als vollständig, wenn alle von der Lehrkraft formulierten Arbeitsaufträge nachvollziehbar festgehalten wurden. Zusätzlich sollten Seitenzahl und Aufgabennummer sowie die Bearbeitungsdauer einer Aktivität notiert werden. Letzteres erfolgte, indem der Beginn und das Ende der Arbeit an einer Aktivität dokumentiert wurden. Das Video wurde einmalig im Plenum ohne Unterbrechung gezeigt, da auch im zu beobachtenden Unterricht im Rahmen der Hospitationen *ad hoc*-Notizen angefertigt werden mussten.

Der Vergleich der Beobachtungsprotokolle des Testdurchlaufs zeigte, dass nicht immer alle von der Lehrkraft zu einer Aktivität aus dem Lehrwerk geforderten Handlungen von allen Test-Beobachter:innen korrekt wiedergegeben wurden. Von den insgesamt 15 zu benennenden Handlungen der Lernenden wurden nur sechs von allen Test-Beobachter:innen festgehalten. Drei der gewünschten Lernendenhandlungen wurden von acht Test-Beobachter:innen notiert, fünf von sieben Test-Beobachter:innen und eine von den Schüler:innen auszuführende Tätigkeit wurde nur von lediglich sechs Test-Beobachter:innen erfasst. Auslassungen wurden in 94,4% der Fälle durch „...“ kenntlich gemacht. Auf dieser Erkenntnis fußend wurden alle Unterrichtsprotokolle aus der Studie ausgeschlossen, in denen durch „...“ auf eine unvollständige Dokumentation der Unterrichtseinheit verwiesen wurde, um die sich aus den Auslassungen in der Dokumentation ergebende verfälschte Darstellung des Unterrichtsgeschehens zu minimieren. Das Beobachtungsinstrument wurde zudem in verschiedenen weiteren Teststufen geringfügig modifiziert.

3 Unter Lehrwerkaktivitäten sind Arbeitsaufträge, die Lehrkräfte zu Materialien (z. B. einem Bild, Text usw.) aus dem Lehrwerk formulieren, oder im Lehrwerk abgedruckte Aufgaben und Übungen gefasst.

3.2.3 Schulung der Beobachter:innen

Um eine möglichst hohe Objektivität bei der Beobachtung zu gewährleisten, wurde vor dem Einsatz des Beobachtungsbogens eine methodologische Schulung der Beobachter:innen abgehalten (Döring & Bortz, 2016, S. 332), die unter anderem folgende Bereiche abdeckte: theoretischer Input zum Thema Unterrichtsbeobachtung, Verwendung des Beobachtungsinstruments am Beispiel einer Videostunde, Reflexion über die Beobachtung, Klärung der Abkürzungen, Sensibilisierung für den Unterschied zwischen Beobachtung und Interpretation.

3.3 Datenauswertung

Im ersten Auswertungsschritt wurden je Unterrichtseinheit folgende Informationen aus den Beobachtungsprotokollen in einer Excel-Tabelle festgehalten: Schultyp, Klassenstufe und Lernjahr, Unterrichtsform, Angaben zum Lehrwerk, Anzahl der Aktivitäten mit dem Lehrwerk pro Stunde, Gesamtzeit des Lehrwerkeinsatzes pro Stunde und die erteilten Hausaufgaben. Anschließend wurden die beobachteten Unterrichtseinheiten nach Aktivitäten mit dem Lehrwerk (z. B. Lesen eines Textes, Lösen einer Aufgabe usw.) sortiert. Als mit dem Lehrwerk durchgeführte Aktivität zählen Arbeitsaufträge, die die Lehrkräfte selbst zu Elementen (z. B. einem Bild, Text usw.) aus dem Lehrwerk formulieren oder im Lehrwerk formulierte Aufgaben und Übungen. Jede Teilaufgabe aus dem Lehrwerk zählt als einzelne Aktivität, da nicht immer alle Teilaufgaben von Lehrkräften eingesetzt bzw. nicht immer alle Teilaufgaben einer Aufgabennummer in derselben Sozialform, mit derselben Differenzierung usw. verwendet werden.

Zu jeder mit dem Lehrwerk durchgeführten Aktivität wurden die Aussage der Lehrkraft, die gewählte Sozialform sowie Informationen zur Differenzierung der Aktivität und zum Feedback zu den Ergebnissen aus der Aktivität vermerkt. Ausgehend von einem breiten Verständnis von Differenzierung, wie es beispielsweise Schulz bereits im Jahr 1969 vorschlägt, wurde bei der Kodierung der Beobachtungsprotokolle jegliche Vorgehensweise der Lehrkräfte, die dazu diente, die Lernenden zu unterstützen, als Differenzierung kodiert, so z. B. auch Verweise auf die Vokabelliste im hinteren Teil des Lehrbuchs. Nicht als Differenzierung erfasst wurden individuelle Zuwendungen von Lehrkräften, wie sie beispielsweise in Einzel- oder Gruppenarbeiten üblich sind. Da diese aus den Daten nicht zweifelsfrei ermittelbar sind, wird durch deren Ausschluss einer Verfälschung der Ergebnisse vorgebeugt.

Die Beobachtungsprotokolle wurden zudem hinsichtlich der eingesetzten Sozialformen zu jeder Unterrichtsaktivität kodiert (siehe Abschnitt 4.1). Es wurde zum einen ausgewertet, in welcher Sozialform Lehrende Aktivitäten aus dem Lehrwerk durchführen, zum anderen, ob sie im Lehrwerk für bestimmte Aktivitäten vorgeschlagene Sozialformen übernehmen oder ändern. Wenn die im Lehrwerk vorgegebene Sozialform im Wesentlichen erhalten blieb, aber beispielsweise

ergänzt wurde (z.B. im Lehrwerk vorgeschlagene Partner:innenarbeit wird als solche durchgeführt und im Plenum gesichert), wurde dies nicht als Veränderung der Sozialform kodiert.

Differenzierungsmaßnahmen wurden immer gemäß der Ausgangsaktivität kodiert, d.h. es wurden nicht die Aktivitäten oder das Material gezählt, die der Differenzierung dienten (z.B. zusätzliche Aktivitäten oder eine Liste mit Redemitteln), sondern diejenigen, zu deren Unterstützung sie herangezogen wurden (siehe Abschnitt 4.2). Es wurden nur Aktivitäten als differenziert bewertet, wenn die Lehrkraft die Differenzierung explizit formulierte. Wenn die Lernenden beispielsweise Begriffe im Lehrwerk nachschlagen, ohne von der Lehrkraft dazu aufgefordert worden zu sein (z.B. [Hinweis zur S-Aktivität: LW wird als Hilfsmittel z.B. Vokabel nachschlagen genutzt] (BoSp.19/20_7.5a+b)), wurde dies nicht als Differenzierung kodiert, da es sich hierbei nicht um eine von der Lehrkraft vorgesehene Maßnahme handelte.

Hinsichtlich des Feedbacks zu Aktivitäten gibt das Lehrwerk selbst kaum Vorgaben. Entsprechend liegt es in der Verantwortung der Lehrkräfte, die Methode (z.B. direkte Ergebnissicherung, Autokorrektur durch *peer-to-peer*-Feedback) zu bestimmen, die am besten zu den Bedürfnissen und Fähigkeiten der Lernenden passt und ihnen die bestmögliche Unterstützung im Lernprozess bietet. In Abschnitt 4.3 wird die vorliegende Stichprobe bezüglich der Häufigkeit und des Zeitpunkts von Feedback sowie der rückmeldenden Person ausgewertet. Ob es sich bei den beobachteten Rückmeldungen um eine Fehlerkorrektur oder um Lob handelt und wie genau dieses gestaltet wird, wird in der qualitativen Auswertung genauere Betrachtung finden und ist nicht Gegenstand des vorliegenden Beitrags.

Mit dem Ziel der Qualitätssicherung wurden 30 % der Kodierungen hinsichtlich dieser Aspekte zur Feststellung der *Intrarater*-Übereinstimmung nach über zwölf Wochen erneut geprüft (Przyborski & Wohlrab-Sahr, 2014). 125 von 133 (93,9 %) Kodierungen stimmten überein. Doppelte oder fehlende Kodierungen wurden nachfolgend identifiziert und gelöscht bzw. ergänzt. Ebenfalls wurden strittige Kodierungen nach Diskussion mit einer Fachkollegin aus der *Peer-Group* im gesamten Material überarbeitet. Nach der in diesem Beitrag präsentierten quantitativen Verarbeitung der Daten wird eine qualitative Analyse nach Kuckartz (2012) einen vertiefenden Einblick in den Einsatz des Lehrwerks (z.B. zur Arbeit mit Lehrwerkstexten oder -bildern) ermöglichen.

4 Ergebnispräsentation und -diskussion

Die nachstehende Analyse hat zum Ziel, einen ersten Einblick zu geben, wie Lehrende das Lehrwerk nutzen, um so Rückschlüsse auf mögliche Fragen zur Strukturierung des Unterrichts zu ziehen (Stichwort: „Auswahl und Thematisierung von Inhalten und Fachmethoden“, Charalambous & Praetorius, 2020; Lindl et al., in diesem Band). Es geht also nicht darum, die Bedeutsamkeit der aus den Lehrwerken ausgewählten Lehrinhalte oder deren fachliche Richtigkeit zu bewerten.

Für diese Analyse wurden lediglich Unterrichtseinheiten ausgewählt, in denen das Lehrwerk genutzt wurde; hierzu zählen auch Unterrichtseinheiten, in denen nur die Hausaufgabe lehrwerkbasierend war, in der tatsächlichen Kontaktzeit mit den Lernenden jedoch ohne Lehrwerk unterrichtet wurde. Im Folgenden wird die Stichprobe hinsichtlich des Einsatzes von Sozialformen (4.1), Binnendifferenzierung (4.2) und Feedback (4.3) in lehrwerkbasierenden Unterrichtsphasen ausgewertet.

4.1 Einsatz von Sozialformen in lehrwerkbasierenden Unterrichtsphasen

4.1.1 Lernjahre 1 bis 5

Die $n = 441$ Aktivitäten mit dem Lehrwerk wurden in Einzel- (EA), Partner:innen- (PA), Gruppenarbeit (GA), Plenum oder in verschiedenen Kombinationen aus den Sozialformen durchgeführt (siehe Abb. 1). Vor allem wurden Aktivitäten mit dem Lehrwerk in Einzelarbeit (86 Aktivitäten, 20,5%) oder Plenum (122 Aktivitäten, 29,1%) sowie einer Kombination aus beiden (112 Aktivitäten, 26,7%) beobachtet.

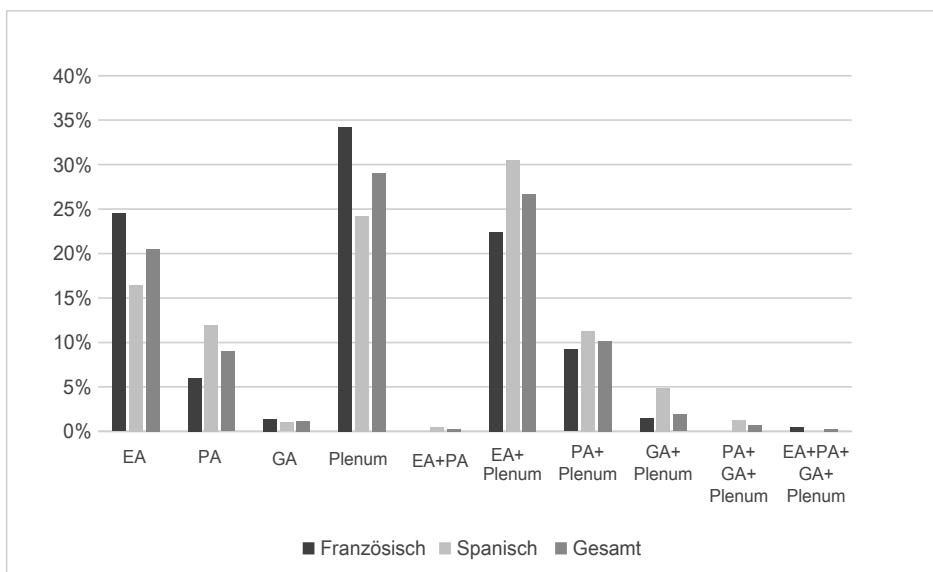


Abbildung 1: Sozialformen in lehrwerkzentrierten Unterrichtsphasen (1. bis 5. Lernjahr)

Insgesamt wurden 320 der mit dem Lehrwerk bearbeiteten Aktivitäten (76,3%) in Einzelarbeit oder im Plenum durchgeführt.

4.1.2 Lernjahre 3 bis 5

Die Vorteile von Einzel- und Plenumsarbeit für niedrige Lernjahre legen die Vermutung nahe, dass die oben dargestellten Ergebnisse (siehe Abschnitt 4.1.1) mit der Tatsache, dass die Stichprobe zu über 60% aus Lernenden der siebten Klasse, erstes Lernjahr, besteht, zusammenhängen. Aus diesem Grund wird im Folgenden eine Auswertung einer Substichprobe ($n = 60$) vorgenommen, bei der das erste und zweite Lernjahr ausgeschlossen werden. Diese ergibt, dass auch von den im dritten bis fünften Lernjahr beobachteten 60 Aktivitäten über 70% in Einzel- (8 Aktivitäten, 13,5%) oder Plenumsarbeit (22 Aktivitäten, 37,4%) bzw. in einer Kombination aus beiden (19 Aktivitäten, 32,3%) durchgeführt wurden, wobei die Arbeit in lehrwerkbasierten Phasen bevorzugt im Plenum oder in einer Kombination aus Einzelarbeit und Plenum stattfand (Abb. 2). Im vierten und fünften Lernjahr wurde keine Aktivität aus dem Lehrwerk nur in Gruppenarbeit durchgeführt; einmal wurde im vierten Lernjahr eine Kombination aus Gruppenarbeit und Plenum für eine Lehrwerksaktivität eingesetzt. Partner:innenarbeit wurde im vierten und fünften Lernjahr immer mit einer Plenumsarbeitsphase kombiniert, wohingegen sie im dritten Lernjahr auch allein eingesetzt wurde (siehe Abb. 2).

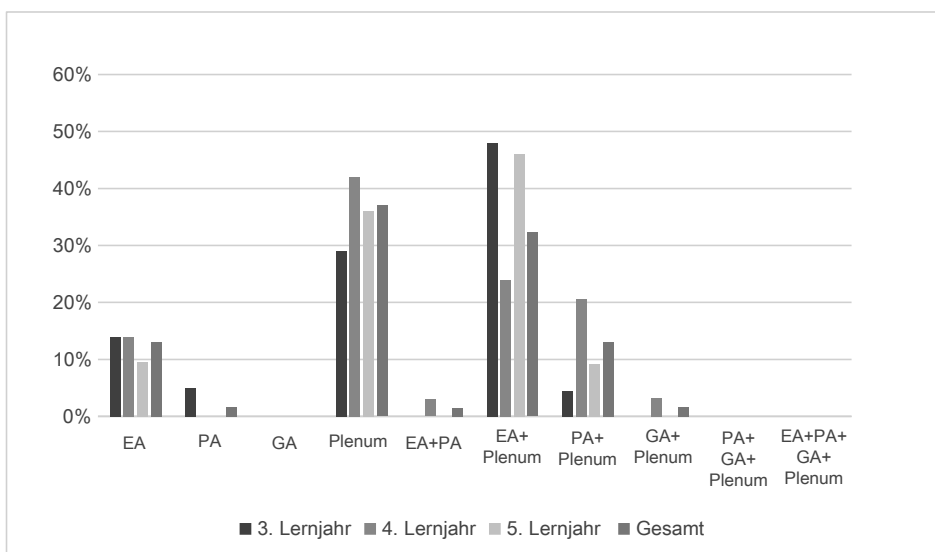


Abbildung 2: Sozialformen in lehrwerkzentrierten Unterrichtsphasen (3. bis 5. Lernjahr)

4.1.3 Veränderung der vom Lehrwerk vorgeschlagenen Sozialform

Insgesamt sind in der vorliegenden Untersuchung 53 Aktivitäten beobachtet worden, bei denen die Sozialform vom Lehrwerk vorgeschlagen wurde. Bei 38 dieser Aktivitäten erfolgte keine Änderung der empfohlenen Sozialform durch die

Lehrkräfte: In Französisch wurden entsprechend der Vorschläge aus dem Lehrwerk fünf Aktivitäten als Einzelarbeit (EA), zwölf als Partner:innenarbeit (PA) und zwei als Gruppenarbeit (GA) durchgeführt. In Spanisch hingegen wurden bei einer Aktivität die vom Lehrwerk vorgeschlagene Sozialform Einzelarbeit, bei 13 Partner:innenarbeit und bei fünf Gruppenarbeit übernommen. Die 15 beobachteten Modifikationen der Sozialform verteilten sich, wie in Tabelle 1 dargestellt ist.

Tabelle 1: Veränderung der im Lehrwerk vorgeschlagenen Sozialform

Änderung der Sozialform	Französisch	Spanisch	Gesamt
PA zu Plenum	4	6	10
PA zu EA	1	1	2
EA zu PA	2	0	2
GA zu EA	1	0	2
Gesamt	8	7	16

Besonders häufig wurden im Lehrwerk vorgeschlagene Partner:innen- bzw. Gruppenarbeiten in Plenums- bzw. Einzelarbeit durchgeführt (insgesamt 14-mal). In zwei Aktivitäten aus dem Französischunterricht wurde eine Einzelarbeit in Partner:innenarbeit geändert (z.B. einen Text statt allein zu zweit schreiben; siehe Tab. 1).

4.2 Binnendifferenzierung mit dem Lehrwerk

Von den beobachteten 44 Lehrkräften haben elf Französisch- und elf Spanischlehrkräfte Differenzierungsangebote zu den Aktivitäten aus dem Lehrwerk gemacht. Nur bei 45 (23 Französisch, 22 Spanisch; zusammen 10,3 %) Aktivitäten greifen die Lehrenden auf eine Differenzierung zurück, 396 Aktivitäten (89,7 %) werden entsprechend nicht differenziert (siehe Tab. 2). In der Stichprobe konnten folgende Arten der Differenzierung beobachtet werden: Lerntempo, Hilfestellung, Neigungs- und Themenwahl sowie eine Kombination aus Lerntempo und Hilfestellung. Hauptsächlich fand eine Differenzierung hinsichtlich des Lerntempos (17-mal, z.B.: „Wenn ihr fertig seid, dann macht die Übungen Nr. 8 und Nr. 9. [...]“ (BoFr.19/20_2.3a+b)) oder durch die Bereitstellung von unterstützenden Materialien (18-mal) statt (siehe Tab. 2): „Wenn ihr nicht weiter wisst [sic], dürft ihr euch vorne eine Hilfe nehmen“ [Hinweis der/s Studierenden: „L. [=Lehrkraft] hält Zettel mit Hilfestellung hoch und legt sie anschließend vorne hin.“ (BoFr.19/20_4.4). Eine Kombination aus Lerntempo und Hilfestellung konnte in drei Fällen verzeichnet werden, z.B.: „Wenn Ihr Wörter nicht wisst, schaut im Buch S. 165. [...] So, allez, commencez. Mindestens vier Fragen müsst ihr machen.“ (BoFr.19/20_8.5a+b). Nach Neigung oder Thema wurde in zehn Fällen differenziert. Hierbei gilt es zu beachten, dass sechs dieser Aktivitäten im Rahmen

eines Stationenlernens frei gewählt werden konnten (BoSp.19/20_10.2), weshalb alle sechs Aktivitäten als differenziert gezählt wurden.

Tabelle 2: Binnendifferenzierung mit dem Lehrwerk

Art der Differenzierung	Französisch	Spanisch	Gesamt
Lerntempo	17	2	19
Hilfestellung	4	9	13
Neigungs- und Themenwahl	1	9	10
Kombination: Lerntempo und Hilfestellung	1	2	3
Gesamt	23	22	45

Zwar ist die Menge an differenzierten Aktivitäten in beiden Sprachen ungefähr gleich, die Art der Differenzierung unterscheidet sich je nach Fremdsprache jedoch gravierend, wobei alle Lehrkräfte zielgleich arbeiten. So werden in Französisch 17 Aktivitäten, in Spanisch hingegen nur zwei Aktivitäten nach Lerntempo differenziert. Auch im Bereich der Hilfestellungen und der Neigungs- und Themenwahl unterscheiden sich die Sprachen grundlegend. Neun Aktivitäten werden im Spanischunterricht mit Hilfestellungen durch die Lehrkraft versehen, vier im Französischunterricht. Nur im Falle einer Aktivität findet in Französisch eine Differenzierung hinsichtlich der Neigungs- und Themenwahl statt, während in Spanisch neun Aktivitäten neigungs- und themendifferenziert sind. Hierbei ist – wie bereits erwähnt – jedoch zu beachten, dass sechs der neigungs- und themendifferenzierten Aktivitäten in derselben Spanischunterrichtseinheit stattfanden und somit als typisch für diese eine Lehrkraft zu interpretieren sind:

„Hoy vamos a trabajar con el autocontrol del cuaderno de actividades. Abrid la p. 42–43. [...] Die Ergebnisse / Lösungen findet ihr hinten im AH [=Arbeitsheft]. Die Lösungen, die nicht hinten stehen, hängen vorne am Smartboard. [Hinweis der/s Studierenden: Die SuS dürfen sich Aufgaben aus dem Autocontrol selbst aussuchen. [...]]“ (BoSp.19/20_10.2).

Die Lernenden können in diesem Fall – in einem sehr eng gesteckten Rahmen (sechs Aktivitäten stehen zur Auswahl) – die Lerninhalte selbst bestimmen. Da es sich hierbei um eine Überprüfung des bisher Gelernten handelt, kann auch diese Differenzierung nicht als zieldifferent bezeichnet werden.

4.3 Feedback zu lehrwerkbasierteren Aktivitäten

Für die Analyse des Feedbacks wurde eine Substichprobe von $n = 406$ gezogen, wobei die folgenden Kriterien angelegt wurden: Aktivitäten, die zwar in der beobachteten Unterrichtsstunde begonnen, aber nicht beendet wurden (14 Aktivitä-

ten) sowie Aktivitäten, die der Vorbereitung einer anderen Aktivität dienen (z. B. das Sammeln von Stichworten, auf dem basierend in der Folge ein kurzer Text geschrieben wurde, 21 Aktivitäten), wurden nicht berücksichtigt, sodass sich eine Substichprobe von $n = 406$ ergibt. Aktivitäten, die im Plenum bearbeitet wurden und auf die im Zuge dessen direkt Feedback erfolgte, werden der Kategorie ‚zeitnah durch Lehrkraft‘ zugeordnet.

In der hier vorliegenden Substichprobe ($n = 406$ Aktivitäten) konnten drei verschiedene Arten der Rückmeldung zu Aktivitäten der Lernenden aus dem Lehrwerk beobachtet werden: a) zeitnahes Feedback durch die Lehrkraft, b) zeitversetztes Feedback durch die Lehrkraft und c) Autoevaluation. In 25,5 % der Fälle (30,0 % Französisch, 20,2 % Spanisch) erfolgte in der jeweiligen beobachteten Unterrichtseinheit keine Rückmeldung.

Passend zu der starken Fokussierung auf die Arbeit im Plenum (20,5 %) bzw. die Kombination aus Einzelarbeit und Plenum (26,7 %) erfolgen die Rückmeldungen zu den Aktivitäten in 68,5 % direkt im Anschluss an die Aktivität durch die Lehrkraft (67,0 % Französisch, 68,0 % Spanisch). Ein zeitversetztes Feedback, ebenfalls durch die Lehrkraft (z. B. Korrektur von Hausaufgaben oder von in vorherigen Unterrichtseinheiten begonnenen Aktivitäten) war nur in 4,7 % der Fälle (2,5 % Französisch, 6,9 % Spanisch) zu beobachten. Aktivitäten, zu denen eine Autoevaluation durchgeführt wurde, bei der Lernende ihre Ergebnisse selbst z. B. anhand von vorgegebenen Lösungen korrigierten, machen lediglich 2,5 % der Stichprobe aus (0,5 % Französisch, 4,4 % Spanisch).

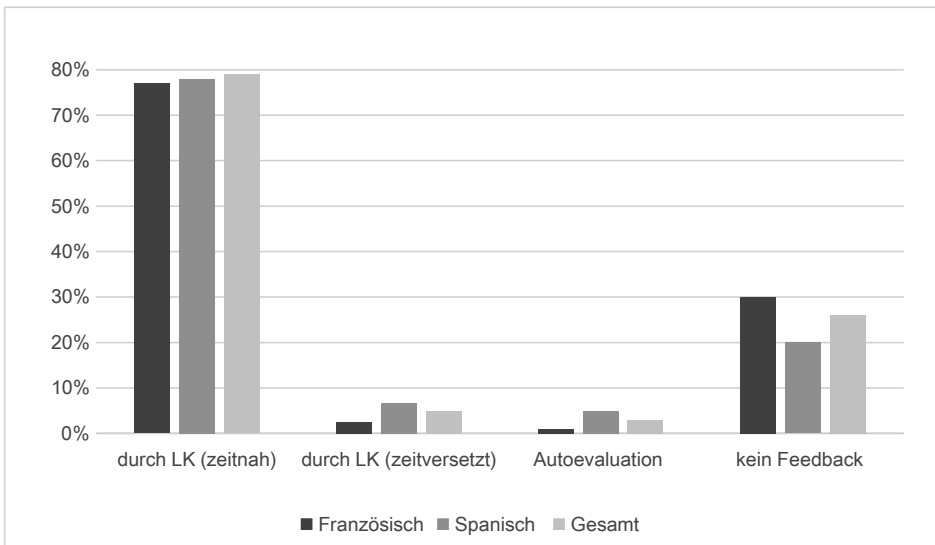


Abbildung 3: Art des Feedbacks zu lehrwerkbasierten Aktivitäten

Das Geben von Rückmeldungen zu lehrwerkbasierten Aktivitäten spielt in der hier vorgestellten Stichprobe also eine zentrale Rolle und erfolgt insgesamt in 95 % der Fälle.

5 Diskussion der Ergebnisse und Limitationen

Die hier vorgestellten Ergebnisse deuten auf einen prinzipiell stark durch die Lehrpersonen gesteuerten Wissensaufbau in lehrwerkbasierten Unterrichtsphasen hin, der sich durch viel Einzel- und Plenumsarbeit auszeichnet. Durch die Fokussierung der Beobachtung auf den Einsatz des Lehrwerks, geht aus der hier vorliegenden Stichprobe nicht hervor, inwieweit der Unterricht mit dem Lehrwerk durch kollaborative Unterrichtsphasen ohne Lehrwerk ergänzt wird. Zudem wurden isolierte Einzeleinheiten beobachtet, sodass keine Aussagen hinsichtlich einer Lehrwerknutzung in aufeinanderfolgenden Unterrichtseinheiten getroffen werden können. Die hier präsentierten Ergebnisse sind unter der Prämisse zu verstehen, dass es durch Schwerpunktsetzungen der verschiedenen Lehrkräfte innerhalb der Beobachtungen – eine Lehrkraft macht mindestens 0,8 %, maximal 7 % der Stichprobe aus – zu Verzerrungen der Daten kommen kann.

Das in allen in der Stichprobe genutzten Lehrwerken inkludierte Angebot binnendifferenzierter Aktivitäten wird nur selten genutzt (10,3 %, siehe auch Korell, 2024, S. 267), was unter Umständen auf die Wahrnehmung der Lehrkräfte hinsichtlich des Differenzierungsangebots in Lehrwerken zurückgeführt werden kann. Korell (2024, S. 231) stellt diesbezüglich fest, dass 38,1 % der 103 von ihr via Fragebogen befragten Spanischlehrkräfte mit dem Differenzierungsangebot in Lehrwerken unzufrieden seien. Gleichzeitig verweist sie darauf, dass 16 Lehrkräfte einige Lehrwerkkomponenten (z. B. das Arbeitsheft oder internetbasierte Tools) nicht kennen bzw. sich nicht damit auseinandergesetzt haben, wodurch die Differenzierungsmöglichkeiten des Lehrwerks zum Teil nicht umfassend beurteilt werden können (Korell, 2024, S. 239, 252). Diese Erkenntnis wird darüber hinaus von der Einschätzung der Lehrkräfte hinsichtlich des Potentials der Verzahnung von Lehrwerkkomponenten oder -elementen für die Differenzierung, das nur von 1,3 % der Befragten als gegeben beurteilt wird (Korell, 2024, S. 241), gestützt. Zusätzlich formulieren die Lehrkräfte ein Unbehagen bzw. „eine wahrgenommene Grundproblematik bei der materialbezogenen Umsetzung“ (Korell, 2024, S. 266) von Differenzierungsangeboten im Unterricht, was mit der Wahl leichter Aufgabenformate auch durch leistungsstarke Lernende und einem Gefühl von „Schlecht-Sein“ leistungsschwächerer Lernender, die leichtere Aufgabenformate nutzen müssen, begründet wird. Der von Korell festgestellte Rückgriff der Lehrkräfte auf „Formate, die weniger offensichtlich und offensiv sowie teils den Lernenden auch nicht bewusst sind (z. B. Hilfe-Seiten im Lehrbuch oder durch Symbole gekennzeichneten Schwierigkeitsgrad der Aufgabe und/oder Übung)“ (2024, S. 267), konnte auch in diesem Beitrag festgestellt werden. Es gilt allerdings zu bedenken, dass es sich bei dieser Art der Differenzierung um Einzelfälle handelt

(im Datensatz dieser Studie 16 Fälle, in Korells Studie basiert diese Erkenntnis auf sieben – zusätzlich zur Fragebogenstudie ($n = 103$) – videografierten Unterrichtseinheiten und dazugehörigen Interviews, wobei die soziale Erwünschtheit bei der Vorbereitung dieser Stunden nicht außer Acht zu lassen ist).

Indem Differenzierung und Adaptivität in lehrwerkbasierten Unterrichtsphasen eine untergeordnete Rolle spielen, fehlt den im MAIN-TEACH-Modell dargestellten Unterrichtsdimensionen Klassenführung, sozio-emotionale Unterstützung, Unterstützung der aktiven Beteiligung (mittlere Ebene) und kognitive Aktivierung, Unterstützung des Übens, Formatives Assessment, Auswahl und Thematisierung von Inhalten und Fachmethoden (oberste Ebene) die erforderliche Basis (Charalambous & Praetorius, 2020; Praetorius & Gräsel, 2021, Ufer & Praetorius, 2022, Lindl et al., in diesem Band). Die Schwerpunkte hinsichtlich der einbezogenen Differenzlinien variieren je nach Fach stark (z. B. 17-mal Berücksichtigung der Differenzlinie „Lerntempo“ im Französischunterricht vs. zweimal Spanischunterricht; Hilfestellungen: viermal im Französischunterricht vs. neunmal im Spanischunterricht). Um ausgehend von diesen Beobachtungen den fächerbedingten Unterschieden auf den Grund zu gehen, bedarf es einer vertieften Beschäftigung mit dem Thema, z. B. anhand einer Befragung von Lehrkräften hinsichtlich der didaktischen Begründung der Differenzierungsmaßnahmen sowie der qualitativen Analyse der differenzierten Aktivitäten.

Auch wenn das durch die Lehrwerke bereitgestellte Material noch andere Formen der Nutzung ermöglicht und zum Teil sogar nahelegt (alle Lehrwerke schlagen auch kollaborative Sozialformen vor und verfügen über – zumindest als solche bezeichnete – Lernaufgaben), findet die fremdsprachliche Interaktion in erster Linie im Plenum statt (mehr als 50%)⁴: Zum Teil ändern Lehrkräfte vom Lehrwerk als kollaborativ gedachte Aktivitäten zudem in Einzel- oder Plenumsarbeit. Obwohl in diesen Sozialformen diverse Modalitäten der reproduktiven Sprachanwendung (z. B. Chorsprechen, Abschreiben) denkbar sind (vgl. hierzu in Bezug auf Plenumsarbeit Haß, 2016, S. 337), ist die individuelle Kommunikationszeit der einzelnen Lernenden in der Gesamtstichprobe und im Kontext der Lehrwerksarbeit als gering einzuschätzen. Auch die Prüfung einer Substichprobe, die nur dritte bis fünfte Lernjahre umfasst, ergab, dass der lehrwerkbasierte Unterricht eher instruktionsorientiert und stark durch die Lehrkräfte gelenkt gestaltet wird. Die Aktivität und insbesondere der Sprechanteil der Lernenden fallen demnach auch in fortgeschrittenen Lernjahren in dieser Stichprobe relativ gering aus. Da sich die hier analysierte Stichprobe jedoch hauptsächlich auf Unterrichtseinheiten der siebten Klasse im ersten Lernjahr bezieht (über 50%), bietet sie keinen repräsentativen Überblick über alle Lernjahre.

Die starke Führung durch die Lehrkräfte kann vor allem in niedrigen Lernjahren, wenn die Schüler:innen noch an ein bestimmtes sprachliches Niveau herangeführt werden müssen, hilfreich sein. Sie kann darüber hinaus zu einem Gefühl

4 Korell (2024, S. 467) stellt im Kontext von Spanischunterricht diesbezüglich fest, dass die Einstellungen und Kenntnisse der jeweiligen Lehrenden maßgeblich für die Wahl der Sozialform sind und nicht die Vorgaben des Lehrwerks.

von Sicherheit auf Seiten der Lernenden wie Lehrenden führen und dadurch ein erfolgreiches Lernen unterstützen (Haß & Kiemweg, 2013, S. 27–28). Im Kontext von kooperativen Lernformen wie vor allem Gruppenarbeit werden zudem die Aspekte der Aufmerksamkeitsfähigkeit der Lernenden sowie das Niveau der Sprachproduktion und -verarbeitung zum Beispiel von Spanischlehrenden als nur eingeschränkt gewinnbringend betrachtet (Korell, 2024, S. 399). Für den Fremdsprachenunterricht ist insbesondere die sprachliche Vorbildfunktion der Lehrenden, die in Plenumsphasen gegeben ist, hervorzuheben (Haß, 2016, S. 337). Zum Aufbau einer kommunikativen Kompetenz bedarf es jedoch auch Unterrichtsphasen, in denen reproduzierbares Wissen „sowohl durch gezieltes sprachorientiertes Üben (*focus on language*) als auch besonders durch Phasen inhaltsorientierter Kommunikation (*focus on communication*) in die Anwendung überführt [Hervorhebungen im Original]“ (Haß, 2016, S. 337) wird. Dies kann in Plenums- und Einzelarbeit jedoch nur bedingt stattfinden. Ein abwechslungsreicher Einsatz verschiedener Sozialformen, wie Haß (2016, S. 339) ihn vorschlägt und der eine „aktive Mitwirkung von allen Schüler:innen“ bzw. ein „Lernumfeld, das produktives Verhalten fördert“ (Praetorius et al., 2020, S. 417–421; Charalambous & Praetorius, 2020; Lindl et al., in diesem Band), kann demnach nicht verzeichnet werden. Dies lässt auf wenig interaktive und kommunikative Lernsettings mit dem Lehrwerk schließen. Diese sowie eine möglichst lange Aktivität möglichst vieler Lernender wären aber für eine optimale Entwicklung sprachlicher Fertigkeiten notwendig (vgl. Dimension V Unterstützung des Lernens aller Schüler:innen; Praetorius et al., 2020, S. 417–421; Charalambous & Praetorius, 2020; Lindl et al., in diesem Band; Wilden, 2021, S. 217).

Kooperative Lernformen bzw. ein ausgewogener Wechsel von Methoden und Sozialformen, die es Schüler:innen ermöglichen, an ihr Vorwissen anzuküpfen und für den eigenen Lernprozess Verantwortung zu übernehmen, gewährleisten „Lernerfolge, erhöhte Leistungsbereitschaft und damit die Möglichkeit zu kumulativem Lernen“ (Steffens & Haenisch, 2019, S. 296). Der vermehrte und einseitige Einsatz von Einzel- und Plenumsarbeit kann hingegen eine angemessene Berücksichtigung heterogener individueller Lernvoraussetzungen verhindern (Haß, 2016, S. 337). Inwieweit sich die Handlungsmuster in den Plenums- und Einzelarbeitsphasen unterscheiden (z. B. fragenentwickelnde Vorgehensweise, Referate, Erzählen von Geschichten usw.), kann dem Datenmaterial nicht entnommen werden. Auch für die Überprüfung der Anwendung von Inszenierungstechniken (z. B. Fragen beantworten lassen, Vor- und Nachsprechen usw.) ist weitere Forschung, z. B. anhand von Videostudien, nötig.

Feedback als Teil der IV Dimension Formatives Assessment (Praetorius et al., 2020, S. 415–416; für die Fremdsprachen siehe Wilden, 2021, S. 217; Charalambous & Praetorius, 2020; Lindl et al., in diesem Band) ist in lehrwerkbasierter Unterrichtsphasen stark durch die Lehrenden gesteuert und nimmt eine zentrale Rolle im Unterrichtsgeschehen ein. Obwohl Lehrwerke auch die Möglichkeit zur Selbstevaluation (z. B. durch das Vorhandensein von Lösungen) bereithalten, wurde auf 68,5% der beobachteten Aktivitäten zeitnah und durch die Lehrenden Feedback

gegeben. Letztere fungieren entsprechend nicht als Organisator:innen bzw. Moderator:innen (z. B. Klieme et al., 2006, S. 131–132), sondern nehmen scheinbar eine eher traditionell geprägte Rolle als Wissensvermittler:innen ein. Durch die häufig frontal gestaltete Ergebnissicherung wird die Entwicklung einer positiven Abhängigkeit der Lernenden (Bonnet & Hericks, 2020, S. 101) z. B. in Partner:innen- oder Gruppenarbeiten (in dieser Stichprobe immerhin 23,7% der Aktivitäten mit dem Lehrwerk) unmöglich. Da keine Datentriangulation stattgefunden hat wie beispielsweise durch eine Befragung der Lehrkräfte zu ihrem Lehrwerkeinsatz, können keine Aussagen zu den Gründen für die Vorgehensweise der Lehrkräfte gemacht werden.

6 Fazit und Ausblick

Die Qualität des Unterrichts ist eng mit der Auswahl und Anwendung der Lehr- bzw. Lerninhalte verknüpft, sodass den Auswahl- und Gestaltungsprozessen der Lehrkräfte eine entscheidende Rolle hinsichtlich der Unterrichtsqualität zukommt: „Von ihrem Funktionsverständnis hängt maßgeblich ab, wie sie das Lehrwerk (in der Verwendung) begreif[en] und welche Aufgabe(n) sie sich selbst bei der Interaktion mit dem Material zuschreib[en]“ (Korell, 2024, S. 467). Indem sie Inhalte und Material aus Lehrwerken auswählen, methodisch aufbereiten und an die Bedürfnisse der Lerngruppe anpassen, bestimmen sie die den Lernenden dargebotenen Lern-Lehr-Settings. Entsprechend kommt der Erforschung des Lehrwerkeinsatzes neben der Analyse der Qualität der Lehrwerke selbst eine zentrale Rolle mit Blick auf Unterrichtsqualität zu und ein möglichst unverstellter Blick in den tatsächlichen Unterrichtsalltag der Lehrkräfte unter Einsatz des Lehrwerks kann Antworten auf die Frage nach Unterrichtsqualität im schulischen Fremdsprachenunterricht der Sekundarstufe I bieten (siehe Kaskadenmodell nach Krauss et al., 2020, Säulen 3 und 4; ausführlich Lindl et al., in diesem Band).

Mit Blick auf die hier schwerpunktartig gewählten Dimensionen von Unterrichtsqualität (III Unterstützung des Übens, IV formatives Assessment, V Unterstützung des Lernens aller Schüler:innen) zeigt sich, dass einerseits das Angebot der Lehrwerke nicht vollends ausgeschöpft wird und in diesen Bereichen wenig von den Lehrkräften initiierte Prozesse integriert werden, die beim kooperativen, selbstbestimmten Lernen oder auch bei der Förderung der Autonomie unterstützen. In diesem Zusammenhang ist die plenumslastige, starke Steuerung des Unterrichts seitens der Lehrkräfte hervorzuheben, wodurch eine hohe Strukturierung in enger Anlehnung an das vom Lehrwerk vorgelegte Material oder Aufgabenarrangement zu beobachten ist.

Für die weitere Auswertung der erhobenen Daten ergeben sich aus den hier vorgestellten Ergebnissen folgende Fragen: Welche Arten von Aktivitäten wählen die Lehrkräfte zur Bearbeitung im Unterricht aus dem Lehrwerk aus? Lässt sich daraus ein Muster aus „Lieblingsaktivitäten“ ableiten? Welche Arten von Aktivitäten werden mit dem Lehrwerk wie durchgeführt (z. B. kognitive Aktivierung)?

Gibt es einen Zusammenhang zwischen der Art der Aktivität und der gewählten Sozialform? Wie wird das Feedback zu den Aktivitäten aus dem Lehrwerk gestaltet? Inwiefern wird Feedback als Grundlage für die Ausrichtung des weiteren Unterrichts genutzt? Diesen und weiteren Fragen wird im Kontext der qualitativen Auswertung der Daten nachgegangen.

Literatur

Lehrwerke

- ¡Apúntate! 1. *Método de español. Spanisch als 2. Fremdsprache.* (2008). Berlin: Cornelsen.
 ¡Apúntate! 1. *Nueva edición. Spanisch als 2. Fremdsprache.* (2016). Berlin: Cornelsen.
 ¡Apúntate! 4. *Nueva edición. Spanisch als 2. Fremdsprache.* (2019). Berlin: Cornelsen.
 À toi! 3. *Lehrwerk für den Französischunterricht.* (2017). Berlin: Cornelsen.
 Découvertes 1. *Série jaune.* (2012). Stuttgart: Klett.
 Horizons. (2017). Stuttgart: Klett.
 ¿Qué pasa? 4. *Nueva edición.* (2019). Braunschweig: Westermann.

Sekundärliteratur

- Allen, H. (2008). Textbook materials and foreign language teaching: Perspectives from the classroom. *NECTFL Review*, 62, 5–28.
- Blume, O.-M. & Nieweler, A. (2018). Raus mit der Sprache! Mündlichkeit stärken im Französischunterricht. *Der fremdsprachliche Unterricht Französisch*, 152, 2–8.
- Bohnsteffen, M. (2011). Englischlehrwerke und ihre unterrichtliche Verwendung – Ergebnisse einer nicht repräsentativen Umfrage. *Fremdsprachen Lehren und Lernen*, 40(2), 120–135.
- Bonnet, A. & Hericks, U. (2020). *Kooperatives Lernen im Englischunterricht. Empirische Studien zur (Un-)möglichkeit fremdsprachlicher Bildung in der Prüfungsschule.* Tübingen: Narr.
- Caspari, D. (2010). Lehrwerksunterricht = Frontalunterricht? Mehr Schülerbeteiligung bei der Arbeit mit dem Lehrbuch! *Praxis Fremdsprachenunterricht*, 1, 3–6.
- Caspari, D. & Holzbrecher, A. (2016). Individualisierung und Differenzierung im kompetenzorientierten Französischunterricht. In L. Küster (Hrsg.), *Individualisierung im Französischunterricht. Mit digitalen Medien differenzierend unterrichten* (S. 7–37). Seelze: Klett & Kallmeyer.
- Charalambous, C. Y. & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, 67, 100894. <https://doi.org/10.1016/j.stueduc.2020.100894>
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin/Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Franke, M. (in Vorbereitung). *Der Lehrwerkeinsatz im Französisch- und Spanischunterricht der Sekundarstufe I: Eine qualitative Studie.*

- Fuchs, E. & Bock, A. (2018). Introduction. In E. Fuchs & A. Bock (Hrsg.), *The Palgrave Handbook of Textbook Studies*. (S. 1–9). New York: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-53142-1_1
- Fuchs, E., Niehaus, I. & Stoletzki, A. (2014). *Das Schulbuch in der Forschung. Analysen und Empfehlungen für die Bildungspraxis*. Göttingen: V&R unipress. <https://doi.org/10.14220/9783737003858>
- Gräsel, C. (2010). Lehren und Lernen mit Schulbüchern – Beispiele aus der Unterrichtsforschung. In Fuchs, E. & Kahlert, J. & Sandfuchs, U. (Hrsg.), *Schulbuch konkret. Kontexte – Produktion – Unterricht* (S. 137–148). Bad Heilbrunn: Klinkhardt.
- Harlan, K. (2000). Foreign Language Textbooks in the Classroom: Bridging the Gap Between Second Language Acquisition Theory and Pedagogy. *Honors Projects*, 5, 1–23.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Haß, F. (2016). Sozialformen im Überblick. In E. Burwitz-Melzer, G. Mehlhorn, C. Riemer, K.-R. Bausch & H.-J. Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht* (S. 335–339). Tübingen: A. Francke Verlag.
- Haß, F. & Kiemweg, W. (2013). *I can make it! Englischunterricht für Schülerinnen und Schüler mit Lernschwierigkeiten*. Stuttgart: Klett/Kallmayer.
- Hoffmann, S. (2010). Heterogenität und Differenzierung. In W. Hallet & F. G. Königs (Hrsg.), *Handbuch Fremdsprachendidaktik* (S. 160–164). Stuttgart: Klett.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projektes „Pythagoras“. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DGF-Schwerpunktprogramms* (S. 127–146). Münster: Waxmann.
- Korell, J. L. (2024). *Das Lehrwerk im Fremdsprachenunterricht der Sekundarstufe I: Eine Mixed-Methods-Studie zu subjektiven Sicht- und Verwendungsweisen von Spanischlehrkräften*. Tübingen: Narr Francke Attempto. <https://doi.org/10.24053/9783381112623>
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N. & Blum, W. (2020). Competence as a continuum in the COACTIV study: The “cascade model”. *ZDM – Mathematics Education*, 52, 311–327. <https://doi.org/10.1007/s11858-020-01151-z>
- Kuckartz, U. (2012). *Qualitative Inhaltsanalyse: Methode, Praxis, Computerunterstützung*. Weinheim: Beltz-Juventa.
- Kurtz, J. (2010). Zum Umgang mit dem Lehrwerk im Englischunterricht. In E. Fuchs, J. Kahlert & U. Sandfuchs (Hrsg.), *Schulbuch konkret. Kontexte, Produktion, Unterricht* (S. 149–163). Bad Heilbrunn: Klinkhardt.
- Kurtz, J. (2011). Zur Einführung in den Themenschwerpunkt. *Fremdsprachen Lehren und Lernen*, 40, 3–14.
- Martin, B. & Hanington, B. (2019). *Universal Methods of Design. 125 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Beverly: Rockport Publishers.
- Martinez, H. (2005). Lehrwerke auf dem Prüfstand. Ist jedes Lehrbuch auch ein Lernbuch? *Französisch heute*, 2, 110–122.
- Nausica Marcos, M. (2015). Textbook consumption in the classroom: Analyzing a classroom corpus. *Procedia – Social and Behavioral Sciences*, 198, 309–319. <https://doi.org/10.1016/j.sbspro.2015.07.449>
- Nieweler, A. (2000). Sprachenlernen mit dem Lehrwerk – Thesen zu Lehrbucharbeit im Fremdsprachenunterricht. In R. Fery & V. Raddatz (Hrsg.), *Lehrwerke und ihre Alternativen* (S. 13–19). Frankfurt am Main: Lang.

- Praetorius, A. K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48, 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Przyborski, A. & Wohlrab-Sahr, M. (2014). *Qualitative Sozialforschung*. München: Oldenbourg Verlag. <https://doi.org/10.1524/9783486719550>
- Rösler, D. (2004). *E-Learning Fremdsprachen. Eine kritische Einführung*. Tübingen: Stauffenburg Verlag.
- Sandfuchs, U. (2010). Schulbücher und Unterrichtsqualität – historische und aktuelle Reflexionen. In E. Fuchs, J. Kahlert & U. Sandfuchs (Hrsg.), *Schulbuch konkret. Kontexte – Produktion – Unterricht* (S. 11–24). Bad Heilbrunn: Klinkhardt.
- Schulz, W. (1969). Differenzierung an Gesamtschulen. In A. Rang & W. Schulz (Hrsg.), *Die differenzierte Gesamtschule. Zur Diskussion einer neuen Schulform* (S. 181–204). München: Piper.
- Sikorová, Z. & Červenková, I. (2014). Styles of textbook use. *The New Educational Review*, 35(1), 112–122. <https://doi.org/10.15804/tner.14.35.1.09>
- Steffens, U. & Haenisch, H. (2019). Gut unterrichten – erfolgreich lernen. Versuch einer Forschungsbilanz unter unterrichtspraktischer Perspektive. In U. Steffens & R. Messner (Hrsg.), *Unterrichtsqualität: Konzepte und Bilanzen gelingenden Lehrens und Lernens* (S. 279–312). Münster: Waxmann.
- Stöber, G. (2010). Schulbuchzulassung in Deutschland. Grundlagen, Verfahrensweisen und Diskussionen. *Eckert.Beiträge*, 3. <http://repository.gei.de/handle/11428/92>, Zugriff: 30.08.2023.
- Thaler, E. (2011). Die Zukunft des Lehrwerks – Das Lehrwerk der Zukunft. *Fremdsprachen Lehren und Lernen*, 2, 15–30.
- Thaler, E. (2014). Lehrwerk-Kompetenz. *Praxis Fremdsprachenunterricht Basisheft*, 1, 4–8.
- Tönshoff, W. (2004). Binnendifferenzierung im lernerorientierten Fremdsprachenunterricht. *Deutsch als Fremdsprache*, 4, 227–231. <https://doi.org/10.37307/j.2198-2430.2004.04.07>
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Wapenhans, H. (2018). Förderung der mündlichen Kommunikationsfähigkeit – Anforderungen an die schulische Praxis und die universitäre fachdidaktische Ausbildung von Russischlehrkräften. In A. Bergmann, O. Caspers & Stadler, W. (Hrsg.), *Didaktisch der slawischen Sprachen – Beiträge zum 1. Arbeitskreis Berlin (12.–14.9.2016)* (S. 255–269). Innsbruck: Innsbruck University Press. <https://doi.org/10.15203/3187-11-5-14>
- Wiater, W. (2005). Lehrplan und Schulbuch – Reflexionen über zwei Instrumente des Staates zur Steuerung des Bildungswesens. In E. Matthes & C. Heinze (Hrsg.), *Das Schulbuch zwischen Lehrplan und Unterrichtspraxis* (S. 41–63). Bad Heilbrunn: Klinkhardt.
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49, 211–219. <https://doi.org/10.1007/s42010-021-00105-y>

Maria Hofrichter, Alfred Lindl, Mara Rader und Annika Krämer

Was macht guten Fremdsprachenunterricht aus?

Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern Englisch und Latein

Abstract (der Herausgeber:innen)

The article “What constitutes good foreign language teaching? Theoretical considerations and first results of an interdisciplinary online survey in the subjects English and Latin” by Maria Hofrichter, Alfred Lindl, Mara Rader and Annika Krämer first presents the status quo of research on instructional quality in English as a Foreign Language and Latin. This provides the basis for the two central questions addressed in their contribution: “What characteristics do pre-service and in-service teachers consider to be good teaching in English and Latin?” and “How do pre-service and in-service teachers, from their practice-oriented perspective, evaluate the characteristics of good teaching that are prominently discussed in educational science and subject-matter didactics?”. These questions are analysed on the basis of an online survey comprising 122 pre-service and in-service teachers (teacher trainees and fully-qualified teachers) from both subjects. Finally, the results are presented and discussed in detail.

1 Forschungsanlass und -hintergrund

1.1 Fachspezifische Lehr- und Lernprozesse im Fokus

Fachbezogene Lehr- und Lernprozesse gelten im bildungswissenschaftlichen Diskurs als größte Einflussfaktoren auf Lernerfolge von Schülerinnen und Schülern (z. B. Reusser & Pauli, 2021; Seidel & Shavelson, 2007). Dennoch werden sie in Modellierungen von Unterrichtsqualität wie beispielsweise den drei Basisdimensionen Klassenführung, konstruktive Unterstützung und kognitive Aktivierung (z. B. Klieme & Rakoczy, 2008), die im deutschsprachigen Raum große Beachtung erfahren, kaum berücksichtigt. Obwohl diese ursprünglich im Rahmen von Videoanalysen zu Mathematikunterricht entwickelt (Klieme et al., 2001) und bisher vorwiegend in diesem sowie im naturwissenschaftlichen Fachkontext überprüft wurden (z. B. Kunter & Ewald, 2016; Rakoczy et al., 2019), nahm man lange Zeit deren universelle Übertragbarkeit und Gültigkeit unter anderem hinsichtlich aller Schularten, Jahrgangsstufen, Fächer etc. an (z. B. Charalambous & Praetorius,

Hofrichter, M., Lindl, A., Rader, M. & Krämer, A. (2024). Was macht guten Fremdsprachenunterricht aus? Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern Englisch und Latein. In V. Lohe, A. Lindl & P. Kirchhoff (Hrsg.), *Unterrichtsqualität in schulischen Fremdsprachen. Theoretische Ansätze und empirische Ergebnisse aus den Fachdidaktiken* (S. 109–138). Münster: Waxmann. <https://doi.org/10.31244/9783830999201.05>

2018; Klieme et al., 2006). Gerade dies stellen aktuelle Forschungsansätze jedoch nicht zuletzt wegen einer – insbesondere interdisziplinär – inkonsistenten empirischen Befundlage zunehmend infrage (zusammenfassend Praetorius et al., 2018; Rothland, 2024). Beispielsweise konnten die drei Basisdimensionen in der DESI-Studie gerade für den Sprachunterricht in Deutsch und Englisch nicht repliziert werden (Helmke, 2022; Helmke & Klieme, 2008).

Um die jeweils wirksamen Lehr- und Lernprozesse in diesen und weiteren Fächern zu identifizieren, wird daher in den letzten Jahren ein stärkerer Einbezug verschiedener Fachdidaktiken sowie vor allem fachkultureller bzw. -spezifischer Aspekte postuliert (z. B. Begrich et al., 2023; Charalambous & Praetorius, 2018; Jentsch & Kirchhoff, in diesem Band; Lindmeier & Heinze, 2020; Lipowsky & Bleck, 2019; Praetorius et al., 2020a). Dieser Appell richtet sich insbesondere an gesellschaftswissenschaftliche sowie (fremd-)sprachliche Schulfächer, die nicht selten eine gewisse Distanz zur empirischen Bildungsforschung wahren (z. B. Helmke, 2022; Praetorius & Gräsel, 2021; Wilden, 2021). Neben den Forschungsaufgaben einer differenzierten fachspezifischen Auslegung, Konkretisierung und Validierung bislang generisch konzipierter Merkmalsdimensionen von Unterricht stehen verschiedene Fragestellungen im Zentrum des Erkenntnisinteresses, deren Klärung einer fachbezogenen Perspektive bedarf: Was macht qualitativ hochwertigen Unterricht im Fach aus? Wird dieser ausschließlich durch fachunabhängige Kriterien determiniert oder existieren darüber hinaus fachspezifische Prozessmerkmale? Welche von diesen erscheinen – gemäß Berliners (2005) Begriffsverständnis – nur normativ als *gut* bzw. welche erweisen sich hinsichtlich kognitiver und affektiver Zielkriterien von Unterricht auch tatsächlich als *effektiv* (vgl. hierzu auch Drechsel & Schindler, 2019; Kunter & Trautwein, 2013)?

Das bisherige Untersuchungsinteresse an derartigen Fragen, der entsprechende wissenschaftliche Kenntnisstand und die jeweilige Forschungslage hierzu variieren von Schulfach zu Schulfach wie auch zwischen den zugehörigen Fachdidaktiken zum Teil erheblich (vgl. Jentsch & Kirchhoff, in diesem Band; Lindl et al., in diesem Band). Nachstehend wird deshalb zunächst der Forschungshintergrund hinsichtlich der schulischen Fremdsprachen Englisch und Latein aufgearbeitet, um darauf fußend die Ansatzpunkte, Fragestellungen und ersten Untersuchungsschritte dieses Beitrags darzulegen.

1.2 Forschungsstand in den Fremdsprachen Englisch und Latein

1.2.1 Englisch

Eine umfassende Dokumentation des aktuellen Stands und künftiger Entwicklungslinien der fachdidaktischen Forschung zu Unterrichtsqualität im Fach Englisch wurde jüngst von Wilden (2021) vorgelegt. Daran anknüpfend und mit Blick auf alle im deutschsprachigen Raum unterrichteten modernen Fremdsprachen, aber nur auf eine einzelne Dimension widmet sich Guttke (2023) in einem sys-

tematischen Review den Einflussfaktoren, Erfassungsmethoden sowie Effekten kognitiver Aktivierung. Weitere Übersichten zu Aspekten von Unterrichtsqualität in der englischen Fachdidaktik finden sich in mehreren Beiträgen dieses Bandes (z. B. Guttke & Porsch; Folkerts; Lindl et al.). Vor diesem Hintergrund lässt sich feststellen, dass die Beschäftigung mit diesem Thema ein noch recht junges Forschungsfeld von bislang überraschend geringer Aktivität ist (Wilden, 2021). Vorherrschend seien vor allem „Studien im qualitativ-interpretativen Spektrum“ (Wilden, 2021, S. 215), die erste Hinweise auf guten Englischunterricht bieten, etwa durch die Untersuchung von Überzeugungen von Lehrkräften hinsichtlich eines effektiven Fremdsprachenunterrichts (Rossa, 2017) oder bezüglich des Einsatzes literarischer Texte (Gardemann, 2021), der Förderung der Lesekompetenz (Gerlach & Lüke, 2020), des Sprechens (Miede, 2019) und des interkulturellen Lernens (Göbel, 2007).

Quantitativ ausgerichtete empirische Forschungsansätze sind hingegen vergleichsweise selten (Wilden, 2021): In einem quasiexperimentellen Längsschnittdesign untersuchte Maag Merki (2011), inwiefern dezentrale bzw. zentrale Prüfungsformate das Niveau kognitiver Aktivierung und konstruktiver Unterstützung in gymnasialen Abschlussklassen beeinflussen, und erfasste hierfür beide Merkmale in drei aufeinanderfolgenden Jahren aus der Perspektive von Schülerinnen und Schülern. Während Kersten et al. (2018) mit dem Teacher Input Observation Scheme (TIOS, vgl. Tab. 1) ein Beobachtungsinstrument (weiter-)entwickelten und validierten, das auf die Qualität und den zielsprachlichen Input des Unterrichts von Lehrkräften in der Primarstufe abhebt, analysierten Wilden und Porsch im Rahmen ihrer Querschnittsstudie TEPS (Teaching English in Primary Schools; vgl. Tab. 1) das Lese- und Hörverstehen von Grundschulkindern am Ende der vierten Jahrgangsstufe in Abhängigkeit von Unterrichtsqualitätsmerkmalen (Wilden & Porsch, 2019; Wilden et al., 2020). Erstmals wurden diese Teilkompetenzen in Bezug auf den Englischunterricht in der DESI-Studie erfasst, deren umfassendes Längsschnittdesign differenzierte Auswertungen zu Lehr- und Lernprozessen bzw. -bedingungen sowie den Erwerb sprachlicher Kompetenzen im Englischunterricht der Sekundarstufe I ermöglichte (Helmke & Klieme, 2008; Helmke et al., 2008). Bis heute birgt DESI sekundäranalytisches Potenzial, etwa in Bezug auf die Zusammenhänge zwischen der Lehrkräftekooperation, Merkmalen der Unterrichtsqualität und Leistung sowie Lernmotivation der Jugendlichen (Hochweber et al., 2012) oder im Hinblick auf die Bedeutsamkeit des Umgangs mit Fehlern im Englischunterricht für den Lernerfolg von Schülerinnen und Schülern (Käfer, 2022). Die DESI-Studie ebenso wie Maag Merkis Längsschnittstudie stellen jedoch weiterhin die einzigen Untersuchungen dar, die sich quantitativ-empirisch mit Qualitätsmerkmalen von Englischunterricht in der Sekundarstufe beschäftigen. Ausführliche Informationen zu den vier erwähnten Studien enthält Tabelle 1, in der diese kontrastiv zusammen- und gegenübergestellt werden, da sie zentrale Anknüpfungspunkte und Anregungen für die vorliegende, ebenfalls quantitativ angelegte Untersuchung bieten.

Tabelle 1: Übersicht über vier quantitativ-empirische Studien zu Merkmalen fachbezogener Unterrichtsqualität im Schulfach Englisch

	DESI (Deutsch-Englisch-Schülerleistungen-International)	<i>Empirical Effects on Math and English Teaching in Secondary Schools</i>	TIOS (Teacher Input Observation Scheme)	TEPS (Teaching English in Primary Schools)
Kontext				
<i>Was war das Design, die Laufzeit, der theoretische Hintergrund der Studie?</i>	<ul style="list-style-type: none"> – bundesweite Längsschnittstudie zur Erfassung der sprachlichen Leistung von Lernenden der 9. Jgst. in Deutsch und Englisch (als erste Fremdsprache) – Schuljahr 2003/2004 (1. Messzeitpunkt im Sept./Okt. 2003, 2. Messzeitpunkt im Mai/Juni 2004) – theoretischer Bezug zum Angebots-Nutzungs-Modell (Fend, 1981) und zum Rahmenmodell der Unterrichtswirksamkeit (Helmke, 2004) 	<ul style="list-style-type: none"> – Längsschnittstudie zu Auswirkungen der Einführung des Zentralabiturs auf die Unterrichtsqualität von Leistungskursen in den Fächern Englisch und Mathematik im letzten Jahr der gymnasialen Oberstufe – Datenerhebung 2007 bis 2009 in Bremen und Hessen – theoretischer Bezug zu Studien angloamerikanischer Länder, welche Effekte landesweiter Abiturprüfungen untersuchen (Abrams, 2007; Amrein & Berliner, 2002) 	<ul style="list-style-type: none"> – Längs-/Querschnittstudie SMILE (Studies on Multilingualism in Language Education) zur Entwicklung eines quantitativen Beobachtungsinstrumentes – Datenerhebung 2015 bis 2018 in Niedersachsen – Entwicklung des TIOS anhand von Befunden der Zweitspracherwerbsforschung im Rahmen der Instructed Second Language Acquisition (ISLA) 	<ul style="list-style-type: none"> – Querschnittstudie zur Untersuchung rezeptiver Englischkenntnisse (Lese- und Hörverstehenskompetenz) von Lernenden am Ende der 4. Jgst. Datenerhebung 2017 in Nordrhein-Westfalen (Beginn des Englischunterrichts in der 1. Jahrgangsstufe) und Niedersachsen (Beginn des Englischunterrichts in der 3. Jahrgangsstufe) – theoretischer Bezug zum Angebots-Nutzungs-Modell der Wirkweise von Unterricht (Helmke, 2015)
Zentrale Publikation(en)				
<i>Wer steht hinter der Studie?</i>	<ul style="list-style-type: none"> – DESI-Konsortium (2008), darin spezifisch für Englisch Helmke & Klieme (2008) bzw. Helmke et al. (2008) – Sekundäranalysen: Hochweber et al. (2012)¹; Käfer (2022)² 	<ul style="list-style-type: none"> – Theoretischer Hintergrund: Maag Merki (2012) – Hauptstudie: Maag Merki (2011) 	<ul style="list-style-type: none"> – Theoretischer Hintergrund: Kersten (2021) – Beobachtungsinstrument und Manual: Kersten et al. (2018) – Befunde aus SMILE: Kersten et al. (2019), Kersten (2020)¹ – Weitere Befunde: Kersten (2022)² 	<ul style="list-style-type: none"> – Pilotstudie: Wilden & Porsch (2019) – Hauptstudie: Wilden et al. (2020)¹ – Weitere Befunde: Porsch & Wilden (2021)²
Stichprobe				
<i>Wer wurde untersucht?</i>	<ul style="list-style-type: none"> – N = 10.639 (1. MZP) bzw. N = 10.632 (2. MZP) Lernende der 9. Jgst. aller Schularten (ohne Förderschule) 	<ul style="list-style-type: none"> – N = 685 Schüler:innen aus Englisch-Grundkursen – N = 2.573 Schüler:innen aus Englisch-Leistungskursen 	<ul style="list-style-type: none"> – N = 218 Lernende der 3. und 4. Jgst. – N = 17 Lehrkräfte; darunter 9 „Immersionslehrkräfte“, die an Schulen mit bilingualen Sachfachunterrichts unterrichten 	<ul style="list-style-type: none"> – N = 776 Schüler:innen der 4. Jgst.

	<i>DESI (Deutsch-Englisch-Schülerleistungen-International)</i>	<i>Empirical Effects on Math and English Teaching in Secondary Schools</i>	<i>TIOS (Teacher Input Observation Scheme)</i>	<i>TEPS (Teaching English in Primary Schools)</i>
(zentrale) Fragestellung und Ziele <i>Welcher Fragestellung und Zielsetzung widmete sich die Studie?</i>	<ul style="list-style-type: none"> - zentrale Fragestellung: Welche Merkmale von Lehrkräften und des Fachunterrichts spielen für den Kompetenzzuwachs von Lernenden eine Rolle? - Ziele: <ul style="list-style-type: none"> - Erklärungsansätze für unterschiedliche Leistungsniveaus bzw. Lernzuwächse im Verlauf der 9. Jgst. - Revision von Curricula, Lehrtexten und Unterrichtsmaterialien, der Lehrkräftebildung und Unterrichtsentwicklung 	<ul style="list-style-type: none"> - zentrale Fragestellung: Inwiefern verändert sich in den Leistungskursen, die im Abiturjahr 2007 dezentral, im Abiturjahr 2008 hingegen zentral geprüft wurden, die Unterrichtsqualität in der gymnasialen Oberstufe? Inwiefern lässt sich diese Veränderung auf die Einführung zentraler Abiturprüfungen zurückführen? - Ziele: <ul style="list-style-type: none"> - Untersuchung der Auswirkungen der Einführung des Zentralabiturs auf die Unterrichtsqualität von Grund- und Leistungskursen im Fach Englisch 	<ul style="list-style-type: none"> - zentrale Fragestellung: Welche Unterrichtstechniken werden im Englischunterricht der Grundschule eingesetzt und welche dieser Techniken zeigen positive Effekte auf den L2-Erwerb der Schüler:innen? - Ziele: <ul style="list-style-type: none"> - Erfassung der Unterrichtsqualität über die Anzahl und Art der eingesetzten Unterrichtstechniken - Untersuchung des Effekts des L2-Inputs in Regel- und Intensivprogrammen auf den rezeptiven L2 Wortschatz- und Grammatikerwerb von Schüler:innen und Schülern 	<ul style="list-style-type: none"> - zentrale Fragestellung: Unterscheiden sich die rezeptiven Englischkenntnisse der Schüler:innen am Ende der 4. Jgst. bezüglich der Anzahl an Lernjahren und inwiefern beeinflussen die Qualifikation der Lehrkräfte und die Qualität des Unterrichts die rezeptiven Englischkenntnisse der Schüler:innen? - Ziele: <ul style="list-style-type: none"> - Erfassung der rezeptiven Englischkenntnisse im Grundschulkontext und Vergleich der Schüler:innen mit früherem und späterem Lernbeginn unter Berücksichtigung von Unterrichtsqualität und Lehrkräftequalifikation
Fokussierte Konstrukte <i>Was wurde untersucht?</i>	<ul style="list-style-type: none"> - Rezeptive und produktive Leistungen im Englischen (Hörverstehen, Leseverstehen, Schreiben, Textrekonstruktion, Sprachbewusstheit, Sprechen) sowie interkulturelle Kompetenz - 15 Skalen, die unterschiedliche Facetten der Unterrichtsqualität aus Schüler:innensicht abdecken (z. B. Strukturiertheit, Unterrichtsklima, positive Fehlerkultur, Verständlichkeit, Motivierung, Wichtigkeit von Kommunikation, Störungsfreiheit, Verwendung der deutschen Sprache) 	<ul style="list-style-type: none"> - Qualität von Unterricht bei Englischlehrkräften in der gymnasialen Oberstufe - Messung kognitiv aktivierenden und unterstützenden Unterrichts (Klieme, 2006) durch vier Skalen 	<ul style="list-style-type: none"> - L2-Input und Qualität von Unterricht bei Grundschullehrkräften im Rahmen des kognitiv-interaktionistischen Ansatzes (Long, 2015) - Weiterentwicklung des Input Quality Observation Scheme (Weitz, 2015), das für den Kita-Kontext entwickelt wurde - 41 Unterrichtstechniken mit fünf Skalen zu Dimensionen von Unterrichtsqualität (kognitiv anregenden Aufgaben/Aktivitäten, verbalem Input, non-verbalem Input, Unterstützung des Outputs der Lernenden und Fehlerkorrektur) 	<ul style="list-style-type: none"> - rezeptive Englischkenntnisse (Hör- und Leseverstehen) - Adaption von zwei Skalen zu Klassenführung und unterstützendem Klima eines Befragungsinstruments zur Erhebung von Unterrichtsqualität im Sachunterricht der Grundschule aus Schüler:innen- und Lehrkräfteperspektive für den Englischunterricht (Fauth et al., 2014) - theoriebasierte Neuentwicklung einer Skala für „kommunikativ-kognitive Aktivierung“ (Thaler, 2014)

	DESI (Deutsch-Englisch-Schülerleistungen-International)	<i>Empirical Effects on Math and English Teaching in Secondary Schools</i>	TIOS (Teacher Input Observation Scheme)	TEPS (Teaching English in Primary Schools)
Instrumente und Auswertungssätze	<ul style="list-style-type: none"> - Tests mit Matrix-Design zur Erfassung der Leistungen von Lernenden - Fragebögen für Lernende (Selbstinschätzung, Lernverhalten, -motivation, -kontext, Einstellung zur Sprache), Lehrkräfte, Schulleitungen, Eltern - Videografie von zwei Unterrichtsstunden pro Klasse die jeweils ein sprachlernorientiertes und ein interkulturelles Thema behandeln - IRT-Modellierung und Mehrebenenstrukturgleichungsanalysen 	<ul style="list-style-type: none"> - Fragebögen für Lernende (18 Items zu den vier Bereichen Elaboration, Motivationsfähigkeit der Lehrkraft, Autonomie und Kompetenzerleben) 	<ul style="list-style-type: none"> - Tests zur Erfassung der Leistungen von Grundschulkindern basierend auf dem British Picture Vocabulary Scale (Dunn et al., 2009) und dem ELIAS Grammar Test II (Kersten et al., 2012) - Videografie von neun Unterrichtsstunden pro Klasse und Analyse anhand des Beobachtungsinstruments TIOS - Mehrebenenstrukturgleichungsanalysen 	<ul style="list-style-type: none"> - Tests zur Erfassung der Leistungen von Grundschulkindern - Fragebögen für Lernende, um Unterrichtsqualität zu bewerten (19 Items zu den Bereichen kommunikativ-kognitive Aktivierung, Klassenführung und unterstützendes Klima) - Fragebögen für Lehrkräfte, um deren Qualifikation für den Englischunterricht zu ermitteln - IRT-Modellierung und Mehrebenenstrukturgleichungsanalysen
Zentrale Ergebnisse	<ul style="list-style-type: none"> - In Klassen mit besonders hohem Zuwachs in der Hörverstehensleistung kommen Schüler:innen häufig zum Sprechen, warten Lehrkräfte mind. drei Sekunden auf deren Antworten, wird bei Unterrichtsgesprächen überwiegend Englisch gesprochen, erhalten Lernende Gelegenheit zur Selbstkorrektur ihrer Fehler, gibt es vergleichsweise wenige „Ein-Wort-Sätze“, kommt es zu Dialogen, die über einfache Frage-Antwort-Sequenzen hinausgehen 	<ul style="list-style-type: none"> - Mehrebenenstrukturgleichungsanalysen 	<ul style="list-style-type: none"> - Die Qualität des fremdsprachlichen Unterrichts sagt signifikant die L2-Kompetenzen der Schüler:innen im Wortschatz und Grammatikverständnis vorher.¹ - Es gibt signifikante Unterschiede zwischen regulären und bilingualen Grundschulen in Bezug auf die L2-Kontaktdauer und -intensität, Einstellung von Lehrkräften und Unterrichtstechniken (kognitive Stimulierung, aktive Wissenskonstruktion, L2-Inputqualität), die Effekte auf den L2-Erwerb sowie die kognitive Entwicklung und die Einstellungen der Lernenden zeigen.¹ 	<ul style="list-style-type: none"> - Es gibt keine Belege dafür, dass Lernende mit frühem Lernbeginn (1. Jgst.) in ihren rezeptiven Englischfähigkeiten am Ende der Grundschulzeit signifikant besser abschneiden als Lernende mit späterem Lernbeginn (3. Jgst.), auch nicht unter Berücksichtigung von Unterrichtsqualität und der Qualifikation der Lehrkräfte.¹
<i>Welche Befunde ergaben sich in Bezug auf Unterrichtsqualität?</i>	<ul style="list-style-type: none"> - In Bezug auf Unterrichtsqualität konnten signifikante Veränderungen nur in den Leistungskursen festgestellt werden, die alle eine positive Entwicklung (in allen vier Bereichen) der Unterrichtsqualität aufwiesen. - Die Unterrichtsqualität verbesserte sich vor allem von 2007 bis 2008 mit der Umstellung von einem dezentralen auf ein landesweites Prüfungssystem, ohne weitere signifikante Veränderungen der Unterrichtsqualität von 2008 bis 2009, als das Prüfungssystem unverändert blieb. 			

<p>DESI (Deutsch-Englisch-Schülerleistungen-International)</p> <p><i>Empirical Effects on Math and English Teaching in Secondary Schools</i></p>	<p>TIOS (Teacher Input Observation Scheme)</p>	<p>TEPS (Teaching English in Primary Schools)</p>
<ul style="list-style-type: none"> - Erfolgreicher Unterricht (bezogen auf Hörverstehens-Zuwachs) ist gekennzeichnet durch intensive Zeitzunutzung, ausgeprägte Aufgabenorientierung, Störungsfreiheit, ein positives Fehlerklima und hohes Schüler:innenengagement. - Die Lehrkraftkooperation im Fachkollegium Englisch hängt positiv mit zwei von drei Basisdimensionen der Unterrichtsqualität (Strukturiertheit, Anforderungen) zusammen.¹ - Die individuelle Leistung und Lernmotivation der Lernenden fallen höher aus, wenn sie das Verhalten der Englischlehrkraft in Fehlersituationen als unterstützend, Fehler insgesamt als nützlich und die Einstellung der Lehrkraft als fehlerfreundlich wahrnehmen.² 	<ul style="list-style-type: none"> - Bilinguale Lernprogramme wirken sich positiv auf den rezeptiven L2-Wortschatzerwerb und die phonologische Bewusstheit der Lernenden aus.² 	<ul style="list-style-type: none"> - Unterstützendes Klima, d. h. das Erleben von Autonomie oder ein konstruktiver Umgang mit Fehlern, korreliert signifikant und positiv mit rezeptiven Englischfertigkeiten von Lernenden.¹ - Die Wahrnehmung der Unterrichtsqualität durch die Lehrkräfte hängt positiv mit ihrem fachspezifischen Enthusiasmus und ihrer zielsprachlichen Kompetenz zusammen.²

Anm.: Informationen, die sich bei DESI, TIOS bzw. TEPS nur in einer bestimmten Publikation finden, werden jeweils durch hochgestellte Ziffern am Ende der Aussage markiert.

1.2.2 Latein

Während zum Schulfach Englisch einige theoretische Überlegungen und empirische Untersuchungen zu fachspezifischer Unterrichtsqualität vorliegen, präsentiert sich das diesbezügliche Forschungsfeld in der altsprachlichen Didaktik als äußerst defizitär (Lindl, 2024). Dies macht eine systematische Literaturrecherche in einschlägigen didaktischen Fachzeitschriften und Datenbanken¹ zu zentralen Begriffen wie guter, effektiver oder qualitativvoller Unterricht, Unterrichtsqualität, aber auch hierzu assoziierten Konzepten wie Klassenführung, kognitiver Aktivierung, Unterstützung etc. deutlich (vgl. Tab. 2). Zu deren überschaubarer absoluter Trefferanzahl gehören zwei Veröffentlichungen, die bildungswissenschaftliche Unterrichtsqualitätsmodelle rezipieren, wobei Wittich (2015) auf die zehn Merkmale guten Unterrichts von Helmke (2015) und Riecke-Baulecke et al. (2021) auf die drei Basisdimensionen nach Klieme et al. (2001) rekurrieren. Zwar erschließen sie damit diese theoretischen Ansätze für die altsprachliche Didaktik, diskutieren aber vor einem fachspezifischen Hintergrund weder deren Übertragbarkeit und Gültigkeit noch hinterfragen sie die zugehörigen bildungswissenschaftlichen Ergebnisse oder nehmen auf den aktuellen Forschungsstand bzw. gegenwärtige Tendenzen Bezug.

Eine zweite Gruppe bildet eine Reihe heuristischer Arbeiten (vgl. Tab. 2), die sich – ähnlich wie im Fach Englisch – mit „der (Weiter-)Entwicklung und Reflexion von unterrichtspraktischen Konzepten“ (Wilden, 2021, S. 215) in den alten Sprachen befassen. Diese enthalten teils anschauliche fachbezogene Deutungsvarianten generischer Prozessmerkmale, teils wertvolle Hinweise auf einzelne fachspezifische Aspekte des Lehrens und Lernens, ohne darin allerdings explizit terminologische oder konzeptionelle Elemente des bildungswissenschaftlichen Unterrichtsqualitätsdiskurses zu berücksichtigen. Dennoch werden bei genauerer Betrachtung nicht nur gewisse Anknüpfungspunkte erkennbar, sondern die bisherigen Ansätze der altsprachlichen Fachdidaktik erscheinen auch mit bestimmten Dimensionen des theoretischen Rahmenmodells des vorliegenden Kapitels bzw. Bandes (Lindl et al., in diesem Band), also dem erweiterten Syntheseframework nach Praetorius et al. (2020b), kompatibel, wie die Zusammenstellung in Tabelle 2 verdeutlicht.

1 Hierzu wurde eine vollständige Schlagwortsuche in der umfangreichen fachdidaktischen Grundlagenbibliothek der Altsprachlichen Didaktik an der Humboldt Universität zu Berlin (<https://www.klassphil.hu-berlin.de/de/fachgebiete/didaktik/lehrmittelsammlung-1>) wie auch in den Indizes der Fachzeitschriften *Der Altsprachliche Unterricht*, *Forum Classicum* und *Pegasus Online* durchgeführt [Stand: März 2024]. Quellenverzeichnisse relevanter Beiträge wurden zudem zur ergänzenden Rückwärtsuche genutzt.

Tabelle 2: Analytische Aufgliederung des (theoretisch-heuristischen) Forschungsstands in der altsprachlichen Fachdidaktik gemäß den Dimensionen des Syntheseframeworks von Praetorius et al. (2020b)

Dimension des Syntheseframeworks	Themenverwandte Schlagworte im fachdidaktischen Diskurs	Publikationen
Angemessene Auswahl und Thematisierung von Inhalten und Fachmethoden	Motivierung von Inhalten und Fachmethoden	Bäcker, 2010, 2019; Häger, 2017
Kognitive Aktivierung	Kognitive Aktivierung	Liebsch, 2020
Unterstützung des Übens	Üben, Hausaufgaben, Fehlerkultur	AU 2/1997; Scholz, 2019b; Steinhilber, 1986; Kuhlmann & Horstmann, 2018; AU 4/2014; AU 6/2000; Kuhlmann, 2014
Formatives Assessment	Diagnose und Leistungsbeurteilung	AU 1/2012; Scholz 2018a, 2019a; Scholz & Weber, 2011
Unterstützung des Lernens aller Schüler:innen	Differenzierung, Inklusion, Heterogenität	AU 1/2008; Müller, 2017b; Scholz, 2018b, 2019a; Scholz & Weber, 2011; Jesper, 2016, 2018; Liebsch, 2019, 2021; Müller, 2017a; Große, 2017; Kipf, 2019
Sozio-emotionale Unterstützung	Lehrkraft-Schüler:in-Interaktion	Klausnitzer, 2018
Klassenführung	Standardsituationen	AU 2/2019

Anm.: Die Abkürzung *AU Nummer/Jahreszahl* verweist auf vollständige Themenhefte der fachdidaktischen Zeitschrift *Der Altsprachliche Unterricht*.

Aus dieser systematischen Recherche und analytischen Auswertung der Literatur zu Qualitätsmerkmalen altsprachlichen Unterrichts (Tab. 2) gehen thematische Beschäftigungsschwerpunkte auf fachbezogenen Aspekten des Übens wie auch auf fachspezifischen Fragen der Differenzierung, Inklusion und Heterogenität zur Unterstützung des Lernens aller Schüler:innen hervor. Zu Teilbereichen der übrigen Dimensionen des erweiterten Syntheseframeworks finden sich vergleichsweise wenige fachdidaktische Publikationen.

Insgesamt ist somit festzuhalten, dass die bildungswissenschaftliche Forschung zu Unterrichtsqualität in der altsprachlichen Fachdidaktik zwar ansatzweise wahrgenommen wird (Riecke-Baulecke et al., 2021; Wittich, 2015). Bislang werden diesbezügliche Konzepte, Modelle und Ergebnisse aber weder theoretisch reflektiert oder unter einer fachspezifischen Perspektive diskutiert noch existieren hierzu eigene empirische Untersuchungen (Lindl, 2024).

1.3 Forschungsfragen dieses Beitrags

Das in beiden Fremdsprachen erkennbare Forschungsdefizit bezüglich Merkmalen fachspezifischer Unterrichtsqualität (Abschnitt 1.2; Lindl et al., in diesem Band) greift die interdisziplinäre Forschungsgruppe FALKO-PV (*Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung*) an der Universität Regensburg auf.² Ihr Fokus liegt auf einer Untersuchung der Zusammenhänge zwischen domänenspezifischen Lehrkraftkompetenzen und der Qualität wie auch kognitiven und affektiven Zielkriterien von Unterricht in den sechs Fächern Deutsch, Englisch, Evangelische Religion, Latein, Mathematik und Musik. Hierzu war in einigen der genannten Disziplinen – so auch in Englisch und Latein, auf die sich dieser Beitrag beschränkt – zunächst zu klären, was *guten* Unterricht aus theoretischer, fachdidaktischer und normativer Sicht ausmacht, bevor eine empirische Überprüfung der *Effektivität* dieser Merkmale überhaupt möglich ist (Berliner, 2005). Als erster Schritt wurde daher eine Onlinebefragung mit (angehenden) Lehrkräften durchgeführt, um die Übertragbarkeit und Passung bildungswissenschaftlicher Modelle zu überprüfen und diese gegebenenfalls um fachspezifische Aspekte zu erweitern oder zu ergänzen. Ein solcher Einbezug einer praxisorientierten Sicht auf Unterrichtsqualität kann, wie Freytag und Theurer (2020) sowie Rothgangel (2021) betonen, in einem kokonstruktiven Forschungsprozess Richtschnur und Impulsgeber bei der Weiterentwicklung theoretischer Ansätze sein (vgl. auch Lindl & Krämer, 2022). Im Zentrum dieses Beitrags stehen somit folgende zwei Fragestellungen:

- 1) Welche Merkmale machen nach Ansicht von (angehenden) Lehrkräften *guten* Unterricht im Fach Englisch bzw. Latein aus?
- 2) Wie bewerten (angehende) Lehrkräfte in der bildungswissenschaftlichen und fachdidaktischen Forschung prominent diskutierte Merkmale *guten* Unterrichts aus ihrer praxisorientierten Perspektive?

Während nachstehend das Studiendesign und die methodische Herangehensweise an diese Fragestellungen beschrieben werden, ist der dritte Abschnitt den Ergebnissen in beiden Fächern gewidmet. Deren Diskussion, die auch einige Unterschiede zwischen moderner und alter Fremdsprache aufzeigt, und eine Darstellung der Limitationen dieser Studie (Abschnitt 4) schließen den Beitrag ab.

2 Die Nachwuchsforschungsgruppe FALKO-PV wird im Rahmenprogramm Empirische Bildungsforschung für fünf Jahre (2021–2026) aus Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter dem Förderkennzeichen 01JG2103 gefördert. Für diese Veröffentlichung sind die Autorinnen und der Autor verantwortlich.

2 Studiendesign

Um eine praxisorientierte Perspektive auf Merkmale guten Englisch- und Lateinunterrichts zu erhalten, wurde im Frühsommer 2022 eine Onlinebefragung mit (angehenden) Lehrkräften durchgeführt. Ihre Stichprobe, Anlage und Auswertung werden im Folgenden vorgestellt.

2.1 Stichprobe und Durchführung

An der Studie nahmen insgesamt 122 Personen aus Deutschland (101) und Österreich (21) teil, von denen 116 ein Lehramtstudium mit dem Fach Englisch oder Latein absolvieren bzw. absolviert haben. Die nachstehenden Analysen beziehen sich nur auf die Daten dieser (angehenden) Lehrkräfte, wobei bei der Ergebnisdarstellung aus Gründen der Vereinfachung und, weil sich deren Antworten nicht substantiell unterscheiden, nicht zwischen Lehramtsstudierenden und Lehrkräften differenziert wird. Von den 55 Personen (76,4% weiblich) im Fach Englisch verfügen 42 mindestens über einen berufsqualifizierenden akademischen Abschluss (z.B. Staatsexamen, Master), unter den 61 in Latein (57,3% weiblich) sind dies 52. Das mittlere Alter der (angehenden) Englischlehrkräfte beträgt 32,31 Jahre ($SD = 9,28$), der Lateinlehrkräfte 38,93 ($SD = 10,80$). Ihre durchschnittliche Unterrichtserfahrung in Englisch ist 6,53 Jahre ($SD = 6,78$), in Latein 13,35 Jahre ($SD = 9,91$).

Die Einladungslinks zur Onlinebefragung wurden im deutschsprachigen Raum über persönliche Kontakte, fachspezifische Netzwerke und Newsletter sowie Fortbildungsangebote und Praxiszeitschriften (Lindl & Krämer, 2022) verbreitet und die Erhebung an privaten Endgeräten durchgeführt, sodass eine Gelegenheitsstichprobe vorliegt. Die Teilnahme unterlag den Bestimmungen der Europäischen Datenschutzgrundverordnung, war freiwillig, konnte an jeder Stelle der Befragung beendet werden und war nicht an finanzielle Anreize oder Studien- bzw. Fortbildungsleistungen geknüpft.

2.2 Instrument: Onlinefragebogen

Die Befragung wurde mittels der Software LimeSurvey auf einem institutseigenen Server durchgeführt und umfasste den obigen Forschungsfragen entsprechend zwei Kernabschnitte. Nach einleitenden Informationen zu Anlass und Hintergrund der Studie erhielten die (angehenden) Lehrkräfte die Instruktion, in drei Schlagworten oder Stichpunkten zu beschreiben, was ihrer Meinung und Erfahrung nach guten Englisch- bzw. Lateinunterricht ausmacht. Ihre Antworten konnten sie in drei Freitextfelder (mit einem Limit von je 100 Zeichen) eintragen. Dies diente im ersten Abschnitt dazu, möglichst intuitive, unbeeinflusste Aussagen zu

Merkmale zu erhalten, die nach Ansicht (angehender) Lehrkräfte mit gutem Unterricht in Englisch oder Latein assoziiert sind.

Im zweiten Abschnitt sollten die Teilnehmer:innen hingegen die Relevanz verschiedener randomisiert präsentierter Kriterien auf einer fünfstufigen Likert-Skala (0 unwichtig, ..., 4 wichtig) in Bezug auf die Qualität von Englisch- bzw. Lateinunterricht bewerten. Die zugehörige Instruktion lautete: Schätzen Sie bitte ein, wie wichtig nachstehende Aspekte Ihrer Erfahrung nach für guten Unterricht im Fach Englisch bzw. Latein sind. Die vorgegebenen Merkmale entsprachen dabei einerseits den 21 generischen Subdimensionen von Unterrichtsqualität nach dem erweiterten Syntheseframework von Praetorius et al. (2020b; vgl. Tab. 2; Lindl et al., in diesem Band). Andererseits wurden literaturbasiert (vgl. Abschnitt 1.2) in Englisch 16 bzw. in Latein 15 zusätzliche Aspekte formuliert, die aus fachdidaktischer Sicht für den Unterricht in der jeweiligen Fremdsprache als bedeutsam sowie insbesondere als fachspezifisch angenommen werden (z. B. Schaffung hoher Sprechanteile bzw. Förderung einer Faszination für das Fach).

Daran schloss eine Zuordnungsaufgabe zur Konstruktvalidierung (auf die hier nicht eingegangen wird) sowie gegen Ende der Befragung die Möglichkeit an, in einem Freitextfeld weitere Merkmale guten Unterrichts zu ergänzen, falls die vorausgehenden Fragestellungen einen bislang noch nicht geäußerten bzw. angeführten Aspekt bei den (angehenden) Lehrkräften evoziert haben sollten. Die Befragung endete mit einer Abfrage von drei Zielkriterien qualitätsvollen Englisch- bzw. Lateinunterrichts (hier ebenfalls nicht berichtet), Optionen zur künftigen Kontaktierung und einer Danksagung.

2.3 Kodierung der Freitextantworten

Zu den Freitextantworten aus dem ersten Abschnitt wurden die wenigen Angaben aus dem Ergänzungsfeld gegen Ende des Fragebogens hinzugenommen und diese gemeinsam anhand eines in allen Fächern einheitlich angelegten, deduktiven Kodierschemas ausgewertet, dessen Grundlage das erweiterte Syntheseframework nach Praetorius et al. (2020b) bildet. Nach der ersten Kodierung aller Antworten wurde dieses Schema bei Bedarf induktiv ergänzt. Dies geschah einerseits unter Rückgriff auf die jeweilige fachspezifische Literatur (Abschnitt 1.2), andererseits in enger Abstimmung mit allen an FALKO-PV beteiligten Fächern, um die fachübergreifende Konsistenz des Kodierschemas sicherzustellen. So wurde nicht nur eine differenzierte Einteilung von unpassenden Nennungen, d. h. Antworten, die sich nicht auf ein tiefenstrukturelles Prozessmerkmal von Unterricht beziehen, vorgenommen, sondern wurden auch fach(gruppen)spezifische Merkmalskategorien ergänzt und in einem zweiten Kodierdurchgang bestimmt (vgl. für Beispiele Abschnitte 3.1 und Tab. 4).

Indem bei diesem in allen Fächern wiederum das identisch weiterentwickelte Manual angelegt wurde, sind die Kodierungen zwischen allen Disziplinen, insbesondere Englisch und Latein, vergleichbar. Zudem wurde die zweite Kodierungs-

runde durch jeweils zwei fachkundige Rater:innen (wissenschaftliche Mitarbeitende mit einem universitären Lehramtsabschluss – 1. Staatsexamen bzw. Master of Education – im jeweiligen Studienfach) unabhängig voneinander durchgeführt, was in Englisch (84 %) und in Latein (78 %) zu einer guten prozentualen Übereinstimmung der Kategorisierungen als Indikator für die Auswertungsobjektivität führte. Fälle, in denen die Einordnung der beiden Rater:innen voneinander abwich, wurden gemeinsam gesichtet und diskutiert, bis eine Einigung erzielt werden konnte, was angesichts der stichpunktartigen Kürze und Prägnanz mancher Freitextantworten gelegentlich nicht einfach war.

3 Ergebnisse

Nachstehend werden die Ergebnisse der Onlinebefragung dargestellt, wobei zunächst auf die Freitextantworten und anschließend auf die Relevanzeinschätzungen eingegangen wird.

3.1 Freitextantworten

Bei Betrachtung der Freitextantworten in Tabelle 3 wird deutlich, dass in Englisch mehr als deren Hälfte, in Latein knapp deren Hälfte unzutreffend sind, d. h., nicht auf die Tiefenstruktur bzw. Prozessmerkmale des Unterrichts Bezug nehmen (z. B. „überzeugende, engagierte Lehrkraft“, „vielfältiger Medieneinsatz“, „Übersetzungstraining“). Dieser erklärungsbedürftige Befund veranlasste zu einer kritischen Revision des Kodiermanuals und einer Neukodierung aller Antworten dieser Kategorie, um einen differenzierten Eindruck zu erhalten, was (angehende) Lehrkräfte als vermeintliche Merkmale guten Unterrichts ansehen. In erster Stelle stehen hier Kompetenzmerkmale der Lehrkraft (10 %) gefolgt von Aspekten der unterrichtlichen Oberflächenstruktur wie Methodenvielfalt (8 %) oder einseitigen Schwerpunktsetzungen auf bestimmten Unterrichtspraktiken (4 %, z. B. „Übersetzen“). Diese stellen aber offensichtlich keine Merkmale des Unterrichts dar oder betreffen zumindest nicht dessen Prozessebene (vgl. Abschnitte 1.1 bzw. 2.3; Lindl et al., in diesem Band).

Generische Merkmale guten Unterrichts

Von den Antworten, die gemäß den sieben Dimensionen und 21 Subdimensionen des Syntheseframeworks von Praetorius et al. (2020b) geordnet wurden, entfallen die meisten in beiden Fremdsprachen (je 16 %) auf die erste Dimension einer angemessenen Auswahl und Thematisierung von Inhalten und Fachmethoden. Während diesbezüglich in Englisch die Subdimension Motivierung von Inhalten (10 %) im Vordergrund steht, beziehen sich in Latein die meisten Antworten auf die Strukturierung der thematisierten Inhalte (11 %). In Englisch findet sich unter 177 Freitextantworten keine einzige Äußerung, die sich klar der Dimension ko-

gnitiver Aktivierung zuordnen ließ, in Latein unter 204 lediglich vier (2%). Aspekte der Unterstützung des Übens nennen Lehrkräfte beider Fächer annähernd gleich häufig (je 3%), auch wenn in Englisch der Fokus tendenziell auf einem konstruktiven Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern (3%), in Latein auf einer Unterstützung bei der Festigung von Lern- und Anwendungsprozessen (3%) liegt. Facetten einer prozessbegleitenden Leistungsbeurteilung spielen in den Antworten der (angehenden) Lehrkräfte in beiden Fremdsprachen eine eher untergeordnete Rolle (je 1%).

Tabelle 3: Absolute Anzahlen und relative Anteile (in Prozent) der Freitextantworten pro generischer Kategorie guten Unterrichts

		Englisch	Latein
	Anzahl der Freitextantworten	177	204
	Kein Prozessmerkmal des Angebots	99	93
		(56 %)	(46 %)
I	Angemessene Auswahl, Thematisierung und Sequenzierung von Inhalten und Fachmethoden	28	32
		(16 %)	(16 %)
I.1	Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten	3	5
		(2 %)	(2 %)
I.2	Motivierung von Inhalten	18	3
		(10 %)	(1 %)
I.3	Strukturierung der thematisierten Inhalte	7	22
		(4 %)	(11 %)
I.4	Akkuratheit und Korrektheit der thematisierten Inhalte	0	2
			(1 %)
II	Kognitive Aktivierung	0	4
			(2 %)
II.1	Auswahl fachlich gehaltvoller und auf das kognitive Niveau der Schüler:innen abgestimmter Aufgaben	0	1
			(1 %)
II.2	Einsatz fachlich gehaltvoller Aufgaben	0	0
II.3	Unterstützung der kognitiven Aktivität der Schüler:innen	0	2
			(1 %)
II.4	Unterstützung des metakognitiven Lernens der Schüler:innen anhand kognitiv aktivierender Aufgaben	0	1
			(1 %)
III	Unterstützung des Übens	6	7
		(3 %)	(3 %)
III.1	Unterstützung bei der Festigung von Lern- und Anwendungsprozessen	1	7
		(1 %)	(3 %)
III.2	Konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern beim Üben	5	0
		(3 %)	
IV	Formatives Assessment	2	1
		(1 %)	(1 %)
IV.1	Klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen	0	0
IV.2	Regelmäßige Überprüfung des Verständnisses der Schüler:innen	0	0
IV.3	Qualitativ hochwertiges Feedback an die Schüler:innen	2	1
		(1 %)	(1 %)
IV.4	Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts	0	0

		Englisch	Latein
V	Unterstützung des Lernens aller Schüler:innen	3 (2%)	6 (3%)
V.1	Bereitstellung eines Lernumfelds, das produktives Verhalten fördert	0	3 (1%)
V.2	Differenzierung und Adaptivität	0	3 (1%)
V.3	Förderung der aktiven Mitwirkung aller Schüler:innen	3 (2%)	0
VI	Sozio-emotionale Unterstützung	6 (3%)	5 (2%)
VI.1	Beziehung zwischen Lehrperson sowie Schülerinnen und Schülern	6 (3%)	5 (2%)
VI.2	Beziehung der Schüler:innen untereinander	0	0
VII	Klassenführung	2 (1%)	1 (1%)
VII.1	Verhaltensmanagement	1 (1%)	1 (1%)
VII.2	Zeitmanagement	1 (1%)	0

Zur Unterstützung des Lernens aller Schüler:innen finden sich nur wenige Erwähnungen (2% bzw. 3%), und zwar in Englisch mit Fokus auf eine Förderung einer aktiven Mitwirkung aller Schüler:innen (2%), in Latein auf die Bereitstellung eines Lernumfelds, das produktives Verhalten fördert wie auch Differenzierung und Adaptivität. Weitere Nennungen, welche die Dimension der sozio-emotionalen Unterstützung betreffen, berühren ausschließlich die Beziehung zwischen Lehrperson und Lernenden (3% bzw. 2%), nicht das Verhältnis der Schüler:innen untereinander. Antworten, die Aspekte der Klassenführung aufgreifen, finden sich in beiden Fächern kaum (je 1%). Dabei mögen diese verhältnismäßig geringen Anzahlen gerade bei tendenziell fachunabhängigen Merkmalen auch dadurch bedingt sein, dass die Instruktion explizit auf guten Englisch- bzw. Lateinunterricht und damit auf fachspezifische Aspekte gerichtet war, die im Folgenden betrachtet werden.

Fachspezifische Merkmale guten Unterrichts

Gemäß dem explorativen Studienansatz wurden aus den Freitextantworten (angehender) Lehrkräfte zusätzlich zu den sieben Dimensionen (und 21 Subdimensionen) des erweiterten Syntheseframeworks (Praetorius et al., 2020b) fachspezifische Kategorien guten Englisch- bzw. Lateinunterrichts dann induktiv ergänzt, wenn diese von Lehrkräften häufiger genannt wurden, in der fachdidaktischen Literatur als Gütekriterium diskutiert werden oder einer neu genannten Kategorie von beiden Kodiererinnen bzw. Kodierern ein lernrelevantes empirisches Wirksamkeitspotenzial zugebilligt wurde. Umgekehrt dokumentiert Tabelle 4 auch die fehlende Erwähnung einiger weiterer Aspekte wie Fremdheitserfahrung und reflektierter Umgang mit (fremd-)kulturellen Schemata oder kognitiv-ästhetische Beschäftigung mit Fachinhalten (Lindl et al., 2024), deren Nennung aus bildungs-

wissenschaftlicher beziehungsweise englisch- oder lateinindidaktischer Sicht zu erwarten gewesen wäre.

Tabelle 4: Absolute Anzahlen und relative Anteile (in Prozent) der Freitextantworten pro fachspezifischer Kategorie guten Unterrichts

		Englisch	Latein
VIII	Fachspezifische Ergänzungen	32 (18 %)	55 (27 %)
VIII.1	Didaktische Reduktion unter Berücksichtigung zukünftiger fachlicher Lernschritte	0	1 (1 %)
VIII.2	Förderung einer Faszination für das Fach	–	6 (3 %)
VIII.3	Adäquate Auswahl und Berücksichtigung bedeutsamer Rezeptionsdokumente	–	1 (1 %)
VIII.4	Regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen	–	24 (12 %)
VIII.5	Kritische Reflexion über die gegenwärtige (und zukünftige) kulturelle wie auch gesellschaftliche Relevanz des Fachs	–	3 (1 %)
VIII.6	Zeitintensive, stark fokussierte und analytisch genaue Betrachtung von Fachinhalten	–	5 (2 %)
VIII.7	Evozierung kognitiver Konflikte und Förderung entdeckenden Lernens	–	1 (1 %)
VIII.8	Kognitiv-ästhetische Beschäftigung mit Fachinhalten	0	0
VIII.9	Kognitiv-kommunikative Auseinandersetzung mit Fachinhalten	0	1 (1 %)
VIII.10	Kognitiv-emotionale Betrachtung von Fachinhalten	0	1 (1 %)
VIII.11	Kognitiv-motorische Auseinandersetzung mit Fachinhalten	0	0
VIII.12	Fremdheitserfahrung und reflektierender Umgang mit (fremd-)kulturellen Schemata	0	0
VIII.13	Aktiver Gebrauch der Fremdsprache als Kommunikationsmedium	7 (4 %)	1 (1 %)
VIII.14	Verwendung der Fremdsprache als Reflexionsgegenstand	1 (1 %)	10 (5 %)
VIII.15	Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses	–	1 (1 %)
VIII.16	Anpassung der Verwendung der Fremdsprache	0	–
VIII.17	Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch	3 (2 %)	–
VIII.18	Gelegenheiten zur Selbstkorrektur	0	–
VIII.19	Interdisziplinäre Bezüge zwischen Fremdsprachen	0	–
VIII.20	Schaffung hoher Sprechanteile	21 (12 %)	–
VIII.21	Repetitives und intelligentes Üben	0	–
VIII.22	Unterstützung bei emotionalen, (inter-)kulturellen (Fremdheits-)Erfahrungen	0	–

Anm.: Der Halbgeviertstrich verweist darauf, dass dieser Aspekt im jeweiligen Fach nicht in die inhaltliche Kodierung einbezogen war.

Wie aus Tabelle 4 hervorgeht, entfallen in Englisch knapp ein Fünftel, in Latein etwas mehr als ein Viertel der Freitextantworten auf fachspezifische Merkmale guten Unterrichts. Das mit Abstand am häufigsten genannte Kriterium in Englisch ist die Schaffung hoher Sprechanteile der Schüler:innen durch ‚echte‘ kommunikative Anlässe (12%); dahinter folgt der aktive Gebrauch der Fremdsprache als Kommunikationsmedium (4%). Im Vordergrund stehen also Unterrichtsprozesse, die auf Sprachproduktion ausgerichtet sind. Guter Lateinunterricht zeichnet sich hingegen dadurch aus, dass regelmäßig Aktualitätsbezüge zur Lebenswelt der Schüler:innen geschaffen werden (12%) und die Fremdsprache als Reflexionsgegenstand verwendet wird (5%). Fachbezogene Schwerpunktsetzungen sind somit klar erkennbar.

3.2 Relevanzeinschätzungen

Inwiefern sich derartige Differenzen auch in den Relevanzeinschätzungen zu Kriterien guten Unterrichts widerspiegeln, die (angehenden) Lehrkräften zur Beurteilung vorgelegt wurden, geht aus Tabelle 5 hervor.

Generische Merkmale guten Unterrichts

Größte Bedeutung wird von (angehenden) Lehrkräften in Englisch den Dimensionen Unterstützung des Lernens aller Schüler:innen und Unterstützung des Übens zugeschrieben, wobei dies insbesondere für die Subdimensionen Beziehung zwischen Lehrperson und Lernenden bzw. konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern beim Üben gilt. Geringe Relevanz weisen aus fachspezifischem Blickwinkel formatives Assessment – hier vor allem die Subdimensionen klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen und Nutzung des Feedbacks als Grundlage für die weitere Unterrichtsausrichtung – und Klassenführung (bzgl. aller Subdimensionen) auf (Tab. 5).

In Latein wird der Unterstützung des Übens (hinsichtlich beider Teilfacetten; Tab. 5) deskriptiv die größte Bedeutsamkeit beigemessen. Die Merkmalsdimensionen kognitive Aktivierung, und zwar vor allem den Einsatz fachlich gehaltvoller Aufgaben wie auch die Unterstützung des metakognitiven Lernens der Schüler:innen anhand kognitiv aktivierender Aufgaben schätzen (angehende) Lehrkräfte als am wenigsten fachspezifisch relevant ein. Ähnlich wie in Englisch werden die Dimensionen formatives Assessment und Unterstützung des Lernens aller Schüler:innen beurteilt. In der ersten Kategorie erhalten die Facetten klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen und Nutzung des Feedbacks als Grundlage für die weitere Unterrichtsausrichtung eine geringere Zustimmung. Beim zweiten Merkmal werden insbesondere die Subdimensionen Bereitstellung eines Lernumfelds, das produktives Verhalten fördert, wie auch Differenzierung und Adaptivität niedrig eingeschätzt.

Tabelle 5: Mittelwerte (M) und Standardabweichungen (SD) der Relevanzeinschätzungen zu generischen Merkmalen guten Unterrichts

	Englisch	Latein
Gesamtanzahl antwortender Personen	51	56
I Angemessene Auswahl, Thematisierung und Sequenzierung von Inhalten und Fachmethoden	3,39 (0,60)	3,51 (0,47)
I.1 Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten	3,43 (0,81)	3,39 (0,85)
I.2 Motivierung von Inhalten	3,64 (0,60)	3,51 (0,88)
I.3 Strukturierung der thematisierten Inhalte	3,12 (0,91)	3,54 (0,66)
I.4 Akkuratheit und Korrektheit der thematisierten Inhalte	3,40 (0,81)	3,59 (0,60)
II Kognitive Aktivierung	3,35 (0,48)	3,19 (0,56)
II.1 Auswahl fachlich gehaltvoller und auf das kognitive Niveau der Schüler:innen abgestimmter Aufgaben	3,60 (0,53)	3,58 (0,63)
II.2 Einsatz fachlich gehaltvoller Aufgaben	3,02 (0,85)	2,91 (1,03)
II.3 Unterstützung der kognitiven Aktivität der Schüler:innen	3,60 (0,70)	3,39 (0,76)
II.4 Unterstützung des metakognitiven Lernens der Schüler:innen anhand kognitiv aktivierender Aufgaben	3,08 (0,68)	2,88 (1,11)
III Unterstützung des Übens	3,54 (0,51)	3,71 (0,44)
III.1 Unterstützung bei der Festigung von Lern- und Anwendungsprozessen	3,30 (0,74)	3,64 (0,62)
III.2 Konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern beim Üben	3,80 (0,49)	3,79 (0,62)
IV Formatives Assessment	3,21 (0,46)	3,18 (0,56)
IV.1 Klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen	2,96 (0,88)	2,79 (1,19)
IV.2 Regelmäßige Überprüfung des Verständnisses der Schüler:innen	3,44 (0,67)	3,59 (0,65)
IV.3 Qualitativ hochwertiges Feedback an die Schüler:innen	3,52 (0,61)	3,52 (0,66)
IV.4 Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts	2,90 (0,87)	2,82 (1,06)
V Unterstützung des Lernens aller Schüler:innen	3,35 (0,49)	2,98 (0,74)
V.1 Bereitstellung eines Lernumfelds, das produktives Verhalten fördert	3,46 (0,68)	2,88 (1,08)
V.2 Differenzierung und Adaptivität	3,08 (0,80)	2,82 (0,92)
V.3 Förderung der aktiven Mitwirkung aller Schüler:innen	3,49 (0,74)	3,27 (0,99)

		Englisch	Latein
VI	Sozio-emotionale Unterstützung	3,58 (0,43)	3,29 (0,67)
VI.1	Beziehung zwischen Lehrperson sowie Schülerinnen und Schülern	3,88 (0,33)	3,73 (0,67)
VI.2	Beziehung der Schüler:innen untereinander	3,27 (0,71)	2,84 (0,99)
VII	Klassenführung	3,08 (0,80)	3,25 (0,78)
VII.1	Verhaltensmanagement	3,06 (0,91)	3,12 (0,94)
VII.2	Zeitmanagement	3,08 (0,91)	3,38 (1,00)

Anm.: Skalierung 0 unwichtig, ..., 4 wichtig.

Bei einem direkten Vergleich der mittleren Relevanzeinschätzungen in Tabelle 5 auf Dimensionsebene tritt also in beiden Fremdsprachen die Bedeutung des Übens deutlich hervor, wohingegen formativem Assessment (und Klassenführung) eher geringe Relevanz beigemessen wird. Die deutlichsten Bewertungsunterschiede zwischen beiden Fächern finden sich hinsichtlich der Merkmalsdimensionen Unterstützung des Lernens aller Schüler:innen und sozio-emotionaler Unterstützung. Weitere Gemeinsamkeiten und Unterschiede auf Ebene der Subdimensionen sind deren durchschnittlichen Relevanzeinschätzungen in Tabelle 5 zu entnehmen.

Fachspezifische Merkmale guten Unterrichts

Die fachbezogenen Merkmale, deren tatsächliche Relevanz (angehende) Lehrkräfte einschätzen sollten, wurden aus der jeweiligen Fachliteratur und Vorüberlegungen abgeleitet. Aus Gründen der Durchführungsökonomie wurden nicht in jedem Fach dieselben bzw. alle fachspezifischen Merkmale angeführt, auch wenn diese eventuell in anderen Fächern zur Disposition standen. Zwischen den Fächern Englisch und Latein können also nur die mittleren Relevanzeinschätzungen derjenigen Merkmale verglichen werden, die auch in beiden Fremdsprachen implementiert waren (siehe hierzu Tab. 6).

Tabelle 6: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Relevanzeinschätzungen zu fachspezifischen Merkmalen guten Unterrichts

		Englisch	Latein
VIII.1	Didaktische Reduktion unter Berücksichtigung zukünftiger fachlicher Lernschritte	3,18 (0,75)	3,23 (1,08)
VIII.2	Förderung einer Faszination für das Fach	–	3,64 (0,67)
VIII.3	Adäquate Auswahl und Berücksichtigung bedeutsamer Rezeptionsdokumente	–	2,91 (0,90)
VIII.4	Regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen	–	3,62 (0,78)
VIII.5	Kritische Reflexion über die gegenwärtige (und zukünftige) kulturelle wie auch gesellschaftliche Relevanz des Fachs	–	2,84 (1,09)
VIII.6	Zeitintensive, stark fokussierte und analytisch genaue Betrachtung von Fachinhalten	–	2,48 (1,04)
VIII.7	Evozierung kognitiver Konflikte und Förderung entdeckenden Lernens	–	2,80 (1,03)
VIII.8	Kognitiv-ästhetische Beschäftigung mit Fachinhalten	2,78 (0,78)	2,73 (0,96)
VIII.9	Kognitiv-kommunikative Auseinandersetzung mit Fachinhalten	3,41 (0,73)	3,24 (0,88)
VIII.10	Kognitiv-emotionale Betrachtung von Fachinhalten	2,78 (0,81)	2,93 (0,97)
VIII.11	Kognitiv-motorische Auseinandersetzung mit Fachinhalten	2,45 (1,02)	1,98 (1,18)
VIII.12	Fremdheitserfahrung und reflektierender Umgang mit (fremd-)kulturellen Schemata	3,20 (0,86)	3,00 (1,07)
VIII.13	Aktiver Gebrauch der Fremdsprache als Kommunikationsmedium	3,84 (0,99)	1,29 (1,12)
VIII.14	Verwendung der Fremdsprache als Reflexionsgegenstand	3,18 (0,99)	2,78 (1,18)
VIII.15	Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses	3,35 (0,72)	2,83 (0,95)
VIII.16	Anpassung der Verwendung der Fremdsprache	3,18 (1,00)	–
VIII.17	Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch	3,70 (0,51)	–
VIII.18	Gelegenheiten zur Selbstkorrektur	3,08 (0,80)	–
VIII.19	Interdisziplinäre Bezüge zwischen Fremdsprachen	2,66 (0,92)	–
VIII.20	Schaffung hoher Sprechanteile	3,69 (0,58)	–
VIII.21	Repetitives und Intelligentes Üben	3,04 (0,86)	–
VIII.22	Unterstützung bei emotionalen, (inter-)kulturellen (Fremdheits-)Erfahrungen	3,22 (0,77)	–

Anm.: Skalierung 0 unwichtig, ..., 4 wichtig. Der Halbgeviertstrich verweist darauf, dass dieser Aspekt im jeweiligen Fach nicht in die Relevanzeinschätzungen einbezogen war.

Wie aus Tabelle 6 hervorgeht, wurden die fachspezifischen Ergänzungen in beiden Fächern tendenziell als relevant eingeschätzt. In Englisch wird Merkmalen, die auf hohe Sprachanteile und -qualität im Unterricht abzielen, die größte Bedeutung zugeteilt (aktiver Gebrauch der Fremdsprache als Kommunikationsmedium, Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch, Schaffung hoher Sprechanteile); von geringerer Relevanz sind eine kognitiv-ästhetische Beschäftigung bzw. eine kognitiv-emotionale Betrachtung von Fachinhalten. In Latein werden Unterrichtsprozesse, die auf eine Förderung einer Faszination für das Fach abzielen und regelmäßig Aktualitätsbezüge zur Lebenswelt der Schüler:innen schaffen, aus fachspezifischer Sicht als am wichtigsten angesehen. Aufgrund der vorherrschenden Unterrichtskultur gelten ein aktiver Gebrauch der Fremdsprache als Kommunikationsmedium und eine kognitiv-motorische Auseinandersetzung mit Fachinhalten hingegen als am wenigsten relevant.

Eine Gegenüberstellung beider Fremdsprachen dokumentiert folglich gerade hier die größten Unterschiede hinsichtlich der Relevanzeinschätzungen und stimmt darin also unter anderem auch mit der Verteilung der Freitextantworten überein. Allerdings erscheint bemerkens- und erwähnenswert, dass auch Aspekte wie die Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses sowie insbesondere die Verwendung der Fremdsprache als Reflexionsgegenstand in Englisch höher bewertet werden als in Latein, was angesichts der unterschiedlichen fachlichen Vermittlungsschwerpunkte auf Sprachproduktion und -rezeption so nicht zu erwarten wäre.

4 Diskussion, Limitationen und Ausblick

Ausgehend von einem Forschungsdefizit zu Unterrichtsqualität in den Fremdsprachen Englisch und Latein (Abschnitt 1.2) wagt die vorliegende Studie in einem explorativen Ansatz einen weiteren bzw. in Latein sogar ersten Schritt zur empirischen Eröffnung dieses Themengebiets aus fachdidaktischer Sicht und damit einen Brückenschlag zur bildungswissenschaftlichen Forschung (Wilden, 2021). Deren Konzepte und Modelle zu Unterrichtsqualität übernimmt sie nicht unreflektiert, sondern beleuchtet sie vor dem Hintergrund des jeweiligen fachdidaktischen Forschungsstands (Abschnitt 1.2) und gleicht sie mit einem praxisorientierten Verständnis von Unterrichtsqualität ab, dessen gewinnbringender Einbezug vielfach gefordert wird (z. B. Freytag & Theurer, 2020; Rothgangel, 2021). Mangels facheigener (Gesamt-)Ansätze und empirischer Studien erscheinen die Übernahme und Adaption theoretischer Grundlagen aus der Bildungswissenschaft als fundierte Ausgangspunkte unabdingbar und zunächst zumindest zweckdienlich, deren weitere Überprüfung und kritische Diskussion aber zugleich erforderlich und angezeigt.

Ihre Validierung unter der praxisorientierten Perspektive von (angehenden) Lehrkräften unterstreicht angesichts insgesamt überdurchschnittlich hoher Einschätzungen die Relevanz der generischen Merkmale auch für die Fächer Englisch

und Latein, für die aber zusätzlich auch einige der zur Beurteilung vorgelegten fachspezifischen Aspekte als bedeutsam erachtet werden. Manche von diesen wurden auch in den intuitiven Freitextantworten von (angehenden) Lehrkräften mit gutem Unterricht in ihrem Fach assoziiert und hervorgehoben. Die Auswertung dieser un gelenkten Antworten war indes in beiden Fächern besonders ergiebig, um bislang unberücksichtigte, innovative und überlegenswerte fachspezifische Merkmale von Unterrichtsqualität zu identifizieren. Hier wie schon bei den Relevanzeinschätzungen lag der Schwerpunkt in Englisch auf sprachproduktiven Prozessen, deren enorme Bedeutung Interviewstudien in ähnlicher Weise indizieren (Folkerts, in diesem Band; Prusse-Hess & Prusse 2018; Rossa, 2017), in Latein eher auf rezeptiven Aspekten. Reflexions- und diskussionswürdig erscheint jedoch, dass in den Freitextantworten beider Disziplinen – wie auch bei Rossas (2017) Untersuchung zu Überzeugungen von Fremdsprachenlehrkräften hinsichtlich effektiven Unterrichts – vor allem Aspekte der Auswahl und Thematisierung von Fachinhalten und -methoden vorkamen. Auf die weiteren sechs Dimensionen entfielen kaum Nennungen. Die Fokussierung dieses Merkmals, das sich nach Ufer und Praetorius (2022) insbesondere auf Unterrichtsplanung und -arrangement bezieht, kann einerseits nahelegen, dass (angehende) Lehrkräfte Aspekte von Unterrichtsqualität gedanklich offenbar eher damit assoziieren. Andererseits mag dies auch ein eingeschränktes oder verzerrtes Bewusstsein für die Vielfalt und Vielschichtigkeit von Merkmalen, die bei der Gestaltung eines guten Unterrichtsangebots berücksichtigt werden könnten, zum Ausdruck bringen oder darauf hinweisen, dass möglicherweise nicht vollständig erkannt wird, worauf es dabei wirklich ankommt. Eine solche Auslegung würde auch die hohe Anzahl an Antworten, die gar keine Merkmale des Unterrichts bzw. nicht dessen Prozessebene betrafen, unterstreichen. Ist es in der fachdidaktischen und bildungswissenschaftlichen Lehre und Forschung zu Unterrichtsqualität – vor allem in den letzten 15 Jahren seit der Veröffentlichung der wegweisenden DESI-Ergebnisse (Klieme et al., 2008) oder der Meta-Metaanalyse von Hattie (2009) – also nicht gelungen, (angehende) Lehrkräfte mit ihren zentralen Aussagen und wichtigen Befunden zu erreichen und diese kokonstruktiv für die Unterrichtspraxis nutzbar zu machen? Wie könnten erfolversprechende Implementationsmaßnahmen künftig angelegt sein, um auf Basis überprüfter empirischer Befunde gemeinsam mit angehenden und berufstätigen Lehrkräften zu einer (Weiter-)Entwicklung und Qualitätssicherung von Unterricht beizutragen?

Als Limitation der vorliegenden explorativen Studie ist ihre nicht repräsentative Gelegenheitsstichprobe anzusehen, die unter anderem in andauernden pandemiebedingten Einschränkungen zum Durchführungszeitpunkt begründet ist. Die oftmals stichpunktartige Kürze und Prägnanz der Freitextantworten stellte deren Auswertung und Interpretation vor große Herausforderungen, die über ein mehrfach revidiertes Kodiermanual, Doppelkodierungen und mehrere Kodierungsdurchgänge bewältigt wurden. Zudem erscheinen Rückschlüsse auf die Bedeutsamkeit generischer Aspekte für die Fächer Englisch und Latein nur unter Berücksichtigung der genauen Instruktion zulässig, da explizit nach Merk-

malen guten Englisch- bzw. Lateinunterrichts gefragt wurde und damit ausdrücklich eine fachspezifische Perspektive eingenommen werden sollte. Bezüglich der Relevanzeinschätzung ist ferner zu überdenken, ob die vorgegebenen Aussagen für alle teilnehmenden (angehenden) Lehrkräfte eindeutig verständlich waren bzw. dieselbe inhaltliche Bedeutung besaßen. Gewisse Unterschiede zwischen den Fächern mögen teilweise durch eine unterschiedliche Wahrnehmung und fachtraditionelle Auffassung einzelner Begriffe und Konzepte wie z. B. Üben bedingt sein, spiegeln aber auch in diesem Fall interessante Fachspezifika wider. Mit der Wahl einer fünfstufigen Likert-Skala geht die Gefahr geringerer Variabilität und einseitiger Einschätzungen bei bestimmten Merkmalen einher. Dies kam mit Blick auf alle fünf an der Onlineumfrage beteiligten Fächer allerdings nicht vor (Lindl et al., eingereicht), sodass offenbar durchaus differenzierte Bewertungen möglich waren und erfolgten. Abschließend gilt es zu betonen, dass alle der hier diskutierten Prozessmerkmale lediglich normative Beschreibungskategorien guten Unterrichts darstellen. Erst zu einigen wenigen liegen belastbare Befunde bezüglich positiver Wirkzusammenhänge mit Lernerfolgen von Schülerinnen und Schülern vor (z. B. Helmke & Klieme, 2008; Helmke et al., 2008); in den meisten Fällen steht eine empirische Überprüfung noch aus.

Gerade dies ist das Ziel der Forschungsgruppe FALKO-PV, in der auf Basis der Ideen anregenden Ergebnisse der Onlinebefragung fachspezifische Merkmale von Unterrichtsqualität in den Fächern Englisch und Latein modelliert und unter anderem in einer Webapp zur digitalen Evaluation des Unterrichts aus Sicht von Schülerinnen und Schülern operationalisiert wurden. Diese steht in einer umfangsreduzierten Version nicht nur zum individuellen Einsatz in der Lehrkräftebildung kostenlos zur Verfügung.³ Vielmehr wird sie in der Hauptstudie von FALKO-PV verwendet, um zu überprüfen, welche der identifizierten und diskutierten Merkmale für Lernerfolge von Schülerinnen und Schülern in den Fremdsprachen Englisch und Latein tatsächlich bedeutsam sind und somit qualitätsvollen Unterricht ausmachen.

Literatur

- Abrams, L. M. (2007). Implications of High-Stakes Testing for the Use of Formative Classroom Assessment. In J. H. McMillan (Hrsg.), *Formative Classroom Assessment. Theory into Practice* (S. 79–98). New York/London: Teacher College, Columbia University.
- Amrein, A. L. & Berliner, D. C. (2002). High-Stakes Testing, Uncertainty, and Student Learning. *Education Policy Education*, 10(18). <https://doi.org/10.14507/epaa.v10n18.2002>
- Bäcker, N. (2010). Motivation. In N. Keip & T. Doepner (Hrsg.), *Interaktive Fachdidaktik Latein* (S. 191–207). Göttingen: Vandenhoeck & Ruprecht.
- Bäcker, N. (2019). Motivation. In N. Keip & T. Doepner (Hrsg.), *Interaktive Fachdidaktik Latein* (4., verb. und erw. Aufl., S. 239–258). Göttingen: Vandenhoeck & Ruprecht.

3 Die Webapp und weitere Informationen hierzu sind unter amadeus.falko-pv.de zugänglich.

- Begrich, L., Praetorius, A.-K., Decristan, J., Fauth, B., Göllner, R., Herrman, C., Klein-knecht, M., Taut, S. & Kunter, M. (2023). Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft. *Unterrichtswissenschaft*, 51(3), 63–97. <https://doi.org/10.1007/s42010-023-00163-4>
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56(3), 205–213. <https://doi.org/10.1177/0022487105275904>
- Charalambous, C. Y. & Praetorius, A.-K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM – Mathematics Education*, 50(3), 355–366. <https://doi.org/10.1007/s11858-018-0914-8>
- Drechsel, B. & Schindler, A.-K. (2019). Unterrichtsqualität. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 353–372). Berlin: Springer. https://doi.org/10.1007/978-3-662-55754-9_18
- Dunn, L. M., Dunn, D. M., Styles, B. & Sewell, J. (2009). *The British Picture Vocabulary Scale* (3. Aufl.). London: GL Assessment.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive. Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3), 127–137. <https://doi.org/10.1024/1010-0652/a000129>
- Fend, H. (1981). *Theorie der Schule*. München: Urban & Schwarzenberg.
- Freytag, V. & Theurer, C. (2020). Ästhetisches Erleben oder Nützlichkeit? – Eine empirische Untersuchung zu Perspektiven von Lehrkräften auf ästhetische Bildung. In E. Pürgstaller, S. Konietzko & N. Neuber (Hrsg.), *Kulturelle Bildungsforschung. Methoden, Befunde und Perspektiven* (S. 23–39). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-30602-1_3
- Gardemann, C. (2021). *Literarische Texte im Englischunterricht der Sekundarstufe I. Eine Mixed Methods-Studie mit Hamburger Englischlehrer*innen*. Heidelberg: J.B. Metzler. <https://doi.org/10.1007/978-3-662-62716-7>
- Gerlach, D. & Lüke, M. (2020). Förderung von Lesekompetenz im Englischunterricht. Ergebnisse einer Interventionsstudie. *Zeitschrift für Fremdsprachenforschung*, 31(2), 159–182.
- Göbel, K. (2007). *Qualität im interkulturellen Englischunterricht. Eine Videostudie*. Münster: Waxmann.
- Große, M. (2017). *Pons Latinus – Latein als Brücke zum Deutschen als Zweitsprache. Modellierung und empirische Erprobung eines sprachsensiblen Lateinunterrichts*. Frankfurt a. M.: Peter Lang. <https://doi.org/10.3726/b11739>
- Guttko, J. (2023). Kognitive Aktivierung im Fremdsprachenunterricht. Ein systematisches Review von Forschungsarbeiten aus dem deutschsprachigen Raum. *Zeitschrift für Fremdsprachenforschung*, 34(2), 145–176.
- Häger, H.-J. (2017). *Motivation im Lateinunterricht. Kompetenzorientiertes Unterrichtsmaterial zu den Briefen Ciceros, Senecas und des jüngeren Plinius*. Bamberg: C. C. Buchner.
- Hattie, J. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge. <https://doi.org/10.4324/9780203887332>
- Helmke, A. (2004). *Unterrichtsqualität. Erfassen, Bewerten, Verbessern*. Seelze: Klett-Kallmeyer.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Aufl.). Seelze: Klett-Kallmeyer.
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung. Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Seelze: Klett-Kallmeyer.

- Helmke, A. & Klieme, E. (2008). Unterricht und Entwicklung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 301–312). Weinheim: Beltz. <https://doi.org/10.25656/01:3149>
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 345–363). Weinheim: Beltz. <https://doi.org/10.25656/01:3149>
- Hochweber, J., Steiner, B. & Klieme, E. (2012). Lehrerverkennung, Unterrichtsqualität und Lernergebnisse im Fach Englisch. *Unterrichtswissenschaft*, 40(4), 351–370.
- Jesper, U. (2016). *Inklusiver Lateinunterricht. Ein methodisch-didaktischer Leitfaden zur Förderung von Schülerinnen und Schülern mit einer Autismus-Spektrum-Störung*. Kronshagen: IQSH.
- Jesper, U. (2018). Inklusiver Lateinunterricht. In M. Korn (Hrsg.), *Latein Methodik. Handbuch für die Sekundarstufe I und II* (S. 213–224). Berlin: Cornelsen.
- Käfer, J. (2022). *Umgang mit Fehlern und seine Bedeutung für den Lernerfolg im Englischunterricht*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-37342-9>
- Kersten, K. (2020). Der Zusammenhang von kognitiven Fähigkeiten und Zweisprachigkeit im frühen L2-Erwerb. Individuelle und externe Faktoren. In H. Böttger, J. Festman & T. Müller (Hrsg.), *Language Education and Acquisition Research. Focusing Early Language Learning* (S. 82–116). Bad Heilbrunn: Julius Klinkhardt.
- Kersten, K. (2021). L2 Input and characteristics of instructional techniques in early foreign language classrooms. Underlying theory and pedagogical practice. *European Journal of Applied Linguistics and TEFL*, 10(2), 27–59.
- Kersten, K. (2022). The Proximity of Stimulation Hypothesis. Investigating the interplay of social and instructional variables with the cognitive-linguistic skills of young L2 learners. In K. Kersten & A. Winsler (Hrsg.), *Understanding Variability in Second Language Acquisition, Bilingualism, and Cognition. A Multi-Layered Perspective* (S. 131–159). London: Routledge. <https://doi.org/10.4324/9781003155683-6>
- Kersten, K., Bruhn, A.-C., Ponto, K., Böhnke, J. & Greve, W. (2018). Teacher Input Observation Scheme (TIOS) and Manual. *Studies on Multilingualism in Language Education*, 4. Universität Hildesheim.
- Kersten, K., Piske, T., Rohde, A., Steinlen, A. K., Weitz, M., Couve de Murville, S. & Kurth, S. (2012). ELIAS Grammar Test II. *Studies on Multilingualism in Language Education*, 1. Universität Hildesheim.
- Kersten, K., Ponto, K., Bruhn A.-C. & Adler, M. (2019). Projekt SMILE – Abschlussbericht. *Studies on Multilingualism in Language Education*, 5. Universität Hildesheim.
- Kipf, S. (2019). Mit Heterogenität produktiv umgehen? Sprachsensibel Latein unterrichten! In N. Keip & T. Doepner (Hrsg.), *Interaktive Fachdidaktik Latein* (4., verb. und erw. Aufl., S. 77–96). Göttingen: Vandenhoeck & Ruprecht.
- Klausnitzer, P. (2018). Lehrer-Schüler-Interaktion im Unterrichtsgespräch beim Umgang mit Schülerfehlern. In M. Korn (Hrsg.), *Latein Methodik. Handbuch für die Sekundarstufe I und II* (S. 205–212). Berlin: Cornelsen.
- Klieme, E. (2006). Empirische Unterrichtsforschung. Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52(6), 765–773. <https://doi.org/10.25656/01:4487>

- Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff, H.-G., Schröder, K., Thomé, G. & Willenberg H. (Hrsg.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz. <https://doi.org/10.25656/01:3149>
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127–146). Münster: Waxmann.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237. <https://doi.org/10.25656/01:4348>
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. Aufgabenkultur und Unterrichtsgestaltung. In BMBF (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Videodokumente* (S. 43–58). Bonn: BMBF.
- Kuhlmann, P. (2014). *Lateinische Grammatik unterrichten – Didaktik des lateinischen Grammatikunterrichts*. Bamberg: C. C. Buchner.
- Kuhlmann, P. & Horstmann, H. (2018). *Wortschatz und Grammatik üben. Didaktische Kriterien und Praxisbeispiele für den Lateinunterricht*. Göttingen: Vandenhoeck & Ruprecht.
- Kunter, M. & Ewald, S. (2016). Bedingungen und Effekte von Unterricht. Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Münster: Waxmann.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Ferdinand Schöningh. <https://doi.org/10.36198/9783838538952>
- Liebsch, A.-C. (2019). Inklusionssensibler Lateinunterricht: Auf die Planung kommt es an. In J. Frohn, E. Brodesser, V. Moser & D. Pech (Hrsg.), *Inklusives Lehren und Lernen. Allgemein- und fachdidaktische Grundlagen* (S. 107–114). Bad Heilbrunn: Julius Klinkhardt.
- Liebsch, A.-C. (2020). Cogitamus ergo sumus. Kognitiv aktivierende Aufgaben im Lateinunterricht. *Pegasus-Onlinezeitschrift*, 19, 1–33.
- Liebsch, A.-C. (2021). Inklusiver Lateinunterricht. In S. Kipf, U. Jesper & T. Riecke-Bauleck (Hrsg.), *Basiswissen Lehrerbildung: Latein unterrichten* (S. 86–94). Hannover: Klett-Kallmeyer.
- Lindl, A. (2024). Lateinunterricht (ars docendi) im Spiegel der Wissenschaft – aktuelle Forschungsperspektiven aus dem Vorhaben FALCO-PV. *Ars docendi*, 18, 1–17.
- Lindl, A., Ehrich, P., Gutmiedl, M., Rader, M., Simböck, L., Gürtner, M., Böhringer, S., Krämer, A., Kirchhoff, J. & Frei, M. (2024). Und wo bleibt die Ästhetik? – Betrachtungen zu einer weiteren Dimension von Unterrichtsqualität aus interdisziplinärer Perspektive. In M. Hemmer, C. Angele, C. Bertsch, S. Kapelari, G. Leitner & M. Rothgangel (Hrsg.), *Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung* (S. 371–388). Münster: Waxmann.
- Lindl, A. & Krämer, A. (2022). Was macht guten Lateinunterricht aus? Eine Delphi-Studie zu Unterrichtsqualität im Fach Latein. *Die Alten Sprachen im Unterricht*, 69(1), 9–15.
- Lindl, A., Simböck, L., Rader, M., Hofrichter, M., Gutmiedl, M., Böhringer, S. & Ehrich, P. (eingereicht). Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften

- Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik. *Zeitschrift für Erziehungswissenschaft*.
- Lindmeier, A. & Heinze, A. (2020). Die fachdidaktisch Perspektive in der Unterrichtsqualitätsforschung: (bisher) ignoriert, implizit enthalten oder nicht relevant? In A.-K. Praetorius, J. Grünkorn & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 255–268). Weinheim: Beltz Juventa. <https://doi.org/10.3262/ZPB2001255>
- Lipowsky, F. & Bleck, V. (2019). Was wissen wir über guten Unterricht? Ein Update. In U. Steffens & R. Messner (Hrsg.), *Unterrichtsqualität. Konzepte und Bilanzen gelingenden Lehrens und Lernens. Grundlagen der Qualität von Schule 3* (S. 219–249). Münster: Waxmann.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley-Blackwell. <https://doi.org/10.1075/tblt.8.01lon>
- Maag Merki, K. (2011). The introduction of state-wide exit examinations. Empirical effects on Math and English teaching in German academically oriented secondary schools. In M. A. Pereyra, H.-G. Kotthoff & R. Cowen (Hrsg.), *Changing knowledge, changing tests, and changing schools* (S. 124–142). Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6091-740-0_9
- Maag Merki, K. (2012). Forschungsfragen und theoretisches Rahmenmodell. In K. Maag Merki (Hrsg.), *Zentralabitur. Die längsschnittliche Analyse der Wirkungen der Einführung zentraler Abiturprüfungen in Deutschland* (S. 11–26). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-531-94023-6>
- Miede, S. (2019). *Förderung des Sprechens im kompetenzorientierten Englischunterricht der gymnasialen Oberstufe. Eine qualitativ-empirische Studie*. Tübingen: Narr.
- Müller, V. (2017a). Inklusion im Lateinunterricht. In M. Janka (Hrsg.), *Lateindidaktik. Praxishandbuch für die Sekundarstufe I und II* (S. 232–238). Berlin: Cornelsen.
- Müller, V. (2017b). Schülerindividualität. In M. Janka (Hrsg.), *Lateindidaktik. Praxishandbuch für die Sekundarstufe I und II* (S. 35–39). Berlin: Cornelsen.
- Porsch, R. & Wilden, E. (2021). Teaching English out-of-field in primary school. Differences in professional characteristics and effects on instructional quality. In L. Hobbs & R. Porsch (Hrsg.), *Out-of-field Teaching across the Teaching Disciplines and Contexts* (S. 117–134). Singapore: Springer. https://doi.org/10.1007/978-981-16-9328-1_6
- Praetorius, A.-K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral. Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020a). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Praetorius, A.-K., Rogh, W. & Kleickmann, T. (2020b). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft*, 48(3), 303–318. <https://doi.org/10.1007/s42010-020-00072-w>
- Prusse-Hess, B. & Prusse, M. (2018). Wirksamer Englischunterricht – Ein Fazit. In B. Prusse-Hess & M. Prusse (Hrsg.), *Wirksamer Englischunterricht* (S. 213–224). Baltmannsweiler: Schneider Hohengehren Verlag.

- Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B. & Besser, M. (2019). Formative assessment in mathematics. Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*, 60, 154–165. <https://doi.org/10.1016/j.learninstruc.2018.01.004>
- Reusser, K & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49(2), 189–202. <https://doi.org/10.1007/s42010-021-00117-8>
- Riecke-Baulecke, T., Liebsch, A.-C. & Jesper, U. (2021). Unterrichtsqualität. In U. Jesper, S. Kipf & T. Riecke-Baulecke (Hrsg.), *Basiswissen Lehrerbildung: Latein unterrichten* (S. 44–59). Hannover: Klett-Kallmeyer.
- Rossa, H. (2017). You teach what you believe in. BELT – Beliefs about effective language teaching. In E. Wilden & R. Porsch (Hrsg.), *The professional development of primary EFL teachers. National and international research* (S. 197–218). Münster: Waxmann.
- Rothgangel, M. (2021). Unterrichtsqualität in der Religionsdidaktik – fachspezifische und fachübergreifende Aspekte. *Unterrichtswissenschaft*, 49, 253–260 (2021). <https://doi.org/10.1007/s42010-021-00107-w>
- Rothland, M. (2024). Auf Sand gebaut? Über den Versuch, die Basisdimensionen der Unterrichtsqualität in erziehungswissenschaftlicher „Theorie“ zu verankern. *Zeitschrift für Bildungsforschung*, 14, 41–53. <https://doi.org/10.1007/s35834-024-00411-2>
- Scholz, I. (2018a). Diagnostik und Leistungsbeurteilung. In M. Korn (Hrsg.), *Latein Methodik. Handbuch für die Sekundarstufe I und II* (S. 185–194). Berlin: Cornelsen.
- Scholz, I. (2018b). Differenzierung. In M. Korn (Hrsg.), *Latein Methodik. Handbuch für die Sekundarstufe I und II* (S. 195–212). Berlin: Cornelsen.
- Scholz, I. (2019a). Diagnose und Differenzierung. In N. Keip & T. Doepner (Hrsg.), *Interaktive Fachdidaktik Latein* (4., verb. und erw. Aufl., S. 193–208). Göttingen: Vandenhoeck & Ruprecht.
- Scholz, I. (2019b). Übung. In N. Keip & T. Doepner (Hrsg.), *Interaktive Fachdidaktik Latein* (4., verb. und erw. Aufl., S. 223–237). Göttingen: Vandenhoeck & Ruprecht.
- Scholz, I. & Weber, K.-C. (2011). *Denn sie wissen, was sie können. Kompetenzorientierte und differenzierte Leistungsbeurteilung im Lateinunterricht* (2., verb. Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade. The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Steinhilber, J. (1986). *Die Übung im lateinischen Sprachunterricht. Grundlagen, Methoden, Beispiele*. Bamberg: C. C. Buchner.
- Thaler, E. (2014). Kognitive Aktivierung & Hattie & Fremdsprachenunterricht. *Praxis Fremdsprachenunterricht*, 6, 5–6.
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Weitz, M. (2015). *Die Rolle des L2-Inputs in bilingualen Kindergärten*. Frankfurt a. M: Peter Lang. <https://doi.org/10.3726/978-3-653-05132-2>
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>
- Wilden, E. & Porsch, R. (2019). The impact of teaching quality and learning time on primary EFL learners' receptive proficiency. Preliminary findings from the TEPS study. *AILA Review*, 32, 160–177. <https://doi.org/10.1075/aila.00025.wil>

Wilden, W., Porsch, R. & Schurig, M. (2020). An early start in primary EFL education and the role of teacher qualification and teaching quality. *Language Teaching for Young Learners*, 2(1), 28–51. <https://doi.org/10.1075/ltyl.19002.wil>

Wittich, P. (2015). *Latein unterrichten: Planen, durchführen, reflektieren*. Berlin: Cornelsen.

Daneben wird in Tabelle 2 vollständig auf folgende Themenhefte der fachdidaktischen Zeitschrift *Der Altsprachliche Unterricht (AU)* verwiesen: 2/1997, 6/2000, 1/2008, 1/2012, 4/2014, 2/2019.

Daniel Reimann

Unterrichtsqualität in den romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen (Französisch, Spanisch, Italienisch, Portugiesisch)

Abstract (der Herausgeber:innen)

In his article “Instructional quality in the Romance languages as second, third and late foreign languages (French, Spanish, Italian, Portuguese),” Daniel Reimann examines the four Romance languages French, Spanish, Italian and Portuguese, focussing on their role as foreign languages that are generally learnt later (than, for example, English). The author explores the question of how instructional quality in this context can be modelled with regard to subject-matter requirements, based on the current state of research in these subjects. First, he derives 13 subject-specific characteristics of the Romance languages and then develops 14 specific features in relation to their role as second, third and late foreign languages.

1 Schulpraktischer Kontext, Fragestellung und Vorgehen

Die vier romanischen Sprachen Französisch, Italienisch, Spanisch und Portugiesisch sind deutschlandweit an allen Positionen der schulischen Fremdsprachenfolge belegt und werden vor allem an Gymnasien und Gesamtschulen, aber auch an Realschulen und im berufsbildenden Bereich unterrichtet. Dabei zeichnen sich folgende Schwerpunkte ab: Französisch wird inzwischen vor allem als zweite und dritte Fremdsprache angeboten und belegt, Spanisch vor allem als zweite, dritte und spät beginnende Fremdsprache, Italienisch als dritte und spät beginnende Fremdsprache, zunehmend aber auch als zweite Fremdsprache (z. B. in Thüringen, Sachsen, Berlin, Nordrhein-Westfalen) und Portugiesisch aktuell nur sehr punktuell als dritte und spät beginnende Fremdsprache sowie neuerdings als zweite Fremdsprache in Hessen ab dem Schuljahr 2023/2024 (Hessischer Landtag, 2021). Aus der folgenden Tabelle wird die Gesamtzahl der Schüler:innen an allgemeinbildenden Schulen in Deutschland, die romanische Sprachen erlernen, ersichtlich (Tab. 1; Statistisches Bundesamt, 2022; Reimann, 2017a, bes. S. 264–278, zum Portugiesischen, da Portugiesisch vom Statistischen Bundesamt nicht eigens ausgewiesen wird). Als Bezugsgröße wird ferner auch die Zahl der Schüler:innen mit Latein erfasst, das in zahlreichen schulischen Sprachenbiografien in Verbindung mit romanischen Sprachen auftritt (s. u., Abschnitt 3.2).

Reimann, D. (2024). Unterrichtsqualität in den romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen (Französisch, Spanisch, Italienisch, Portugiesisch). In V. Lohe, A. Lindl & P. Kirchhoff (Hrsg.), *Unterrichtsqualität in schulischen Fremdsprachen. Theoretische Ansätze und empirische Ergebnisse aus den Fachdidaktiken* (S. 139–155). Münster: Waxmann. https://doi.org/10.31244/9783830_999201.06

Tabelle 1: Schüler:innen der romanischen Fremdsprachen und Latein an allgemeinbildenden Schulen in Deutschland (eigene Darstellung)

Sprache	Zahl der Schüler:innen
Französisch	ca. 1,3 Mio.
Latein	ca. 540.000
Spanisch	ca. 500.000
Italienisch	ca. 50.000
Portugiesisch	< 1.000

Mit Bezug auf die Sprachenfolgen ist stark vereinfacht festzustellen, dass Französisch eine etablierte zweite und dritte Fremdsprache ist, mit der Tendenz, die Position der zweiten Fremdsprache einzunehmen. Die Position der ersten Fremdsprache kann das Französische seit flächendeckender Einführung des Englischen an den Grundschulen nur noch selten halten oder sogar neu einnehmen. Spanisch ist inzwischen beinahe als zweite Fremdsprache neben Französisch und Latein etabliert, aber weiterhin auch in den Positionen der dritten und spät beginnenden Fremdsprachen anzutreffen. Italienisch ist als dritte und spät beginnende Fremdsprache fester Bestandteil des schulischen Angebots. Dabei lässt sich die leichte Tendenz erkennen, dass Italienisch zunehmend auch als zweite Fremdsprache angeboten wird. Portugiesisch wird bisher nur sehr punktuell angeboten; Hessen hat als erstes Bundesland dezidiert auf die Bedeutung des Portugiesischen als Weltsprache reagiert (s.o.). Vor diesem Hintergrund verfolgt der Beitrag folgende Fragestellung:

Wie kann Unterrichtsqualität für die romanischen Sprachen mit Blick auf ihre Rolle als in der Regel zweite, dritte und spät beginnende Fremdsprache(n) auf der Grundlage des aktuellen Forschungsstandes spezifisch modelliert werden?

Der Beitrag verfolgt einen hermeneutischen Zugriff und rekapituliert dabei einschlägige Modellierungen benachbarter Disziplinen (Schulpädagogik im Allgemeinen, Allgemeine Didaktik) und Fremdsprachendidaktiken (Didaktik der englischen Sprache und Literatur). Er prüft diese auf ihre Anschlussfähigkeit für die Didaktik der romanischen Sprachen und Literaturen und legt infolgedessen einen theoretisch-konzeptionellen Entwurf zur Unterrichtsqualität speziell in den romanischen Sprachen als zweiten, dritten und spät beginnenden Fremdsprachen vor.

2 Theoretischer Kontext

2.1 Schulpädagogische und allgemein-didaktische Modellierungen

An anderer Stelle (Reimann 2023, bes. S., 248–260) wurde der Versuch unternommen, die in der Literatur mitunter unsystematisch und miteinander konkurrierend verwendeten Begrifflichkeiten wie Unterrichtskonzeptionen, Unterrichtsprinzipien, didaktisch-methodische Prinzipien usw. systematisch zueinander in Beziehung zu setzen und daraus ein Stufenmodell der Unterrichtsmethodik für den Fremdsprachenunterricht vor allgemein schulpädagogischem Hintergrund zu entwickeln. Man kann erkennen, dass es Überschneidungen und Wechselwirkungen zwischen reiner (spekulativer) Theorie und empirisch basierten Aussagen sowie zwischen den Konzepten Unterrichtsprinzipien und Unterrichtsqualität gibt: Zu Unterrichtsprinzipien werden häufig solche grundlegenden Orientierungen erhoben, die (ggf. empirisch begründet) als Merkmale von Unterrichtsqualität angenommen werden (allgemein einführend zur Unterrichtsqualität aus der Perspektive der Schulpädagogik und der Bildungswissenschaften vgl. z. B. Gröschner & Kleinknecht, 2013; vertiefend z. B. Helmke, 2017). Bei Integration der verschiedenen Beschreibungsversuche von Unterricht, die auf unterschiedlichen Ebenen ansetzen, kann ein Stufenmodell mit fünf Ebenen vorgeschlagen werden: Auf den Ebenen 1 und 2, der schulpädagogischen bzw. allgemeindidaktischen Ebene, bewegen sich grundlegende Überlegungen zur Konzeption von Unterricht, die fächerübergreifend gültig und daher nicht für den Fremdsprachenunterricht spezifisch sind, gleichwohl aber den Fremdsprachenunterricht betreffen. Dabei betrifft die Ebene 1 Unterrichtskonzeptionen (z. B. lehrer:innen- und schüler:innengesteuerter Unterricht), wohingegen auf Ebene 2 Unterrichtsprinzipien bzw. Unterrichtsgrundsätze angesiedelt sind, die häufig auch mit Qualitätsmerkmalen von Unterricht gleichgesetzt wurden. Wiater (2020a und b) etwa unterscheidet zwischen fundierenden Unterrichtsprinzipien wie der Schüler:innenorientierung (vgl. S. 131–133; S. 22–31) und regulierenden Unterrichtsprinzipien wie der Differenzierung (vgl. S. 134–139; S. 32–130). Ebene 3 stellt eine Ebene des Übergangs von der schulpädagogischen bzw. allgemeindidaktischen Ebene zur fachdidaktischen und im konkreten Fall fremdsprachendidaktischen bzw. romanistisch-didaktischen Ebene dar. Es handelt sich hierbei um didaktisch-methodische Prinzipien des (Fremdsprachen-)Unterrichts wie Mehrsprachigkeitsorientierung. In diesem Bereich werden einerseits Prinzipien erfasst, die auch in anderen Fächern Gültigkeit beanspruchen können, andererseits aber auch Prinzipien, die durchaus als spezifisch für den Fremdsprachenunterricht gelten dürfen (wie z. B. Mehrsprachigkeitsorientierung). Ebene 4 betrifft die Organisation von Unterricht. Darunter werden die Aktions- und Arbeitsformen einerseits und die Sozialformen andererseits in ihrer jeweiligen fachbezogenen Realisierung gefasst, nunmehr regelmäßig ergänzt um Formate des offenen Unterrichts, wie z. B. Gruppenarbeit oder Lernzirkel. Ebene 5 ist die Ebene der konkreten Unterrichtstechniken wie z. B. der *one-minute talk* (Reimann, 2023, S. 248–250).

Um sich der Frage nach Unterrichtsqualität in den romanischen Sprachen, gerade auch der zweiten, dritten und spät beginnenden Fremdsprachen, zu nähern, bietet es sich weiterhin an, bekannte schulpädagogische Kriterienkataloge ‚guten Unterrichts‘ zur Kenntnis zu nehmen (z. B. Meyer, 2020; Lipowsky, 2020; weiterhin Helmke, zuletzt 2022). Eine Konzentration des daraus letztlich hervorgehenden Konsenses findet sich in den TIMS-Studien seit 1995, in deren Rahmen als ebenfalls hinlänglich bekannte Qualitätsdimensionen von Unterricht benannt werden: Klassenführung und Strukturierung, Schüler:innenorientierung und Unterstützung, kognitive Aktivierung (vgl. Klieme et al., 2001, S. 51).

Ausgehend von der Annahme der genannten drei Basisdimensionen von Unterrichtsqualität können die folgenden Reflexionen zu Tiefenstrukturen und Prozessmerkmalen von Unterricht der romanischen Sprachen im Kontext der in der Einleitung von Lindl et al. dargestellten Modelle von Krauss et al., 2020 („Kaskadenmodell“, vgl. Abb. 1 im zitierten Beitrag) und von Charalambous & Praetorius, 2020 (MAIN-TEACH-Modell, vgl. Abb. 2 im zitierten Beitrag) verstanden werden. Sie wären dort insbesondere im Bereich der Säulen 3 bis 5 (Kaskadenmodell) bzw. auf der mittleren Ebene (Klassenführung, sozio-emotionale Unterstützung, Unterstützung der aktiven Beteiligung) und der innersten Ebene der Auswahl und Thematisierung von Inhalten und Fachmethoden, der kognitiven Aktivierung, der Unterstützung des Übens und des formativen Assessments (MAIN-TEACH-Modell) zu verorten.

2.2 Fremdsprachendidaktische Modellierungen

Im Bereich der Fremdsprachendidaktik gibt es bisher nur sehr wenige Entwürfe zu Unterrichtsqualität. Diese stammen vor allem aus der Englischdidaktik. Auf fundierter empirischer Basis ist hier nach wie vor an die DESI-Studie von 2003/2004 zu erinnern, die mit Blick auf Unterrichtsqualität im Englischunterricht unter anderem zu folgenden Ergebnissen gelangt:

„In erfolgreichen Klassen ...

- kommen Schülerinnen und Schüler häufig zum Sprechen [...],
- warten Lehrerinnen und Lehrer mindestens drei Sekunden auf Schülerantworten,
- ist bei Unterrichtsgesprächen Englisch die überwiegende Unterrichtssprache,
- erhalten Schülerinnen und Schüler Gelegenheit zur Selbstkorrektur ihrer Fehler,
- gibt es vergleichsweise wenige „Ein-Wort-Sätze“,
- kommt es zu Lehrer-Schüler-Dialogen (mehrere Gesprächsstationen, über einfache Frage-Antwort-Sequenzen hinaus)“ (Helmke et al., 2008, S. 361, vgl. Wilden, 2021, S. 214).

Weiterhin kann die Publikation *Wirksamer Englischunterricht* von Prusse-Hess und Prusse (2018) angeführt werden. Datengrundlage sind hier schriftliche In-

interviews mit Fachdidaktiker:innen sowie erfahrenen Unterrichtspraktiker:innen (insgesamt 18 bzw. 9 aus jeder der beiden Gruppen). Nach deren Ansicht zeichnet sich „wirksamer (Englisch-) Unterricht“ in der Zusammenfassung von Prusse-Hess und Prusse (2018) durch folgende Merkmale aus:

- hohe Sprachkompetenz der Lehrkraft,
- Methodenvielfalt,
- Balance zwischen offenen und geschlossenen Lernformen,
- differenzierte Arbeitsaufträge / Aufgaben,
- Zielsprache als Unterrichtssprache,
- anspruchsvolle und kommunikationsfördernde Aufgaben,
- sorgfältige Unterrichtsplanung,
- vielfältiger Medieneinsatz,
- bilingualer Sachfachunterricht,
- „eine angemessene Repräsentation der Breite des Wissensgebiets“ (Thaler, 2019, S. 5)

Nach Thaler (2014) und Wilden (2021) erfährt die kognitive Aktivierung etwa der TIMS-Studien und nachfolgender Veröffentlichungen zur Unterrichtsqualität in anderen Fächern für den Fremdsprachenunterricht eine spezifische Ausprägung als „kommunikativ-kognitivierende Aktivierung“ bzw. allgemeiner „kommunikative Aktivierung“ (Wilden, 2021, bes. S. 214, 216).

3 Vorschläge zur spezifischen (Weiter-) Entwicklung des Konzepts Unterrichtsqualität für die romanischen Schulsprachen

3.1 Romanistisch-didaktische Beiträge zu Unterrichtsprinzipien

Forschung zu Unterrichtsqualität im engeren Sinne fehlt innerhalb der romanistischen Fachdidaktik bisher weitgehend. Erst recht liegen praktisch keine entsprechenden empirischen Forschungen vor. Zu didaktisch-methodischen Unterrichtsprinzipien der o. g. fachdidaktischen Ebene 3 finden sich in diversen unterrichtspraktisch orientierten, einführenden Handbüchern zu den Fachdidaktiken und Fachmethodiken der romanischen Sprachen verschiedene Listen, die in Tabelle 2 synoptisch zusammengetragen werden (vgl. analog für das Fach Latein Hofrichter et al., in diesem Band).

Tabelle 2: Didaktisch-methodische Prinzipien (Ebene 3) des Fremdsprachenunterrichts in diversen Überblicksdarstellungen (eigene Darstellung, Reimann, 2023, S. 256–257)

(Fremdsprachen-)Didaktisch-methodische Prinzipien (Ebene 3)	Krechel, 2016	Müller & Rohling, 2016	Plikat, 2018	Reinfried, 2017b	Steveker, 2019
Lerner:innenorientierung/ Schüler:innenorientierung	X	X		X	X
Differenzierung			X		
nachhaltiger Umgang mit Fehlern	X				
Handlungsorientierung	X	X		X	X
Kommunikationsorientierung		X			
Anwendungsorientierung	X				
Aufgabenorientierung	X		X	X	
Kompetenzorientierung		X			
Standardorientierung					X
Outputorientierung	X	X			
Prozessorientierung				X	X
Inhaltsorientierung, Bildungsanspruch			X	X	
funktionale Zwei- oder Mehrsprachigkeit	X	X			
Mehrsprachigkeit/ Mehrsprachigkeitsdidaktik			X		
Sprachbewusstheit			X		
Ganzheitlichkeit				X	
Rituale	X				

Betrachtet man die Zusammenstellung in Tabelle 2, so ergeben sich daraus wenig grundlegend neue Aspekte, allerdings kann man feststellen, dass folgende didaktisch-methodische Prinzipien offensichtlich besondere Beachtung im aktuellen romanistisch-didaktischen Diskurs genießen: Es ist einerseits eine dezidierte Betonung von Handlungs- und Aufgabenorientierung festzustellen, andererseits ist allmählich eine (postulierte) Rückbesinnung auf Inhaltsorientierung bzw. einen weiter gefassten Bildungsanspruch des Unterrichts der romanischen Sprachen feststellbar. Weiterhin ist eine Beachtung der funktionalen Zwei- bzw. Mehrsprachigkeit und nicht zuletzt ein explizites Aufgreifen mehrsprachigkeitsdidaktischer Aspekte zu beobachten.

Weitere Ansatzpunkte für die Bestimmung von Gütekriterien für den Unterricht der romanischen Sprachen können Diskursfelder der neueren romanistisch-fachdidaktischen Debatte sein, die Marcus Reinfried im Jahr 2001 auf der Grundlage einer systematischen Auswertung fachdidaktischer Beiträge empirisch fundiert ausmachen konnte und die er als konstituierend für einen „neokommunikativen Fremdsprachenunterricht“ bezeichnet hat (Reinfried, 2001; überarbeitet in Reinfried & Volkman, 2012 und in Reinfried, 2017a und b). Diese Beschreibung einer „neokommunikativen Phase“ des Fremdsprachenunterrichts wurde von Reimann theoretisch weiterentwickelt (bes. Reimann, 2014; Reimann, 2015a und b; Reimann, 2023). Zu weit hat sich der Fremdsprachenunterricht insbesondere seit der Jahrtausendwende unter anderem in der Folge der Publikation des *Gemeinsamen europäischen Referenzrahmens für Sprachen* im Jahr 2001 und der fremdsprachlichen Bildungsstandards von 2003 vom kommunikativen Paradigma der 1970er Jahre entfernt bzw. weiterentwickelt, als dass man noch immer einfach von „kommunikativem“ Fremdsprachenunterricht sprechen könnte. Zugleich wird das Grundanliegen der „kommunikativen Methode“, die kommunikative Kompetenz (vgl. z.B. zur Grundlegung in der deutschen fremdsprachendidaktischen Debatte Piepho, 1974) keineswegs in Frage gestellt – ja der „neokommunikative Fremdsprachenunterricht“ schreibt diese unter veränderten Vorzeichen fort und setzt sie im Sinne der Kompetenzorientierung konsequent um. War die bevorzugte Bezugsdisziplin der „kommunikativen Methode“ die linguistische Pragmatik (z.B. Reinfried, 2006, S. 41–42) und waren entsprechende Lehrwerke noch deutlich von strukturalistischen Einflüssen geprägt, so sind wesentliche Diskussionsfelder der neokommunikativen Fremdsprachenforschung, mithin distinktive Merkmale, durch welche sich ein neokommunikativer Fremdsprachenunterricht vom Unterricht der kommunikativen Methode der 1970er Jahre unterscheidet, nach Reinfried (2001; Reinfried & Volkman, 2012): Lerner:innenorientierung, Handlungsorientierung, Ganzheitlichkeit, Prozessorientierung, Interkulturalität und fächerübergreifendes Lernen einschließlich Mehrsprachigkeitsdidaktik. Reinfried (2001) gliedert letztere in drei Teilbereiche: „interlinguale Koordination des Sprachinputs“, „interlinguale Lernstrategien“ und „kontrastive Sprach- und Kulturbewusstheit“ (Reinfried, 2001, S. 10). Aus heutiger Sicht könnte man vom erweiterten Konzept einer „aufgeklärten Mehrsprachigkeit“ sprechen (s. u., vgl. z. B. Reimann, 2015a und b; 2016). Darüber hinaus sind die Diskursfelder Interkulturalität um Transkulturalität (z. B. Reimann, 2013, 2017b und c) und Schüler:innenorientierung um Differenzierung und Inklusion zu erweitern sowie Aufgaben- und Standardorientierung, Multimedialität, Kognitivierung und Metakognition, die implizit teilweise bereits in Reinfried (2001) und Reinfried und Volkman (2012) angelegt sind, zu ergänzen (vgl. Reimann, 2014, bes. S. 90–91, 2015b). Marcus Reinfried selbst hat die inzwischen auch virulente Diskussion um eine „Inhaltsorientierung“, die der offensichtlichen Gefahr einer inhaltlichen Verflachung des Fremdsprachenunterrichts bei allzu ausgeprägter Konzentration auf die Vermittlung der leicht messbaren funktionalen kommunikativen Kompe-

tenzen gefolgt ist, als Charakteristikum des neokommunikativen Fremdsprachenunterrichts hinzugefügt (Reinfried, 2017b, bes. S. 79).

In Ergänzung zu den teilweise doch eher willkürlich wirkenden Zusammenstellungen in den oben ausgewerteten unterrichtspraktisch orientierten Einführungen in die Fachdidaktiken und -methodiken der einzelnen romanischen Sprachen und zur Übersicht über die Prinzipien des neokommunikativen Fremdsprachenunterrichts kann an dieser Stelle zur Orientierung über aktuelle Diskurse über den Unterricht in den romanischen Sprachen weiterhin auf die systematisch begründeten 13 didaktisch-methodischen Prinzipien des Landes Schleswig-Holstein in den *Fachanforderungen Französisch* bzw. *Spanisch* (2015a und b) und *Italienisch* (2018) hingewiesen werden. Die genannten Lehrpläne gehen diesbezüglich weit über die genannten Übersichten zu didaktisch-methodischen Prinzipien, aber auch über die bis dahin veröffentlichten Lehrpläne anderer Bundesländer hinaus. Es ist festzuhalten, dass die Konzepte nicht in exakt dieser Form in den Fachanforderungen für alle drei Sprachen aufgelistet werden, sondern dass die Lehrpläne der romanischen Sprachen diesbezüglich geringfügig voneinander abweichen. An dieser Stelle (Tab. 3) werden 13 didaktisch-methodische Prinzipien aus den Lehrplänen bzw. Fachanforderungen für die drei romanischen Schulsprachen in Schleswig-Holstein integrierend zusammengestellt und wiedergegeben:

Tabelle 3: Vorschlag von 13 didaktisch-methodischen Prinzipien (Ebene 3) für den schulischen Unterricht der romanischen Sprachen (auf der Grundlage von Ministerium für Schule und Berufsbildung, 2015a und b, S. 17–18 bzw. 16 sowie Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2018, S. 13–15, eigene Darstellung)

Prinzip	Erläuterung
Kompetenzorientierung	Der Französisch- / Spanisch- / Italienischunterricht vermittelt fremdsprachliches, interkulturelles und methodisches Wissen und bietet vielfältige Gelegenheiten, dieses Wissen in Performanzsituationen handelnd anzuwenden, wodurch die entsprechenden Kompetenzen entwickelt und gefördert werden. Die Unterrichtsplanung stellt die angestrebten Kompetenzen ins Zentrum, z. B. durch die Arbeit mit komplexen Lernaufgaben.
Kommunikationsorientierung	Zentraler Bestandteil des Französisch- / Spanisch- / Italienischunterrichts ist der Ausbau der mündlichen Kommunikationsfähigkeit. Bei der Wahl der Unterrichtsmethoden ist daher darauf zu achten, dass ein möglichst hoher Sprechanteil der Schüler:innen im Unterricht erreicht wird und sprachlich anspruchsvolle und inhaltlich adäquate Sprechansätze geboten werden. Im Französisch- / Spanisch- / Italienischunterricht erfolgt tatsächliches Interagieren, das sich in einem zielgerichteten und kommunikationsbezogenen Gebrauch des Französischen / Spanischen / Italienischen manifestiert. Der Unterricht orientiert sich am Sprachgebrauch der alltäglichen Kommunikation, in der Sprechen und Hören dominieren. Dem Mündlichen wird daher der Vorrang vor dem Schriftlichen gegeben. Dies spiegelt sich im unterrichtlichen Geschehen, bei der Text- und Materialwahl, in den Aufgabenstellungen und bei der Leistungsbewertung wider.

Prinzip	Erläuterung
Inhaltsorientierung	Der Französisch- / Spanisch- / Italienischunterricht stellt die Kommunikation über Themen und Inhalte ins Zentrum; diese ergeben sich weitestgehend aus dem an der Schule eingeführten Lehrwerk, dessen Inhalte nur als Angebote zu verstehen sind, und orientieren sich an der Lebenswelt der Schüler:innen. Spracharbeit erfolgt stets integrativ; die sprachlichen Mittel erfüllen in diesem Sinne eine dienende Funktion.
Integrative Spracharbeit	Die sprachlichen Mittel (Wortschatz, Grammatik, Aussprache, Intonation, Orthografie) haben dienende Funktion. Sie sollten ausgehend von situativ angelegten Sachinhalten prinzipiell induktiv erarbeitet werden. Da die Schüler:innen bereits über Sprachlernstrategien verfügen, kann aber auch eine deduktivere Herangehensweise zur Erschließung sprachlicher Strukturen gewinnbringend sein. Die sichere Beherrschung sprachlicher Mittel ermöglicht die erfolgreiche Realisierung von Kommunikationsabsichten, wie sie typischerweise in alltäglichen sowie sach- und problembezogenen Kommunikationssituationen auftreten.
Funktionale Einsprachigkeit	Der Französisch- / Spanisch- / Italienischunterricht orientiert sich am muttersprachlichen Vorbild. Bei der Vermittlung des Spanischen stehen das Kastilische und die lateinamerikanischen Varianten gleichberechtigt nebeneinander, wobei die Lehrkraft eine Norm konsequent anwendet. Unterrichtssprache ist Französisch / Spanisch / Italienisch. Die deutsche Sprache wird in Ausnahmefällen zielführend eingesetzt, z. B. bei methodenorientiertem Arbeiten.
Funktionale Fehlertoleranz	Im Französisch- / Spanisch- / Italienischunterricht werden Fehler als Teil des Lernprozesses verstanden. Der Umgang mit ihnen erfolgt reflektiert und mit angemessenem, nach Lernstand und Anspruchsniveau differenziertem Vorgehen. Ziel ist primär die Aufrechterhaltung der Kommunikation, nicht die Fehlerfreiheit der Aussage. Zugleich erhalten die Schüler:innen die für den Lernfortschritt erforderliche Orientierung.
Kriterienorientierung	Im Französisch- / Spanisch- / Italienischunterricht wird zwischen Lern- und Leistungssituationen unterschieden. Die Anforderungen in den Lernsituationen und bei der Überprüfung des Lernzuwachses in mündlichen und schriftlichen Leistungssituationen sind an transparenten Kriterien ausgerichtet, zu deren Erreichen die Lehrkraft den Schüler:innen im Sinne einer Lernhilfe explizit Rückmeldung gibt.
Authentizität	Im Französisch- / Spanisch- / Italienischunterricht begegnen die Schüler:innen einer größtmöglichen Breite an altersgemäßen authentischen Texten im Sinne eines erweiterten Textbegriffs, ggf. in didaktisierter Form. Im Laufe der Sekundarstufe I wird mindestens eine Lektüre behandelt. Die Lernsituationen werden auf der Basis anschaulicher Materialien möglichst lebensnah gestaltet. Der Einsatz von Materialien und Medien bietet Lerngelegenheiten für den selbstständigen, kritischen und kreativen Umgang mit ihnen.
Lernen am anderen Ort	Außerschulische Lernorte und vorhandene internationale Kontakte werden zielführend und unterstützend in den Unterricht eingebunden.
Individualisierung/Differenzierung	Der Französisch- / Spanisch- / Italienischunterricht berücksichtigt die Erkenntnis, dass Lernen ein individueller Aneignungsprozess ist. Er zielt auf die kognitive Aktivierung aller Schüler:innen ab und berücksichtigt deren individuelle Lernvoraussetzungen. Dies bedingt die Vermittlung von fachbezogenen Lernstrategien und die Bereitstellung von differenzierten Lernhilfen.

Prinzip	Erläuterung
Methodenvielfalt	Der Französisch- / Spanisch- / Italienischunterricht ist ein methodisch vielfältiger Unterricht mit variablen Organisationsformen, in denen sich individuelle, kooperative und plenare Arbeitsphasen finden. Der Unterricht lässt auch Raum für offenes und fächerübergreifendes Lernen.
Selbstständigkeit und Eigenverantwortung	Die Schüler:innen werden dazu befähigt, zunehmend Verantwortung für den eigenen Lernprozess und Spracherwerb zu übernehmen, und über die Sprache und den eigenen Sprachlernprozess zu reflektieren. Selbstständigkeit und Eigenverantwortung zum lebenslangen Sprachenlernen werden gefördert.
Mehrsprachigkeitsschulung / Sprachvernetzung	Der Französisch- / Spanisch- / Italienischunterricht nutzt Kenntnisse und Fertigkeiten, die die Schüler:innen bereits in weiteren Fremdsprachen erworben haben, zeigt Verbindungen zwischen Sprachen (auch verschiedenen Erstsprachen) auf, regt zu sprachvernetzendem Lernen an und bereitet durch die Vermittlung von Strategiewissen auf das Erlernen weiterer Fremdsprachen vor. Es wird grundsätzlich für Varietäten sensibilisiert.

3.2 Vorschläge zur spezifischen (Weiter-) Entwicklung des Konzepts „Unterrichtsqualität“ für die romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen

Vor dem Hintergrund der bisherigen Ausführungen, die als Synthese des Forschungsstandes bzw. des Diskurses um Unterrichtsqualität in den romanischen Sprachen gelten dürfen, soll in mehreren Schritten der Versuch unternommen werden, auf der Grundlage der Abschnitte 2.1, 2.2 und 3.1 die für die romanischen Sprachen anzunehmenden Kriterien der Unterrichtsqualität weiter auszudifferenzieren. In einem ersten Schritt soll der Vorschlag einer Weiterentwicklung des Konzepts „Unterrichtsqualität“ für die romanischen Sprachen auf Grundlage der in der Romanistik festgehaltenen Unterrichtsprinzipien erfolgen: Angesichts der in Abschnitt 3.1 rekapitulierten, im romanistisch-didaktischen Diskurs dieser Jahre belegten Unterrichtsprinzipien können zu den in 2.1 und 2.2 festgehaltenen Aspekten folgende Nuancierungen bzw. Ergänzungen aus romanistisch-didaktischen Modellierungen hinzutreten:

- Inhaltsorientierung / Bildungsanspruch (vgl. Ministerium für Schule und Berufsbildung Schleswig-Holstein, 2015ff.; Reinfried, 2017; Plikat, 2018)
- fächerübergreifendes Lernen und Elemente bi- / multilingualen Sachfachunterrichts (vgl. Reinfried, 2001ff.; Reimann, 2014ff.)
- „aufgeklärte Einsprachigkeit“ / funktionale Ein- bzw. Zweisprachigkeit (vgl. Ministerium für Schule und Berufsbildung / Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2015ff.; Krechel, 2016; Müller & Rohling, 2016)
- Handlungsorientierung und Aufgabenorientierung (Krechel, 2016; Müller & Rohling, 2016; Reinfried, 2017b; Plikat, 2018; Steveker, 2019)

- Authentizität und außerschulische Lernorte (vgl. Ministerium für Schule und Berufsbildung / Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2015ff.)
- Selbständigkeit / Eigenverantwortung einschließlich Sprachlernkompetenz (vgl. Ministerium für Schule und Berufsbildung / Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2015ff.)
- Mehrsprachigkeitsschulung / Sprachvernetzung (vgl. Reimann, 2015ff.; Ministerium für Schule und Berufsbildung / Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2015ff.; Plikat, 2018)
- einschließlich (rezeptive) Varietätenkompetenz (vgl. Reimann, 2015ff., vgl. Reimann, 2011, 2017d; Ministerium für Schule und Berufsbildung / Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2015ff.)
- Inklusion (Reimann, 2014ff.; 2019).

Weiterhin sollen ergänzende Vorschläge zur spezifischen Weiterentwicklung des Konstrukts Unterrichtsqualität für die romanischen Sprachen mit Blick auf ihre Rolle als zweite, vor allem aber als dritte und spät beginnende Fremdsprachen in die Diskussion eingebracht werden. Folgende Kriterien können als Spezifika der Unterrichtsqualität in den romanischen Sprachen als zweite, dritte und spät beginnende Fremdsprachen angenommen werden:

1. Der Unterricht einer romanischen Sprache muss, je nachdem, ob es sich um eine zweite, dritte oder spät beginnende Fremdsprache handelt, unterschiedlich gestaltet werden. Dem zunehmenden Lebensalter, der kognitiven Entwicklung und dem jeweiligen Sprachenprofil ist Rechnung zu tragen.
2. Mehrsprachigkeitsdidaktische Aspekte müssen proaktiv und retroaktiv eingebracht werden, d. h., der Unterricht einer romanischen Sprache muss ihre Rolle als wenigstens zweiter, häufig dritter oder spät beginnender Fremdsprache gerecht werden. Dabei fungiert die jeweilige romanische Sprache auch als Sprungbrett zu weiteren romanischen Sprachen und trägt zugleich insgesamt zur Entwicklung von Sprachlernkompetenz und Sprachbewusstsein ab der zweiten Fremdsprache bei. Während Englisch tendenziell nur proaktiv auf weitere Fremdsprachen hinweisen kann – was die romanischen Sprachen auch können und sollen – so kann der Unterricht in den einzelnen romanischen Sprachen jeweils auch retroaktiv mindestens auf das Englische, ggf. auch auf das Lateinische (vgl. Hofrichter et al., in diesem Band) oder weitere gelernte Fremdsprachen zurückgreifen.
3. Unterrichtsqualität im Sinne einer mehrsprachigkeitsdidaktischen Orientierung wird insbesondere in der Sensibilisierung für und Nutzung von Lernerleichterungen durch Transfer (beispielsweise in den Bereichen Lexik, Morphologie, *consecutio temporum*, Modi usw.) manifest (retroaktive Mehrsprachigkeit; vgl. für ähnliche Potenziale des Lateinunterrichts auch Hofrichter et al., in diesem Band).

4. Insbesondere die Sensibilisierung für weitere romanische Sprachen (z.B. Portugiesisch, Rumänisch) stellt darüber hinaus ein Merkmal von Unterrichtsqualität in den romanischen Sprachen dar, wobei sich eine solche Sensibilisierung idealerweise durch eine tatsächliche, systematische Progression in einer weiteren romanischen Fremdsprache auszeichnet (proaktive Mehrsprachigkeit).
5. Die Berücksichtigung des lateinisch-römischen sprachlichen und kulturellen Erbes scheint ein weiteres Qualitätsmerkmal eines bildend konzipierten Unterrichts der romanischen Sprachen zu sein, der Brückenschlag zum Lateinischen, punktuell auch zum Griechischen, ist in den verschiedenen Ausbildungszügen der einzelnen romanischen Sprachen möglich und stellt ein eigenes, spezifisches Merkmal von Unterrichtsqualität in den romanischen Schulsprachen dar (traditionell retroaktiv bei vorgelerntem Latein oder Griechisch, zunehmend auch proaktiv bei einer romanischen Sprache als zweiter und Latein als dritter Fremdsprache; vgl. dazu auch Hofrichter et al., in diesem Band).
6. Im Sinne der beschriebenen mehrsprachigkeitsdidaktischen Orientierung stellt nicht nur die Initiierung des Rückgriffs auf sprachliches und inhaltliches, sondern auch auf strategisch-methodisches Vorwissen und entsprechende Kompetenzen aus der ersten, zweiten und ggf. dritten Fremdsprache ein weiteres Merkmal guten Unterrichts in den romanischen Sprachen dar.
7. Mit Blick auf sich abzeichnende mehrsprachige Lernbiografien, aber auch mit Blick auf das höhere Lebensalter zu Beginn eines Sprachlehrgangs (i. d. R. 13/14 Jahre) können sprachsystembezogene Kognitivierung und Sprachreflexion generell als Qualitätsmerkmal v. a. im Unterricht der dritten und spät beginnenden Fremdsprachen angenommen werden.
8. Aufgrund des höheren Lebensalters und fortgeschrittener kognitiver Entwicklung kann in den romanischen Sprachen als dritte, insbesondere aber als spät beginnende Fremdsprachen bei Aktivierung vorhandener sprachlicher Ressourcen und Sprachlernkompetenzen eine höhere Progression erzielt werden als z.B. im Englischen als erster Fremdsprache und einer romanischen Sprache als zweiter Fremdsprache. Diese Progression stellt eine Folge der oben beschriebenen Qualitätsmerkmale des Unterrichts in den romanischen Sprachen dar (insbesondere als dritte und spät beginnende Fremdsprachen).
9. Eine Betonung der rezeptiven Kompetenzen kann gerade in der spät beginnenden Fremdsprache dazu beitragen, ein noch niedriges sprachliches Niveau mit einer fortgeschrittenen kognitiven Entwicklung der Lernenden und ausgeprägterem (Welt-)Wissen in Einklang zu bringen. Ansprechende und komplexe Inhalte können so bereits relativ früh im Sprachlehrgang behandelt werden. Auch dies stellt ein Qualitätsmerkmal des Unterrichts der romanischen Sprachen, hier speziell als spät beginnende Fremdsprachen, dar.
10. Auch im Bereich des kulturellen Lernens kann eine Didaktik der Mehrkulturalität proaktiv und retroaktiv ausgestaltet werden, indem einerseits bekannte kulturelle Phänomene aus anderen Sprachen aufgegriffen werden und andererseits proaktiv etwa auf Kulturen in anderen romanophonen Sprachgemein-

- schaften hingewiesen wird. Inter- und transkulturelle Kompetenzentwicklung im Unterricht einer romanischen Sprache sollte anhand spezifisch romanistischer Inhalte aus den verschiedenen Kontinenten erfolgen.
11. Die Vermittlung der romanischen Sprachen als Kultur- und Bildungssprachen sollte als ein weiteres Qualitätsmerkmal des Unterrichts in den romanischen Sprachen angesehen werden, aus dem ggf. auch ein Profilvermerkmal der romanischen Sprachen neben dem eher utilitaristisch wahrgenommenen Englischen entstehen kann.
 12. Differenzierung und Inklusion sollten noch stärker als Merkmal für Unterrichtsqualität betont werden. Hier ist auch der Spezifik der zweiten und dritten bzw. spät beginnenden Fremdsprachen Rechnung zu tragen, in denen mitunter sehr heterogene Lerngruppen neu gebildet werden. Man denke nur an den Fall einer spät beginnenden Fremdsprache als zweite Fremdsprache an einer Gesamtschule oder einem Oberstufenzentrum, an der/dem neben Schüler:innen, für die die spät beginnende Fremdsprache eine zweite Fremdsprache darstellt, Schüler:innen treten, für die der gleiche Kursus einen Sprachlehrgang in einer vierten Fremdsprache darstellen kann.
 13. Als Sonderfall von Inklusion kann die Berücksichtigung von Herkunfts- bzw. Familiensprachen im Fremdsprachenunterricht zum Qualitätskriterium erhoben werden (wobei dies grundsätzlich auch für den Englischunterricht zutrifft, mithin nicht als spezifisches Qualitätsmerkmal des Unterrichts der romanischen Sprachen gelten kann). Dabei kommt der Berücksichtigung einer besonderen Gruppe von Herkunftssprecher:innen, namentlich der Schüler:innen mit zielsprachlichem Hintergrund, herausragende Bedeutung zu (vgl. Reimann, 2020, 2021). Diese stellt für den Unterricht der romanischen Sprachen als Sprachen einer traditionellen, historischen, aber auch aktuellen Migration nach Deutschland (aber beispielsweise auch als Landessprachen in der Deutschschweiz) nicht nur ein Qualitätsmerkmal, sondern ein besonderes Gütekriterium dar.

Insgesamt sind diese Merkmale im Rahmen des Konzepts einer „aufgeklärten Mehrsprachigkeit“ realisierbar (vgl. Reimann, 2015a und b; 2016). Mit Blick auf die dritten und vor allem auf die spät beginnenden Fremdsprachen kommen sicherlich insbesondere die Merkmale 2, 8, 9 und 11 in besonderem Maße zur Geltung, da sich der Unterricht gerade in den spät beginnenden Fremdsprachen dadurch auszeichnet, einerseits eine hohe sprachliche Progression zu ermöglichen, andererseits schon vor Erreichen eines hohen produktiven Zielniveaus inhalts- und bildungsorientiert gearbeitet werden kann.

Die genannten Aspekte haben mit Blick auf die in Abschnitt 2.1 resümierten schulpädagogischen Modellierungen ihren Bezug vor allem in den beiden Basisdimensionen der Schüler:innenorientierung und Unterstützung der Schüler:innen sowie der kognitiven Aktivierung. Sie sind Manifestationen insbesondere der Säule 3 des Kaskadenmodells von Krauss et al. (2020) in ihrer jeweiligen fach- und ausbildungszugspezifischen (zweite, dritte oder spät beginnende Fremd-

sprache) Realisierung. Dabei gehen die Merkmale 4, 5, 10 und 11 insofern auch darüber hinaus, als sie bildungsorientierte inhaltliche Aspekte gerade auch mit Blick auf die dritten und spät beginnenden Fremdsprachen fokussieren. Im Rahmen des MAIN-TEACH-Modells (Charalambos & Praetorius, 2020) könnten die genannten Kriterien als fach- und ausbildungszugspezifische Modellierungen insbesondere der Aspekte „Auswahl und Thematisierung von Inhalten und Fachmethoden“ (Merkmale 2, 4, 5, 9, 10, 11, 14), „kognitive Aktivierung“ (Merkmale 1, 2, 3, 4, 6, 7, 8, 9, 14) und „Unterstützung des Übens“ (Merkmale 1, 2, 3, 6, 9, 12, 13, 14) angesehen werden. Anhand dieser Zuordnungen zeigt sich zugleich, dass die zitierten Modelle aus den Bildungswissenschaften bzw. anderen Fächern zwar anschlussfähig, aber nur bedingt geeignet sind, um der Spezifik der romanischen Sprachen als zweiten, dritten und spät beginnenden Fremdsprachen gerecht zu werden. Es zeichnet sich insgesamt ab, dass das MAIN-TEACH-Modell mit seiner Berücksichtigung der Auswahl von Inhalten und fachspezifischen Methoden insgesamt geeigneter sein könnte, die inhaltlichen und unterrichtsmethodischen Spezifika eines fachmethodisch reflektierten, inhaltsorientierten Fremdsprachenunterrichts zu erfassen.

4 Perspektiven

Bei dem vorliegenden Entwurf handelt es sich, wie in den methodischen Vorbemerkungen des Abschnitts 1 dargelegt, um hermeneutisch fundierte, aber letztlich subjektive Vorschläge. Diese Vorschläge haben innerhalb der romanistischen Fachdidaktik, in der es bisher noch keine systematischen Entwürfe, geschweige denn empirische Studien zur Unterrichtsqualität gibt, sondierenden Charakter. Weitere theoretische und empirische Forschung ist erforderlich, um das Konstrukt der Unterrichtsqualität für die romanischen Sprachen insbesondere in ihrer Position als zweite und gerade auch dritte und spät beginnende Fremdsprachen zu fassen (vgl. dazu auch die exemplarischen Fragestellungen bei Lindl et al., in diesem Band). Dabei könnten unter anderem auch die Auswahl und die Darstellung sprachlicher Inhalte, aber beispielsweise ebenfalls die Auswahl der kulturellen Inhalte oder die Entwicklung einer mehrsprachigkeitsdidaktischen Progression reflektiert und weiter ausgestaltet werden. Mit Blick auf Unterrichtsqualität durch kulturelle Inhalte würde sich auch die Frage nach einer (Re-)Kanonisierung der Inhalte und ihrem Verhältnis zur aktuellen Kompetenzorientierung stellen. Zudem könnten und müssten Qualitätsmerkmale mit Blick auf romanische Sprachen als Herkunftssprachen im Fremdsprachenunterricht weiter definiert werden (vgl. dazu z. B. für das Fach Russisch Tark, in diesem Band). Nicht zuletzt wäre die Qualität von herkunftssprachlichem Unterricht in den romanischen Sprachen, der in den letzten Jahren wieder an Bedeutung zu gewinnen scheint, spezifisch zu modellieren. Weiterhin scheint eine Inbezugsetzung der Dimensionen „Unterrichtsqualität“ und „Kompetenzen der Lehrkräfte“ mit Bezug auf den Unterricht

in den romanischen Sprachen sinnvoll (allgemein vgl. z.B. Blömeke et al., 2015; Krauss et al., 2020; Lindl et al., in diesem Band).

Literaturverzeichnis

- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies. Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Charalambous, C. Y. & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, 67, 100894. <https://doi.org/10.1016/j.stueduc.2020.100894>
- Gröschner, A. & Kleinknecht, M. (2013). Qualität von Unterricht – Ansätze aus der Perspektive der Unterrichtsforschung. In L. Haag, S. Rahm, H. J. Apel & W. Sacher (Hrsg.), *Studienbuch Schulpädagogik* (S. 162–177). Bad Heilbrunn: Klinkhardt.
- Haag, L. & Streber, D. (2013). Klassenführung. In L. Haag, S. Rahm, H. J. Apel & W. Sacher (Hrsg.), *Studienbuch Schulpädagogik* (S. 221–242). Bad Heilbrunn: Klinkhardt. <https://doi.org/10.36198/9783838540580>
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (7. Aufl.). Seelze: Kallmeyer-Klett.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 345–363). Weinheim: Beltz. <https://doi.org/10.25656/01:3149>
- Hessischer Landtag (2021). *Entschließungsantrag Fraktion CDU, Fraktion Bündnis 90 / Die Grünen: Mehrsprachigkeit als Gewinn – Hessen baut Angebot der 2. und 3. Fremdsprachen an Schulen aus*. Hessischer Landtag, Drucksache 20/5547 vom 20.04.2021. Verfügbar unter <https://starweb.hessen.de/cache/DRS/20/7/05547.pdf>
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–58). Bonn: BMBF.
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N. & Blum, W. (2020). Competence as a continuum in the COACTIV study: The “cascade model”. *ZDM – Mathematics Education*, 52, 311–327. <https://doi.org/10.1007/s11858-020-01151-z>
- Krechel, H. L. (2016). Erfolgreiche und ergebnisorientierte Unterrichtsgestaltung. In H. L. Krechel (Hrsg.), *Referendariat Französisch. Kompaktwissen für Berufseinstieg und Examensvorbereitung* (S. 19–52). Berlin: Cornelsen.
- Lipowsky, F. (2020). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (3. Aufl., S. 69–118). Heidelberg: Springer. https://doi.org/10.1007/978-3-662-61403-7_4
- Meyer, H. (2020). *Was ist guter Unterricht?* (15. Aufl.). Berlin: Cornelsen.
- Ministerium für Bildung, Wissenschaft und Kultur Schleswig-Holstein (2018). *Fachanforderungen Italienisch*. Kiel.
- Ministerium für Schule und Berufsbildung Schleswig-Holstein (2015a). *Fachanforderungen Spanisch*. Kiel.
- Ministerium für Schule und Berufsbildung Schleswig-Holstein (2015b). *Fachanforderungen Französisch*. Kiel.

- Müller, K. & Rohling, B. (2016). *Referendariat Spanisch. Kompaktwissen für Berufseinstieg und Examensvorbereitung*. Berlin: Cornelsen.
- Piepho, H.-E. (1974). *Kommunikative Kompetenz als übergeordnetes Lernziel im Englischunterricht*. Dornburg-Frickhofen: Frankonius.
- Plikat, J. (2018). Gegenwärtige Leitkonzepte. In A. Grünewald & L. Küster (Hrsg.), *Fachdidaktik Spanisch* (S. 131–150). Stuttgart: Klett.
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality. The German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Prusse-Hess, B. & Prusse, M. C. (2018). *Wirksamer Englischunterricht*. Baltmannsweiler: Schneider Hohengehren.
- Reimann, D. (2011). Diatopische Varietäten des Französischen, Minderheitensprachen und Bilinguismus im transkulturellen Fremdsprachenunterricht. In M. Frings & F. Schöpp (Hrsg.), *Varietäten im Französischunterricht* (S. 123–168). Stuttgart: ibidem.
- Reimann, D. (2013). Transkulturelle kommunikative Kompetenz – ein neues Paradigma für den Fremdsprachenunterricht? In M. Franke & F. Schöpp (Hrsg.), *Auf dem Weg zu kompetenten Schülerinnen und Schülern. Theorie und Praxis eines kompetenzorientierten Fremdsprachenunterrichts im Dialog* (S. 165–180). Stuttgart: ibidem.
- Reimann, D. (2014). Italienisch – ein vielschichtig thematisches Schulfach. In H. Lange & S. Sinning (Hrsg.), *Forschungs- und Lehrzusammenhang Themenkonstitution 10: Kommunikation und Verstehen. Fachdidaktik und Themenkonstitution in den sprach- und kommunikationsbezogenen Fächern und Lernbereichen* (S. 79–111). Baltmannsweiler: Schneider Hohengehren.
- Reimann, D. (2015a). Aufgeklärte Mehrsprachigkeit – neue Wege (auch) für den Spanischunterricht, Basisartikel. *Der fremdsprachliche Unterricht Spanisch*, 51, 4–11.
- Reimann, D. (2015b). ELE (Spanisch als Fremdsprache) für Lernende mit L1 Deutsch aus kontrastiv-linguistischer Perspektive. In M. Meliss & B. Pöll (Hrsg.), *Aktuelle Perspektiven der kontrastiven Sprachwissenschaft: Deutsch-Spanisch-Portugiesisch* (S. 193–217). Tübingen: Narr.
- Reimann, D. (2016). Aufgeklärte Mehrsprachigkeit – Sieben Forschungs- und Handlungsfelder zur (Re-)Modellierung der Mehrsprachigkeitsdidaktik. In M. Rückl (Hrsg.), *Sprachen & Kulturen vermitteln und vernetzen. Mehrsprachigkeit und Inter-/Transkulturalität im Sprachenunterricht und in der Lehrer_innenbildung* (S. 15–33). Münster: Waxmann.
- Reimann, D. (2017a). Portugiesischunterricht in Deutschland als Beitrag zu Mehrsprachigkeit und transkultureller Identitätsbildung. Historische Entwicklung und gegenwärtige Perspektiven. In ders., *Transkulturelle kommunikative Kompetenz in den romanischen Sprachen. Theorie und Praxis eines neokommunikativen und kulturell bildenden Französisch-, Spanisch-, Italienisch- und Portugiesischunterrichts* (S. 247–286). Stuttgart: ibidem.
- Reimann, D. (2017b). Transkulturelle kommunikative Kompetenz im Unterricht der romanischen Sprachen. In ders., *Transkulturelle kommunikative Kompetenz in den romanischen Sprachen. Theorie und Praxis eines neokommunikativen und kulturell bildenden Französisch-, Spanisch-, Italienisch- und Portugiesischunterrichts* (S. 11–94). Stuttgart: ibidem.
- Reimann, D. (2017c). Inter- und transkulturelles Lernen im romanischen Tertiärsprachen- und im altsprachlichen Unterricht (Fokus: Spanisch, Italienisch, Griechisch). In M. Frings, S. Paffenholz & K. Sundermann (Hrsg.), *Vernetzter Sprachunterricht. Die Schulfremdsprachen Englisch, Französisch, Griechisch, Italienisch, Latein, Russisch und*

- Spanisch im Dialog. Akten einer Fortbildungsreihe des Bildungsministeriums und des Pädagogischen Landesinstituts Rheinland-Pfalz* (S. 217–235). Stuttgart: ibidem.
- Reimann, D. (2017d). Rezeptive Varietätenkompetenz im Spanischen – Modellierung einer Teilkompetenz zwischen funktionaler kommunikativer Kompetenz und Sprachbewusstheit. In E. Leitzke-Ungerer & C. Polzin-Haumann (Hrsg.), *Varietäten des Spanischen im Fremdsprachenunterricht* (S. 69–95). Stuttgart: ibidem.
- Reimann, D. (2019). Inklusion im Fremdsprachenunterricht unter besonderer Berücksichtigung der romanischen Sprachen. Eine Standortbestimmung zehn Jahre nach Ratifizierung der UN-Behindertenrechtskonvention (2009–2019). *Aggiornamenti – Rivista dell' Associazione Docenti d'Italiano in Germania*, 15, 15–26.
- Reimann, D. (2020). Schülerinnen und Schüler mit zielsprachlichem Hintergrund im Fremdsprachenunterricht. Ergebnisse einer qualitativen Pilotierung (am Beispiel des Spanischen, mit Ausblicken auf weitere romanische Sprachen). In M. García García, M. Prinz & D. Reimann (Hrsg.), *Mehrsprachigkeit im Unterricht der romanischen Sprachen* (S. 211–257). Tübingen: Narr.
- Reimann, D. (2021). ‚Muttersprachler/innen‘ im Fremdsprachenunterricht. Ergebnisse einer Studie zu den Fächern Spanisch, Portugiesisch und Italienisch mit Blick auf die Perspektive der Schüler/innen, *state of the art* der unterrichtsmethodischen Forschung und Anregungen für die Praxis. In C. Helmchen, S. Melo-Pfeifer & J. von Rosen (Hrsg.), *Mehrsprachigkeit in der Schule. Ausgangspunkte, unterrichtliche Herausforderungen und methodisch-didaktische Zielsetzungen* (S. 169–196). Tübingen: Narr.
- Reimann, D. (2022). Porquê uma didática do Português como língua estrangeira (Fachdidaktik Portugiesisch) na Alemanha? Para a fundação de uma disciplina. In L. Coelho (Hrsg.), *Portugal – Alemanha. Convergências e divergências* (S. 257–283). São Paulo/Lisboa: Lisbon International Press.
- Reimann, D. (2023). *Kompendium Fachdidaktik Romanistik. Französisch – Italienisch – Spanisch. I: Grundlagen*. Tübingen: Narr. <https://doi.org/10.24053/9783381102426>
- Reinfried, M. (2001). Neokommunikativer Fremdsprachenunterricht: Ein neues methodisches Paradigma. In F.-J. Meißner & M. Reinfried (Hrsg.), *Bausteine für einen neokommunikativen Französischunterricht. Lernerzentrierung, Ganzheitlichkeit, Handlungsorientierung, Interkulturalität, Mehrsprachigkeitsdidaktik* (S. 1–20). Tübingen: Narr.
- Reinfried, M. (2006). Im Rückspiegel: Die „großen“ Methoden. In A. Nieweler (Hrsg.), *Fachdidaktik Französisch* (S. 38–43). Stuttgart: Klett.
- Reinfried, M. (2017a). Ein Rückblick auf die „großen“ Methoden des Fremdsprachenunterrichts. In A. Nieweler (Hrsg.), *Fachdidaktik Französisch* (S. 68–73). Stuttgart: Klett.
- Reinfried, M. (2017b). Didaktisch-methodische Prinzipien heute. In A. Nieweler (Hrsg.), *Fachdidaktik Französisch* (S. 74–84). Stuttgart: Klett.
- Reinfried, M. & Volkmann, L. (2012). Medien im neokommunikativen Fremdsprachenunterricht: Einsatzformen und Nutzungsmöglichkeiten. In dies. (Hrsg.), *Medien im neokommunikativen Fremdsprachenunterricht* (S. 9–39). Frankfurt am Main: Lang. <https://doi.org/10.3726/978-3-653-02288-9>
- Statistisches Bundesamt. (2022). *Schüler mit Fremdsprachen-Unterricht: Deutschland, Schuljahr, Fremdsprachen, Schularzt, Jahrgangsstufen* [Ergebnis 21111-0006]. Verfügbar unter <https://www-genesis.destatis.de/genesis/online?operation=abrufabelleBearbeiten&levelindex=2&levelid=1673953254902&auswahloperation=abrufabelleAuspraegungAuswaehlen&auswahlverzeichnis=ordnungsstruktur&auswahlziel=werteabruf&code=21111-0006&auswahltext=&werteabruf=Werteabruf#abreadcrumb>
- Steveker, W. (2019). Zeitgemäß unterrichten. In K. Sommerfeldt (Hrsg.), *Spanisch Methodik. Handbuch für die Sekundarstufe I und II* (6. Aufl., S. 23–48). Berlin: Cornelsen.

- Thaler, E. (2014). Kognitive Aktivierung & Hattie & Fremdsprachenunterricht. *Praxis Fremdsprachenunterricht – Basisheft*, 6, 5–6.
- Thaler, E. (2019). Wirksamer Fremdsprachenunterricht. *Praxis Fremdsprachenunterricht – Basisheft*, 6, S. 4–5.
- Wiater, W. (2020a). *Unterrichtsplanung* (4. Aufl.). Augsburg: Auer.
- Wiater, W. (2020b). *Unterrichtsprinzipien* (8. Aufl.). Augsburg: Auer.
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>

Sonja Brunsmeier

Zusammenspiel von generischen und fachspezifischen Aspekten

Ein Konzept für aufgabenbasiertes Lehren und Lernen
im digital gestützten Fremdsprachenfrühbeginn

Abstract (der Herausgeber:innen)

In her article “Interplay of generic and subject-specific aspects – A concept for task-based teaching and learning in digitally-supported early foreign language learning,” Sonja Brunsmeier presents the findings of the research project ‘EMMa’, which yielded a concept for digitally-supported task-based teaching and learning. The author’s focus is on early English language learning and the role of digital media and materials in advancing English teaching and learning at the primary level. In terms of instructional quality, her contribution aims to demonstrate how subject-specific (= English competences) and generic competences (e.g., media competences) can be promoted through the appropriate design of early foreign language teaching.

1 Einleitung

Digitale Medien und deren Einsatz im schulischen Kontext sind in aller Munde. Gemäß den Vorgaben der Kultusministerkonferenz (2016, 2021) ist die Ermöglichung einer selbstbestimmten Teilhabe aller Schüler:innen an unserer digital geprägten Gesellschaft Aufgabe aller Schulstufen und -fächer gleichermaßen. Dabei werden häufig insbesondere allgemeine Aspekte – wie z. B. die Motivation der Schüler:innen (z. B. Alhinty, 2015, S. 26; Aufenanger, 2017, S. 126) – als Argumente für den Medieneinsatz im Unterricht angeführt. Wie digitale Medien und Materialien fachspezifisch ausgewählt und qualitativvoll eingesetzt werden können, um fachliche Kompetenzen zu entwickeln, ist jedoch bislang wenig erforscht (Welling, 2017, S. 17–18). Auf der Basis von empirischen Erkenntnissen (Brunsmeier, 2023) aus dem Aktionsforschungsprojekt ‚EMMa‘ (Englisch lehren und lernen mit digitalen Medien und Materialien) versucht das Kapitel – unter Berücksichtigung der Fachkultur und -spezifik – aufzuzeigen, welche Dimensionen von Unterrichtsqualität besonders relevant für die Gestaltung des Grundschulenglischunterrichts sind. Das Forschungsprojekt ‚EMMa‘ generiert ein Konzept für digital gestütztes aufgabenbasiertes Lehren und Lernen, das illustriert, wie fachspezifische

(hier Englischkompetenzen) und generische Kompetenzen (hier Medienkompetenzen) im Fremdsprachenfrühbeginn entwickelt werden können.

2 Generische und fachspezifische Aspekte

Eine stärkere Berücksichtigung unterschiedlicher Fachkulturen in der Forschung zu Unterrichtsqualität ermöglicht, die drei Basisdimensionen um sowohl generische als auch fachspezifische Merkmale zu erweitern und zu konkretisieren (Lindl et al., in diesem Band). Durch einen fachlichen Fokus und Betrachtung einer bestimmten Unterrichtssequenz zu einem gewissen Thema oder einer konkreten Unterrichtsaktivität entsteht zwangsläufig ein Spannungsfeld zwischen Spezifität und Generalisierbarkeit. Lindl et al. (in diesem Band) schlagen vor, dies als Kontinuum zu begreifen und die daraus resultierenden Implikationen in der Forschung und bei der Unterrichtsgestaltung zu beachten. Dieses Kapitel leistet einen Beitrag aus der fachlichen Perspektive des Grundschulenglischunterrichts. Es wird das Forschungsprojekt ‚EMMa‘ vorgestellt und anhand dessen diskutiert, wie sowohl fachspezifische Kompetenzen (hier Englischkompetenzen) als auch fächerübergreifende Kompetenzen (hier Medienkompetenzen) durch einen fachunspezifischen Ansatz (hier aufgabenbasiertes Lehren und Lernen) in einem bestimmten Kontext (hier Grundschule) entwickelt werden können. Dabei eröffnet das Kapitel eine zusätzliche Perspektive auf Unterrichtsqualität, indem es Oberflächenstrukturen und deren etwaige Wechselwirkungen auf Tiefenstrukturen erkundet.

2.1 Ziel: Medienkompetenzen

„Die Digitalisierung und Mediatisierung haben in allen Bereichen unserer Lebens- und Arbeitswelt zu entscheidenden Veränderungen geführt. Diese gehen über einen rein technischen Fortschritt hinaus und führen zu einem breit angelegten kulturellen und gesellschaftlichen Wandel, der sich auf das schulische Lehren und Lernen und auf die Bewältigung und Gestaltung von Lebens- bzw. Arbeitsprozessen von Kindern, Jugendlichen und jungen Erwachsenen auswirkt“ (KMK, 2021, S. 3). Mit diesen Sätzen eröffnet die Kultusministerkonferenz ihre Ergänzung zur Strategie der Kultusministerkonferenz *Bildung in der digitalen Welt* (KMK, 2016) und bekräftigt damit nochmals die Notwendigkeit, sich im gesamten schulischen Bereich fortlaufend mit der sich stetig verändernden digitalen Realität auseinanderzusetzen. Dabei nimmt die Kultusministerkonferenz alle Schulstufen und Schularten gleichermaßen in die Verantwortung, digital gestützte Lehr-Lern-Prozesse zu planen und zu gestalten. Das Ziel ist, „den Lernenden eine mündige, souveräne und aktive Teilhabe an der digitalisierten Lebens- und Arbeitswelt [zu] ermöglichen“ (KMK, 2021, S. 6). Explizit wird in dem Strategiepapier betont, dass digital gestützte Lernprozesse „schon für Unterrichtsprozesse ab

Beginn der Primarstufe relevant sind“ (KMK, 2021, S. 3). Sowohl nationale Studien (z. B. mpfs, 2021; vbw, 2021; Vodafone Stiftung, 2020) als auch internationale Studien (z. B. Fraillon et al., 2018; Laranjeiro, 2021) bestätigen die digital geprägte Lebenswelt der jungen Lernenden. Gleichzeitig wird jedoch auch darauf hingewiesen, „that there is a serious digital skills deficit amongst Europe’s children, despite the popular view that they are „digital natives“ (European Commission, 2012, S. 5). Der Zugang zu digitalen Medien und Materialien ist also nicht automatisch gleichzusetzen mit einer kompetenten, zielgerichteten und kritisch-reflexiven Nutzung dieser. Gerade die entwicklungspsychologischen Voraussetzungen dieser jungen Altersgruppe müssen in einer digital geprägten Welt bedacht werden: „While the Internet, the devices used to access it and the services available today were not created specifically for use by children, research shows that they are increasingly prominent users. Children therefore form a specific group requiring attention“ (European Commission, 2012, S. 15). Ein Blick in die Forschungslandschaft zeigt, dass insbesondere für den Grundschulbereich und das fachspezifische Lehren und Lernen in digital gestützten Lernsettings erheblicher Forschungsbedarf besteht: „[R]esearch tend to neglect young learners“ (Whyte & Cutrim-Schmid, 2019, S. 345) „[and s]o far, the effectiveness of tablet PCs for FL learning has not been extensively studied“ (Golonka et al., 2014, S. 86). Insgesamt wird (fast) durchgängig gefordert, den Fokus von den Technologien an sich auf die Einbettung dieser im Unterricht zu verschieben und darauf, dass sich zukünftige Forschungsarbeiten der Komplexität schulischer Lehr- und Lernprozesse widmen: „Aus forschungsmethodischer Sicht ist anzumerken, dass weitere, differenzierte Studien notwendig sind, um Lehrkräfte sowie Schülerinnen und Schüler besser von den Tablets profitieren zu lassen. Das heißt zum Beispiel, lernrelevante Schülerdaten einzubeziehen, Lern- und Lehrstile zu untersuchen oder insgesamt mehr Unterrichtsbeobachtungen durchzuführen, um konkrete Einsatzszenarien besser bewerten zu können“ (Aufenanger, 2017, S. 134). Insbesondere für die Erforschung fachspezifischer Lehr- und Lernprozesse (Lindl et al., in diesem Band) bedarf es der Verschränkung von Sicht- und Oberflächenstrukturen mit unterrichtlichen Tiefenstrukturen und Prozessmerkmalen (z. B. Decristan et al., 2020; Drechsel & Schindler, 2019; Kunter & Trautwein, 2013). Das Forschungsprojekt ‚EMMa‘ beschäftigt sich mit Auswahl und Einsatz digitaler Technologien und der Konzeption von Aufgaben (= Sicht- und Oberflächenstrukturen) sowie der tatsächlichen Implementierung der Aufgaben im Englischunterricht und der Beobachtung daraus resultierender Lehr- und Lernprozesse im kommunikativen Englischunterricht der Grundschule (= Tiefenstrukturen und Prozessmerkmale).

2.2 Ansatz: Aufgabenbasiertes Lehren und Lernen

Im Fremdsprachenunterricht wird Aufgabenorientierung (*task-based language learning and teaching*) international nicht nur als „a well-recognised approach to L2 learning and pedagogy“ (Butler, 2019, S. 307), sondern sogar als „State of

the Art“ (Dreßler, 2018, S. 76) des Lehrens und Lernens einer fremden Sprache angesehen (z. B. Kolb & Schocker, 2021; Meyer, Volkmann & Grimm, 2022; Pinter, 2015). Lindl et al. (in diesem Band) führen Untersuchungen an, die beispielhafte fachspezifische Aspekte von Unterrichtsqualität in Englisch vorlegen. Unter anderem werden explizit Aufgaben (siehe „Aufgabenorientierung“, Helmke et al., 2008, bzw. „Cognitively Stimulating Tasks/Activities“, Kersten et al., 2018) als fachspezifisches Unterrichtsmerkmal, das mit der Qualität von Englischunterricht in Verbindung steht, genannt. Dreßler (2018, S. 119–120) identifiziert die folgenden Aufgabenmerkmale als relevant für den Grundschulenglischunterricht in Deutschland: Aufgaben

- sind lernendenzentriert, basieren auf bedeutungsvollen Erfahrungen und persönlichen Interessen der Lernenden und haben einen besonderen Bezug zu deren Lebenswelt.
- sind Aktivitäten, bei denen die Sprache auf natürliche Weise zur Bewältigung kommunikativer Situationen in einem für die Lernenden bedeutungsvollen Kontext eingesetzt wird.
- fördern/entwickeln die persönliche Entwicklung auf mehreren Ebenen (kognitiv, pädagogisch, persönlich, emotional) und das aktive Engagement in einer positiven Lernumgebung.
- helfen den Lernenden, unabhängiger und selbstgesteuerter zu lernen (persönliche Interessen, Lernendenzentrierung) und lassen verschiedene Lösungsmöglichkeiten zu. Dadurch wandeln sich traditionelle Rollen der Lernenden (z. B. statt passiver Teilhabe, müssen die Lernenden für ihr Lernen verantwortlich werden) und Lehrenden (z. B. statt ausschließlich als Wissensvermittler:innen zu agieren, wirken sie als Motivator:innen und Vermittler:innen).
- verlangen kritisches Denken und Reflexion, Problemlösung und die Teilnahme an gesellschaftlichen Diskursen.
- beziehen die Lernenden in die Aushandlung von Bedeutungen ein und erfordern, dass die Sprache als Mittel zur Kommunikation genutzt wird (z. B. in Gruppen/Partner:innenarbeit, *information gap activities*).
- basieren auf allgemeinen grundschulpädagogischen und fremdsprachenspezifischen Prinzipien (z. B. Spiralcurriculum, Wiederholung, Routinen, Übungen und Transferaktivitäten, Übergang vom Unterricht auf konkreter zu abstrakter Ebene entsprechend der kognitiven Entwicklung der Kinder).
- forcieren das Primat *meaning before accuracy*.
- motivieren durch bedeutungsvolle Erfahrung und den Einbezug persönlicher Interessen der Lernenden.
- bieten Kohärenz und Einheitlichkeit.
- definieren klare (Sprachlern-)Ziele.
- haben einen Anfang und ein Ende.

Der Sequenzierung von Aufgaben kommt beim aufgabenbasierten Lehren und Lernen eine entscheidende Rolle zu: „In general, the different TBLT advocates (...) agree on a basic tripartite approach“ (Dreßler, 2018, S. 84). Viel zitiert und

als „groundbreaking“ (Schocker-von Ditfurth & Müller-Hartmann, 2010, S. 91) angesehen, wird das TBA-Modell von Willis (2005, S. 38, siehe auch Abb. 1), das aufgrund seiner eingängigen Form auch als Stoppzeichen (*Stop-Sign*) bezeichnet wird. Auch Willis' Adaption ihres eigenen Modells für junge Lernende (Willis, 2005, S. 116–131, siehe auch Abb. 2) gliedert sich in drei Phasen: *Pre-Task Phase*, *Task Cycle*, welcher in anderen Publikationen als *Core Activity* (z. B. Cameron, 2001, S. 32) bezeichnet wird, und *Post-Task Phase*. Ziel der *Pre-Task Phase* ist die sprachliche und inhaltliche Vorbereitung der Schüler:innen auf die zu bearbeitende Aufgabe (Willis, 2005, S. 42–50). In der *Core Activity* / dem *Task Cycle* bearbeiten die Kinder die Aufgabe(n) (Willis, 2005, S. 52–65). Für die *Post-Task Phase* sieht Willis (2005, S. 101–115) einen *Language Focus* vor, der sprachliche Analysen und Übungen beinhaltet. Digitale Medien und Materialien verändern Lernsettings.

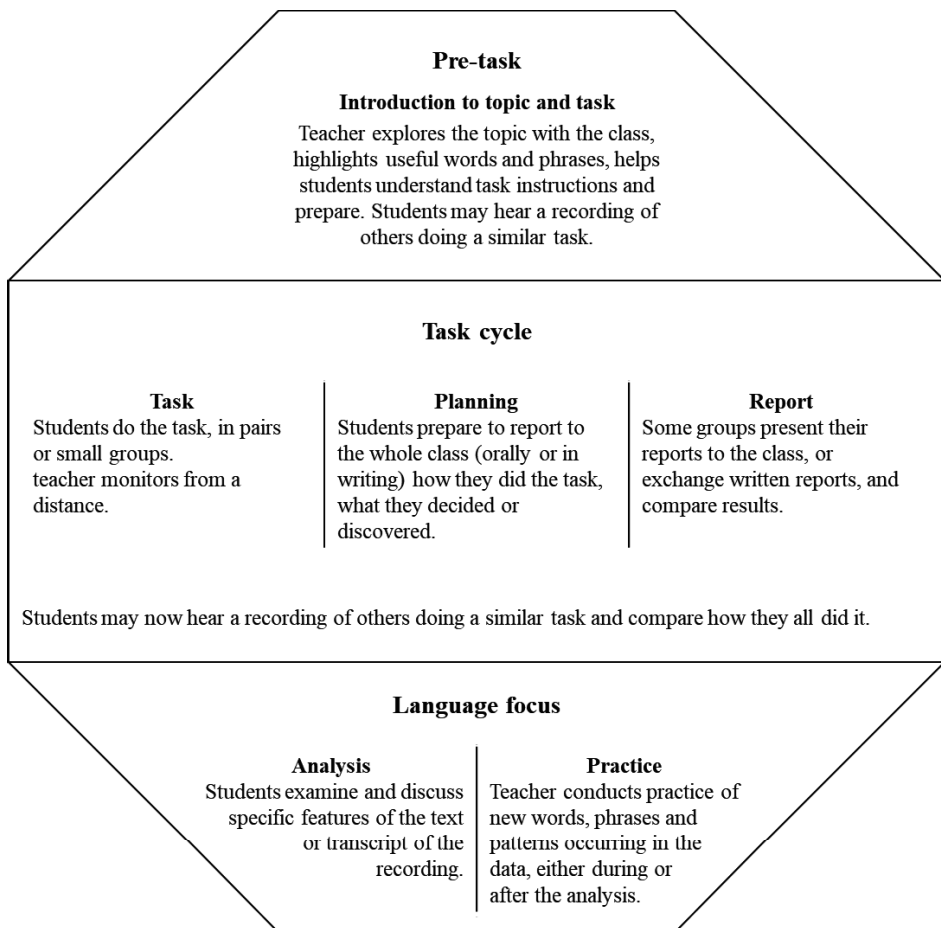


Abbildung 1: Elemente des TBA-Modells (Willis, 2005, S. 38)

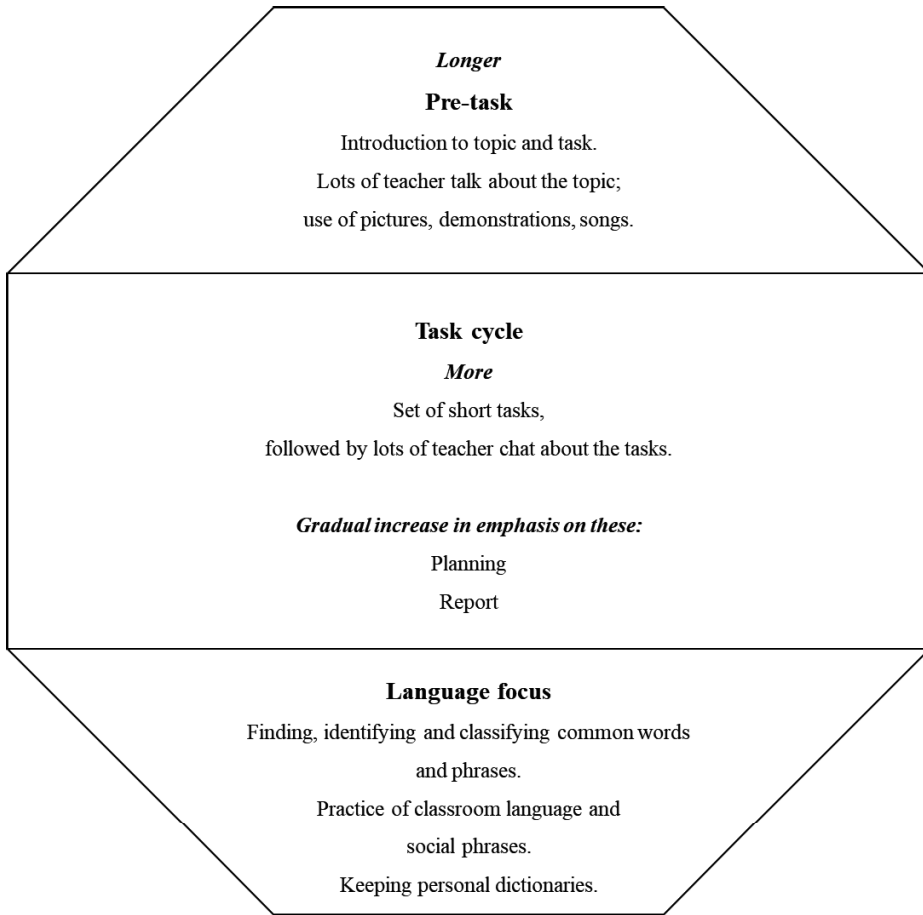


Abbildung 2: TBA-Modell für Anfänger:innen und junge Lernende (Willis, 2005, S. 116)

Das bereits 1996 entwickelte TBA-Modell (Willis, 2005) wurde bislang noch nicht in Bezug auf Möglichkeiten und Herausforderungen, die durch digitale Prozesse bei der Vermittlung fremdsprachlicher Kompetenzen im Grundschulenglischunterricht bedacht werden müssen (Brunsmeier & Holberg, 2022), reflektiert. „To maximize digital technology for language education, understanding technology’s pros and cons is as important as using it properly and strategically for a given purpose“ (Butler, 2022, S. 142). Das beinhaltet auch, etablierte Ansätze wie das aufgabenbasierte Lehren und Lernen weiterzuentwickeln. Genau hierzu besteht aktuell dringender Forschungsbedarf:

„[T]he contexts for learning and the resources available have expanded dramatically over the last few decades. In order to realize the full potential of technology-mediated TBLT, tasks need to be designed (...) [and] future research should examine (...) task design and implementation using new and emerging media“ (Ziegler, 2016, S. 140, 142).

Dies gilt insbesondere auch in Bezug auf junge Lernende: „Technology-mediated ELT research with young learners is in its infancy” (Pinter, 2019, S. 151).

2.3 Kontext: Fremdsprachenfrühbeginn

Der Fremdsprachenunterricht hebt sich insofern von anderen Schulfächern (wie z.B. Mathematik, Sport, Musik) ab, als die entsprechende Fremdsprache Medium und Ziel des Unterrichts zugleich ist. Aspekte, die erheblich zur Effizienz des Englischunterrichts in der Grundschule beitragen, wurden beispielsweise im Jahr 2015 in der deutschlandweiten Studie des BIG-Kreises (2015) ermittelt. Dazu wurden die Fertigkeiten der Schüler:innen in den Bereichen Hör- und Leseverstehen, Schreiben und Sprechen untersucht. Auf Basis dieser empirischen Erkenntnisse betonen die Verantwortlichen (BIG-Kreis, 2015, S. 72) insbesondere die Bedeutung des Einsatzes fremdsprachlich ausgebildeter Lehrpersonen mit sehr guten fremdsprachlichen und didaktischen Qualifikationen sowie die Erhöhung der Kontaktzeit mit der Fremdsprache (d.h. eine Erhöhung der Stundenzahl bzw. kleinerer Englischeinheiten wie z.B. Integration englischer Rituale, wie englischsprachige Lieder, englischsprachige Reime etc. in den Schulalltag). Auch Frisch (2017, S. 59, 61) sieht diese beiden Aspekte als bedeutsam für die Qualität des Grundschulenglischunterrichts an. Darüber hinaus sind aber auch noch „die Qualität des Zugangs zur englischen Sprache, die Sprachlernmotivation der Lernenden, (...) die didaktisch-methodische Ausrichtung des Unterrichts und die Kontinuität des Lehrprogramms“ (Frisch, 2017, S. 55) entscheidend. Zudem macht sie darauf aufmerksam, dass standardisierte Abschlussprofile, deren Entwicklung für den Englischunterricht in der Grundschule nach wie vor aussteht, einen maßgeblichen Beitrag zu einheitlichen Leistungsanforderungen und Zieltransparenz schaffen könnten (Frisch, 2017, S. 56–59) und es insgesamt einer Berücksichtigung der „Bedürfnisse und Besonderheiten der Lerner im Alter von sechs bis zehn Jahren“ bedarf (Frisch, 2017, S. 61).

2.4 Unterrichtsqualität als ein komplexes Zusammenspiel generischer und fachspezifischer Aspekte begreifen

Wilden (2021, S. 212), die als eine der ersten fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch beleuchtet hat, kommt zu dem Schluss, dass es „[e]ine umfassende, kohärente und für unterschiedliche Lernsituationen und Lernende passende Theorie für ‚guten Englischunterricht‘ (...) wohl nicht geben [wird,] angesichts dessen, dass es eine Reihe von Einflussfaktoren gibt, die in wechselnden Konstellationen unterschiedlich einflussreich sind.“ Statt also auf die Suche nach den Aspekten und dem ‚Erfolgsrezept‘ zu gehen, stellt sich vielmehr die Aufgabe fortlaufend und immer wieder aufs Neue zu verstehen, zu erkennen und zu reflektieren, (1) welche Einflussfaktoren (2) wie und (3) unter welchen

Umständen und (4) in welchen Kontexten (nicht) wirken. Genau diesen Versuch unternimmt dieses Kapitel: Anhand des Forschungsprojekts ‚EMMa‘ diskutiert dieser Beitrag, wie sowohl fachspezifische Kompetenzen (d.h. Englischkompetenzen) als auch fächerübergreifende Kompetenzen (d.h. Medienkompetenzen) durch einen fachunspezifischen Ansatz (aufgabenbasiertes Lehren und Lernen) in einem bestimmten Kontext (d.h. Grundschule) entwickelt werden können.

Gemäß dem MAIN-TEACH-Modell (Multi-layered and integrated in conceptualizing the quality of teaching; Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022; Lindl et al., in diesem Band) werden im Forschungsprojekt ‚EMMa‘ explizit generische und fachspezifische Aspekte von Unterrichtsqualität thematisiert und hinsichtlich ihrer Funktionen und Wechselwirkungen im Lernprozess beleuchtet. Unterricht, der im MAIN-TEACH-Modell durch das Wechselspiel zwischen unterrichtlichem Angebot auf Lehrkraftseite und dessen Nutzung auf Schüler:innenseite dargestellt wird, findet aber immer in einem spezifischen Kontext (z.B. Schulart, Fach, Ausstattung des Klassenzimmers, Persönlichkeit der Lehrperson, individuelle Lerngruppe etc.) statt. Diese kontextspezifischen Aspekte werden im Aktionsforschungsprojekt ‚EMMa‘ besonders berücksichtigt.

3 EMMa

‚EMMa‘ ist ein Aktionsforschungsprojekt (Burns, 2010), welches sich den drei folgenden Fragestellungen widmet:

- 1) Was sind geeignete Kriterien zur Auswahl digitaler Medien und Materialien für die Entwicklung fremdsprachlicher Kompetenzen im Englischunterricht der Grundschule?
- 2) Welche Aufgaben und Aktivitäten eignen sich, um das Lernpotenzial digitaler Medien und Materialien auszuschöpfen?
- 3) Welche Funktionen digitaler Medien und Materialien haben das Potenzial, welche fremdsprachlichen Kompetenzen (weiter) zu entwickeln?

Dieses Kapitel konzentriert sich ausschließlich auf die zweite Forschungsfrage. Dabei sind wesentliche Motive der Aktionsforschung, die Qualität der Arbeit in einem Praxisbereich zu verbessern (Altrichter et al., 2018, S. 11) und „Theorie und Praxis in der Forschung untrennbar miteinander zu verbinden“ (Caspari, 2016, S. 72). Entsprechend wird die Studie im schulischen Kontext durchgeführt, nämlich in einer Englisch-AG an einer Grundschule. Denn „bei der Bearbeitung komplexer praktischer Probleme [zeigt sich, dass] die Spezifität des Kontextes besonderes Augenmerk erfordert. Den Aktionsforscher:innen geht es zunächst einmal um situatives Verstehen und nicht sogleich darum, allgemeine Aspekte der Situation heraus zu präparieren“ (Altrichter et al., 2018, S. 13). Die Aktionsforschung „ist als längerfristiger, zyklischer Prozess angelegt, innerhalb dessen (...) theoretische Annahmen zur Veränderung der Praxis im praktischen Handeln überprüft werden und nach erneuter Reflexion in revidierten Praxisvorschlägen bzw. Ver-

änderungen der theoretischen Annahmen münden“ (Caspari, 2016, S. 73). Abbildung 3 zeigt das Forschungsdesign des ersten Aktionszyklus.

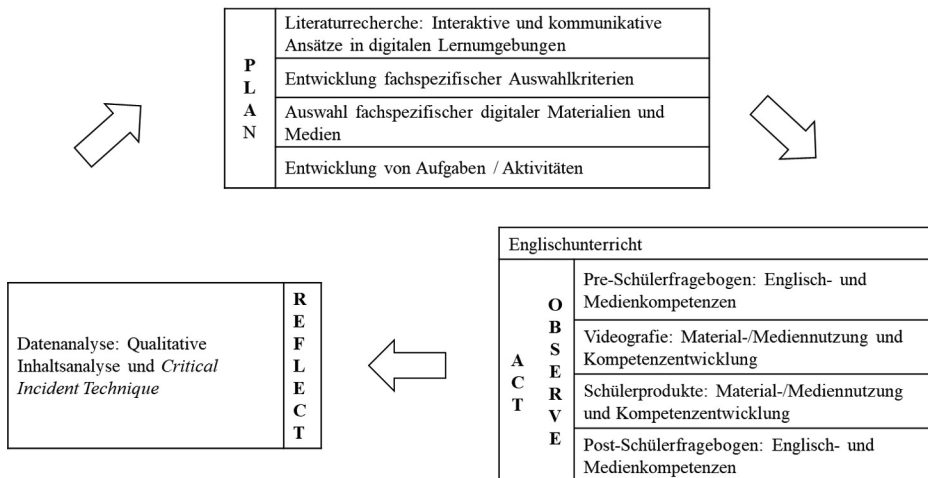


Abbildung 3: Forschungsdesign erster Aktionszyklus (Brunsmeier, 2023, S. 128)

Der erste Aktionszyklus fand im Zeitraum von Januar 2022 bis Mai 2022 als wöchentliches Nachmittagsangebot in Form einer Englisch-AG statt. Zum damaligen Zeitpunkt waren der Schulalltag, die Lehrer:innen und die Schüler:innen durch die pandemische Situation mit den vielen Unwägbarkeiten äußerst belastet: Aus diesem Grund führte die Autorin selbst die Englisch-AG durch. 16¹ Viertklässler:innen aus vier verschiedenen Parallelklassen im Alter zwischen neun und zehn Jahren nahmen an der Englisch-AG teil. Zum Startzeitpunkt der Englisch-AG hatten die Kinder bereits 1,5 Jahre Englischunterricht (Beginn des Englischunterrichts: Klasse 3).

3.1 Datenerhebung: 1. Aktionszyklus

Das explorativ-qualitativ angelegte Forschungsprojekt ‚EMMA‘ bedient sich ganz gezielt verschiedener Erhebungsinstrumente, um möglichst reichhaltige und vielfältige Einblicke zu gewinnen (Tabelle 1). Die verschiedenen Datensätze wurden dann in einer ‚Rückwärtsanalyse‘ (*Backward Analysis*; Brunsmeier, 2024) triangulierend ausgewertet (siehe Abschnitt 2.2).

¹ Die Anzahl der Plätze der AG war durch die Anzahl der zur Verfügung stehenden Tablets vorgegeben.

Tabelle 1: Datenerhebungsinstrumente/-dokumente und deren Ziele im 1. Aktionszyklus (vgl. Brunsmeier, 2023)

Instrumente / Dokumente	Ziele
Vorab-Fragebogen	Zu Beginn der Englisch-AG füllten die Schüler:innen einen Fragebogen aus. In diesem wurden die Kinder gebeten, Auskünfte zu ihrem Englischunterricht zu geben und darzulegen, welche Medien sie wie in ihrem (Schul-)Alltag nutzen. Auf diese Weise wurde erhoben, wie die Viertklässler:innen ihren Englischunterricht wahrnehmen sowie ihre Englisch- und Medienkompetenzen einschätzen.
Videografie	Jede Stunde des Nachmittagsangebots wurde mit zwei Videokameras aufgezeichnet. Dadurch konnte das komplexe Geschehen im Klassenzimmer (z. B. Interaktionen, nonverbales Verhalten, Reaktionen und Bearbeitungsprozesse der Schüler:innen) dokumentiert werden. Die aufgezeichneten Englischstunden bieten eine reichhaltige Datengrundlage (ca. 1.080 Minuten Videomaterial).
Schüler:innenprodukte	Alle Schüler:innenprodukte (z. B. Arbeitsblätter, Reflexionsbögen zu jeder thematischen Einheit, digitale Aufnahmen, digitale Produkte) wurden gesammelt, um Einblicke in Arbeitsprozesse und -ergebnisse der Schüler:innen zu erhalten.
Abschluss-Fragebogen	In der letzten Stunde der Englisch-AG erhielten die Viertklässler:innen einen Abschlussfragebogen. In diesem wurden die Kinder um eine Rückmeldung zur Englisch-AG sowie eine Einschätzung zur (Weiter) Entwicklung ihrer Englisch- und Medienkompetenzen gebeten.

3.2 Datenauswertung: 1. Aktionszyklus

Um den innovativen und lernförderlichen Einsatz digitaler Materialien und Medien für fremdsprachliche Lern- und Lehrprozesse im Englischunterricht der Grundschule zu analysieren, wurden die verschiedenen Datensätze (siehe Tabelle 1) mit der qualitativen Inhaltsanalyse (Mayring & Fenzl, 2019), der *Critical Incident Technique* (Grünewald, 2012, S. 59) und einfachen deskriptiven statistischen Verfahren (z. B. Häufigkeitsberechnungen) ausgewertet bzw. analysiert. Zudem wurden die Ergebnisse triangulierend (Elsner & Viebrock, 2014) in Beziehung zueinander gesetzt. Die verschiedenen Auswertungsverfahren wurden in einer ‚Rückwärtsanalyse‘ (*Backward Analysis*; Brunsmeier, 2024) kombiniert (siehe Abbildung 4).

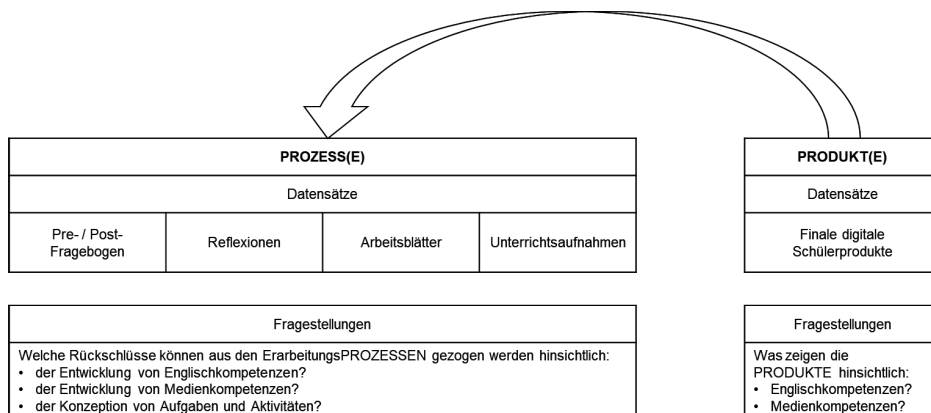


Abbildung 4: Ablauf der einer ‚Rückwärtsanalyse‘ (Backward Analysis; Brunsmeier 2024)

Ausgangspunkt der ‚Rückwärtsanalyse‘ waren die finalen (digitalen) Schüler:innenprodukte einer jeden thematischen Einheit. Für die Auswertung der Schüler:innenprodukte wurden Excel-Tabellen erarbeitet, in der die Englisch- und Medienkompetenzen für die entsprechende thematische Einheit sowie die technischen Eigenschaften der jeweils genutzten digitalen Medien und Materialien aufgeführt wurden. Entlang dieser Kriterien wurden die Schüler:innenprodukte ausgewertet.

Um zu verstehen, warum und wie diese Ergebnisse im Prozess (nicht) erreicht wurden, wurden weitere Daten, die im Unterrichtsprozess entstanden (d.h. Videoaufnahmen, Arbeitsblätter, Reflexionen etc.), ausgewertet und dann triangulierend (Elsner & Viebrock, 2014) in Beziehung zueinander gesetzt. Dazu wurden die Video- und Audioaufnahmen zunächst aufbereitet (d.h. mittels der Software F4 vollständig transkribiert). Grundlage der qualitativen Datenanalyse (Software MaxQDA) war ein deduktiv entwickeltes Kategoriensystem, in dem Erkenntnisse der Aufgabenforschung, Englisch- und Medienkompetenzen integriert wurden. Tabelle 2 zeigt einen Auszug aus dem Kategoriensystem.

Im Auswertungsprozess wurde das Kategoriensystem durch das Herausfiltern von *Critical Incidents* induktiv erweitert und ggf. verändert. Alle Datensätze wurden sowohl von der Autorin selbst als auch von zwei weiteren geschulten Hilfskräften kodiert und dann in Kodierkonferenzen kommunikativ validiert (Flick, 2019, S. 475–478).

Tabelle 2: Exemplarischer Auszug aus dem Kategoriensystem

1 PRE-TASK PHASE		
Kategorie	Unterkategorie	Beschreibung / Definition der Unterkategorie
Language	Activation of Prior Language	Das sprachliche Vorwissen (z. B. Wörter, Phrasen, Sätze) der Kinder wird aktiviert.
	Introduction of New Language	Neue Sprache (z. B. Wörter, Phrasen, Sätze) wird von der Lehrperson eingeführt und/oder gemeinsam mit den Kindern erarbeitet.
	Language Practice	Im Rahmen von verschiedenen (spielerischen) Übungen haben die Kinder die Gelegenheit, die (neuen) Wörter, Phrasen, Sätze etc. anzuwenden, zu sichern und zu festigen.
Content	Activation of Prior Knowledge	Das Vorwissen (z. B. Inhalte, Aufgabenstruktur) der Kinder wird aktiviert.
	Provision of Knowledge	Wissen (z. B. Inhalte, Aufgabenstruktur) wird von der Lehrperson bereitgestellt und/oder gemeinsam mit den Kindern erarbeitet.
	Introduction of Topic	Das Thema wird eingeführt und Motivation und Interesse für dieses geweckt.
	Setting the Task	Die Aufgabe wird gestellt, demonstriert und/oder besprochen. Dabei werden das Ziel und Produkt geklärt und/oder ausgehandelt.
Navigation and Handling of Digital Medium/Material	User Language	Sprache wird für die Navigation und Handhabung des digitalen Mediums (z. B. Tablet) und/oder digitalen Materials (z. B. App) vorentlastet und/oder anhand/mit diesen erarbeitet, um beispielsweise Buttons, Settings, Funktionen etc. zu verstehen.
	User Knowledge	Informationen zu den <i>features</i> , die für die Navigation und Handhabung des digitalen Mediums (z. B. Tablet) und/oder digitalen Materials (z. B. App) zur Erstellung des digitalen Produkts relevant sind, werden bereitgestellt und/oder erarbeitet.
	Organising Technology	Die Handhabung des Mediums (z. B. Tablet) und/oder des digitalen Materials (z. B. App) wird erläutert, demonstriert und/oder besprochen. Dabei wird der Umgang organisiert und/oder geklärt.

4 Ein Konzept für digital gestütztes aufgabenbasiertes Lehren und Lernen im Fremdsprachenfrühbeginn

Das etablierte TBA-Modell (Willis, 2005) wurde bislang jedoch noch nicht für den digital gestützten Englischunterricht adaptiert (siehe Abschnitt 1.2). Prozesse der Digitalisierung bedingen, dass neue Elemente ergänzt werden (z. B. ‚Navigation and Handling of Digital Medium/Material‘), bestehende Elemente sich

verändern (z. B. ‚Active Involvement‘) und neue Formen annehmen (z. B. ‚Task Support‘). Im Folgenden wird das adaptierte TBA-Modell (Abbildung 5) – und damit jede der drei Phasen – für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn detailliert beschrieben. Dabei wird auch modelliert, wie generische Aspekte – sowohl des aufgabenbasierten Lehrens und Lernens als auch von Digitalität – fachspezifisch (Englischunterricht) ausgestaltet werden können.

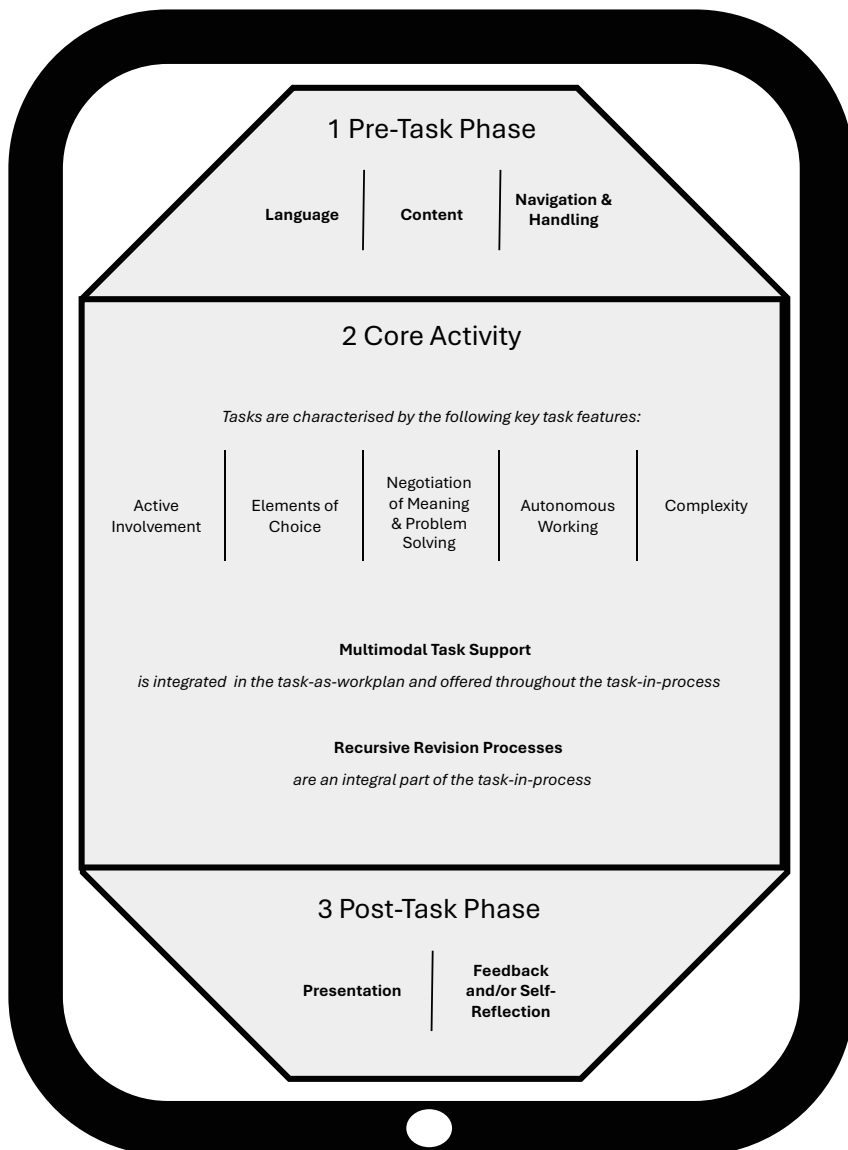


Abbildung 5: Konzept für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn

4.1 Pre-Task Phase

Hinsichtlich der *sprachlichen* Vorbereitung geht es darum, sprachliches Vorwissen der Kinder zu aktivieren, neue Sprache einzuführen oder gemeinsam mit den Kindern zu erarbeiten sowie die (neuen) Wörter, Phrasen, Sätze etc. im Rahmen von verschiedenen (spielerischen) Aktivitäten zu üben.

Die *inhaltliche* Vorbereitung zielt darauf ab, das Vorwissen der Kinder zu aktivieren, Wissen bereitzustellen oder gemeinsam mit den Kindern zu erarbeiten. Somit wird das Thema eingeführt und werden Motivation sowie Interesse für dieses geweckt. Zudem wird die Aufgabe gestellt, demonstriert und besprochen. Dabei werden auch das Ziel und Produkt/Ergebnis der Aufgabe geklärt oder ausgehandelt.

Analog zu den sprachlichen sowie inhaltlichen Kenntnissen und Fertigkeiten, die für eigenständiges, erfolgreiches Bearbeiten der Aufgabe(n) im *Task Cycle* entscheidend sind, erfordert ein digital gestütztes Setting Sprache und Handhabung des digitalen Materials. So kommt in der *Pre-Task Phase* nun ein drittes neues Element hinzu: Die *mediale* Vorbereitung (,Navigation and Handling of Digital Medium/Material'). Dies beinhaltet, die Fremdsprache, die für die Navigation und Handhabung des digitalen Mediums (z.B. Tablet) und digitalen Materials (z.B. App) nötig ist, vorzuentlasten und anhand des Mediums und Materials zu erarbeiten und zu verstehen (,User Language'). Dies kann beispielsweise die Klärung von Buttons, Settings, Funktionen etc. umfassen. Weiterhin werden im Bereich der medialen Vorbereitung auch Informationen, die für die Navigation und Handhabung des digitalen Mediums (z.B. Tablet) und digitalen Materials (z.B. App) zur Erstellung des digitalen Produkts relevant sind, bereitgestellt, demonstriert oder erarbeitet (,User Knowledge'). Im Rahmen von ,EMMa' wurden ,User Language' und ,User Knowledge' beispielsweise durch selbstgedrehte Tutorials zu den Apps, Handouts (siehe Abbildung 6), Demonstrationen, ,Ausprobierphasen' und Unterrichtsgespräche umgesetzt. Außerdem werden der Umgang und der Einsatz des digitalen Mediums und digitalen Materials organisiert (,Organising Technology'). Dies beinhaltet z.B. das Ausgeben und Einsammeln der iPads, Klärung von Rahmenbedingungen zur Nutzung des Internets, ,Aufräumen' des Home-Bildschirms etc.

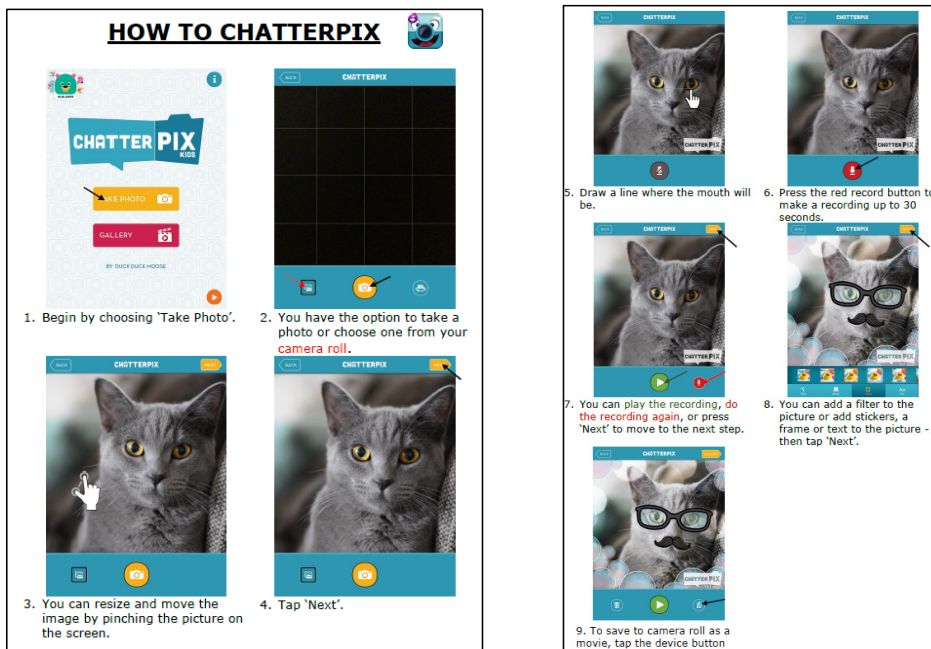


Abbildung 6: Handout zur App Chatterpix Kids (Duck Duck Moose LLC) (Quelle: selbst-erstellt)

4.2 Core Activity / Task Cycle

Die *aktive Involvierung* ist ein grundlegendes Element des aufgabenbasierten Arbeitens. Aufgrund des multimodalen Charakters digitaler Medien und Materialien findet dies in einem digital gestützten *Lernsetting*, aber auch durch das digitale Medium und Material statt: Dabei liegt der Fokus zum einen auf einem produktiven und nutzenden Umgang („Interactive Features“). Durch das Ausprobieren von verschiedenen *features* (beispielsweise Drücken des Buttons ‚Next‘, um weiterzukommen, Drücken des Mikrofons zur Audioaufnahme etc.) setzen sich die Kinder aktiv mit diesen auseinander. Es ist insbesondere die taktile Dimension, die einen produktiven und nutzenden Umgang fördert. Zum anderen halten die verschiedenen Modi (z. B. visuelle Stimuli, auditiver Input) die Aufmerksamkeit und die Motivation der Kinder aufrecht. Dabei ist nun ein konsumierender, rezeptiver Umgang gefordert („Multimodality“). In der Englisch-AG nutzten die Schüler:innen produktiv verschiedene *features* der App Chatterpix Kids (Duck Duck Moose LLC), um ihre Karnevalsfigur sprechen zu lassen:

(Paul zieht einen neuen Mund für die Figur.)

(Emil scrollt durch die Rahmen und Sticker.)

(Paul klickt auf weiter und verziert seine Figur mit Filtern und Stickern.)

(2022.03_Fasching2_Paul&Emil_Z.101-103)

Durch die Multimodalität eines digital gestützten Lernsettings erweitern sich auch die *Elements of Choice*: Um Sprache und Inhalte zu realisieren, können die Kinder im Bearbeitungsprozess verschiedene visuelle (z. B. Emoticons, Symbole, (selbstgemalte/fremde) Bilder, (eigene/fremde) Fotos etc.), auditive (z. B. (eigene) Audios, Geräusche etc.), audio-visuelle (z. B. (eigene/fremde) Videoaufnahme, Animationen etc.) und schriftliche (z. B. Sprechblasen, Sätze, Phrasen, Wörter, Überschriften etc.) Formen wählen. Der Screenshot des Schülerproduktes (siehe Abbildung 7) zeigt, dass die Kinder auf visuelle Elemente zurückgriffen, um sowohl Bedeutungen zu vermitteln als auch die Buchseiten zu ihrem Lieblingssport in der App Book Creator (Tools for Schools Limited) zu gestalten.



Abbildung 7: Screenshot der Titelseite von Bens multimodalem Book Creator Buch zu seinem Lieblingssport

Negotiation of Meaning und *Problem Solving* sind in einem digital gestützten Lernsetting stets kontextualisiert. Im Erarbeitungsprozess können sich die Kinder zum einen mit ihren Mitschüler:innen absprechen (z. B. hinsichtlich der Inhalte und wie diese zu versprachlichen sind, welche *features* der App hierzu genutzt werden sollen etc.). Zum anderen erhalten sie auch immer wieder Feedback durch die App (z. B. sprachlich, Buttons, die nicht reagieren, inhaltlich falsche Antworten, *trial-and-error*-Prozesse etc.). Sowohl Aushandlungs- als auch Problemlöseprozesse führen dazu, dass die Kinder ihren Erarbeitungsweg und das digitale Produkt fortlaufend überdenken und ggf. ändern. Das nachfolgende Transkript gibt einen Einblick in einen Programmierungsprozess mit der App Scratch-Junior (Scratch Foundation).

Christina: (Klickt auf das Bewegungsfeld nach unten). Gehe hier nach oben, einen Schritt. (Zieht die Bewegungsfelder in die Reihenfolge). Nein nicht einen, wir brauchen mehrere Schritte. Drei.

Petra: Ja, wir können auch drei machen.

Christina: Drei Schritte nach unten. (Lacht). Das sind zu viele. * So. Das sind also drei Schritte. Es soll da unten bleiben.

Petra: Hä, braucht man mehrere? Wir brauchen mehr. Das waren nur drei bis hier. Wir brauchen vielleicht fünf Schritte.

Christina: Klick. So fünf. Oder machen wir gleich sechs?

Petra: Ja, das geht ja. (Zieht das Tablet zu sich).

Christina: (Lässt die Figur sich nach oben bewegen, indem sie auf das Bewegungssymbol drückt). Ja, so passt es. Oder machen wir sieben?

Petra: Ja, machen wir sieben.

Christina: (Verändert die Anzahl der Schritte).

(2022.05_Kinderhotel3_Petra&Christina_Z.129-137)

Autonomous Working wird durch die multimodalen *features* erlaubt. Die Kinder können sich durch bildliche Darstellungen, Animationen, Audioangebote etc. Inhalte und Funktionen erschließen und somit eigenständig arbeiten. Wichtig dabei ist die Vertrautheit der Kinder mit den digitalen *features* des Mediums (Tablet) und des Materials (App). Die beiden untenstehenden Transkripte illustrieren, dass die Vertrautheit mit den digitalen *features* die Grundlage für eigenständiges Arbeiten ist:

Christina: Okay, dann machen wir lieber eine Sprechblase. (Klickt auf das Symbol für die Sprechblase). Wow, das sind anscheinend Sprechblasen.

(2022.05_Kinderhotel3_Petra&Christina_Z.45)

(Paul wählt einen violetten Block aus.)

Paul: Damit wird das kleiner.

(Paul wählt den gelben Block mit der grünen Flagge aus.)

(Paul wählt den roten Endblock aus.)

(Paul startet die Programmierung mehrmals. Das Kaninchen wird immer kleiner.)

Paul: Perfekto.

(2022.05_Kinderhotel2_Paul&Emil_Z.44-50)

In Bezug auf den Aspekt der *Complexity* kann in digital gestützten Lernsettings die Multimodalität sowohl förderlich als auch herausfordernd sein. Die Art der digitalen *features* (z. B. mündlich, schriftlich, rezeptiv, produktiv etc.) wirkt sich reduzierend oder erhöhend auf die Komplexität und Anforderungen aus („Modes of Language“). Der Typ und die Anzahl der digitalen *features* (z. B. intuitive Nutzung, Ablenkungen, viele/wenige etc.) minimiert oder maximiert die Komplexität und Anforderungen („Type/Amount of Features“). Bei Unklarheiten wandten sich manche Kinder direkt an die Lehrperson:

Paul (an die Lehrkraft gerichtet): Wie ändere ich denn da die Figur? Back?

L: Click on back. (Paul klickt auf ‚back‘) You may click on back. (Paul klickt auf ‚back‘). And then you can delete it. And then / You can now click on this one and you can choose a different figure.

Paul: Ah.

(2022.03_Fasching2_Paul&Emil_Z.68-70)

Andere Schüler:innen versuchten, Herausforderungen eigenständig zu lösen:

Petra: Ja wie löscht man den?

Christina: Ja, löschen. (...)

(Petra tippt und zieht die Katze. Sie wischt diese über den Hintergrund.)

Christina: Wir machen die halt mal kleiner. * Ganz klein machen.

(Christina versucht die Katze durch das Zusammenschieben der Finger zu verkleinern.)

(Christina tippt sich durch die verschiedenen Bereiche der Blöcke.)

Petra: Vielleicht ja hier. Hier gibt es (unv.) zum Löschen.

(Christina tippt mehrfach auf die linke Figurenleiste und geht dann auf den Editor.)

Christina: Nein, die soll ja weg.

(Christina zoomt auf das Gesicht der Katze. Christina macht die Katze weiß und lässt sie somit im weißen Hintergrund verschwinden.)

Christina: Ja. Ich hab sie weg.

(2022.05_Kinderhotel2_Petra&Christina_Z.212-226)

In einem digital gestützten Lernsetting wird während des *Task Cycles Task Support* sowohl für die Navigation und die Handhabung des digitalen Mediums und der digitalen Materialien (z. B. in Form eines Handouts, siehe Abbildung 6) als auch für die Erstellung des eigenen digitalen Produkts angeboten bzw. genutzt (z. B. in Form von multimodalen *word webs*, digitalen Wörterbüchern etc.). Die Schüler:innen griffen beispielsweise auf die *word webs*, zurück, um ihre Produkte zu versprachlichen:

(Emil öffnet zum ersten Mal seine Mappe und schlägt das MindMap auf.)
(2022.03_Fasching2_Paul&Emil_Z.105)

Zudem – in diesem Punkt unterscheidet sich die Konzeptualisierung für digitale Lernsettings vom ursprünglichen TBA-Modell – finden fortlaufend Überarbeitungsprozesse und Verbesserungen bereits während des *Task Cycles* (und nicht erst in der *Post-Task Phase*) statt. Die Kinder verbesserten, veränderten und erweiterten ihr digitales Produkt (sprachlich) von sich aus („Self-Initiated“), aufgrund von Hinweisen eines anderen Kindes („Peer(s)“), Hinweise der Lehrperson („Teacher“) oder aufgrund von Hinweisen der App („App-Feedback“). Beispielsweise verbesserten Schüler:innen ihre Audioaufnahmen bei der Erstellung ihrer Karnevalsfigur mit der App Chatterpix (Duck Duck Moose LLC):

(Drückt auf den Aufnahmebutton und spricht ein) It is carnival and today, there is a fancy dress party at school. Guess what I am. I am wearing a blue eyes and green skin colour. I look spooky and ugly. Guess what I am. I am an alien.

(Drückt den Button zum Stoppen der Aufnahme.) (...)

(Drückt auf den Play Button und hört sich die Aufnahme seiner Chatterpix Figur an.)

(Drückt auf den Zurück Button.)

(Verändert den Mund seiner Figur noch einmal und versucht, eine neue Linie zu ziehen.)

(Drückt auf den Zurück Button und zieht eine neue Mundlinie.)

(Drückt auf den Play Button und hört sich den Anfang seiner Aufnahme an.)

(...)

(Drückt auf den Button zum Aufnehmen und spricht ein:) It is carnival and today, here is a fancy dress party at school. Guess what I am. I am wearing a blue eyes and green skin colour. I look spooky and ugly. Guess what I am. I am an alien. (Drückt den Button zum Stoppen der Aufnahme.)

(Drückt auf den Weiter Button und gestaltet sein Alien mit Accessoires, indem er die Symbole mit den Fingern antippt und verschiebt. Er drückt auf das Symbol der Mülltonne, um Elemente wieder zu löschen.)

(Sucht sich eine Schriftart aus und gibt das Wort „Alien“ ein.) (...)

(Drückt auf den Play Button und hört sich die Aufnahme seiner Chatterpix Figur an.)

(2022.03_Fasching2_Fynn_Z.92-110)

4.3 Post-Task-Phase

Auch in einem digital gestützten Lernsetting wird in der *Post-Task Phase* (in Ergänzung zur bereits im *Task Cycle* erfolgten Rückmeldung) *Feedback* zu sprachlichen und inhaltlichen Aspekten gegeben oder werden die Kinder explizit zur *Self-Reflection* angeregt. Darüber hinaus wird aber auch auf die digitale Qualität des Produktes/Ergebnisses eingegangen (z. B. Qualität der Tonaufnahme, Umsetzung der Aufgabe mit/durch digitale *features* etc.). Im Forschungsprojekt ‚EMMa‘ wurden die Kinder nach dem Abschluss einer jeden thematischen Einheit unter anderem gebeten, auch den Mehrwert der App für ihren fremdsprachlichen Lernprozess zu reflektieren:

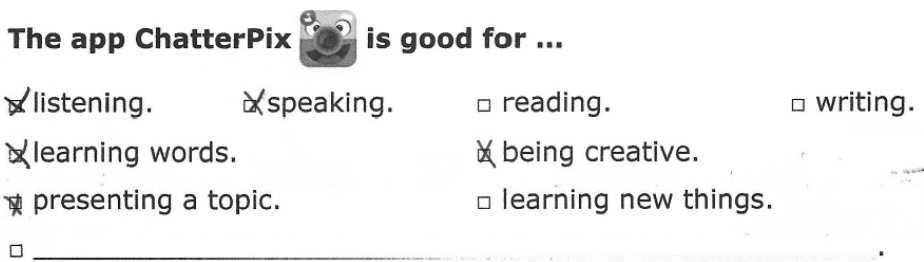


Abbildung 8: Ausschnitt aus Julias Reflexion zum Potenzial der App ChatterPix Kids (Duck Duck Moose LLC) für den eigenen fremdsprachlichen Lernprozess (Reflexionsblatt selbst erstellt)

Für die *Presentation* des digitalen Produkts/Ergebnisses gilt es, in einem digital gestützten Lernsetting einen kommunikativen Rahmen zu geben. So regen Höraufträge (z. B. Fragestellung) und Höranlässe (z. B. Quizze) die Kinder an, der jeweiligen Präsentation aufmerksam zu folgen (‚Active Listening/Watching‘). Zudem bieten gezielt geplante Sprechkanäle (z. B. Fragen, Rätsel) Raum für einen angeleiteten Austausch über das vorgestellte digitale Produkt/Ergebnis (‚Interaction‘). In einem interaktiven und kommunikativen Setting wurden die multimodalen Book Creator Bücher präsentiert und gemeinsam diskutiert.

L: And after sports, obviously, he seems to very, very hungry. What does he like to eat? Fynn?

Fynn: What?

L: What does he like to eat?

Fynn: Er eat a sandwich and drink water.

(2022.02_Sport2_Z.288-291)

5 Zusammenspiel generischer und fachspezifischer Aspekte

Im ersten Aktionszyklus des Forschungsprojekts ‚EMMa‘ wurde deutlich, dass mit, wegen und/oder durch digitale Medien und Materialien vielfältige Lerngelegenheiten entstehen können, die es bereits jungen Lernenden erlauben, Englischkompetenzen (= fachspezifische Kompetenzen) und Medienkompetenzen (= generische Kompetenzen) gleichermaßen zu entwickeln. Dabei wurden Oberflächenstrukturen thematisiert (siehe Abschnitt 1), deren etwaiges Zusammenspiel mit Tiefenstrukturen dann durch die Videodaten (z. B. Aufzeichnung von Arbeitssequenzen) und Schüler:innenperspektiven (z. B. Reflexionen) diskutiert wurden (siehe Abschnitt 3). Durch die Betrachtung dieser Wechselwirkungen konnte die Forschungsfrage ‚Welche Aufgaben und Aktivitäten eignen sich, um das Lernpotenzial digitaler Medien und Materialien auszuschöpfen?‘ mit der Entwicklung des ‚Konzepts für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn‘ beantwortet werden (siehe Abbildung 5).

5.1 Untersuchung generischer und fachspezifischer Aspekte zur Verbesserung der Unterrichtsqualität im Grundschulenglischunterricht

Zu dem Zusammenspiel fachspezifischer und generischer Aspekte besteht weiterer Forschungsbedarf, dem sich das Aktionsforschungsprojekt ‚EMMa‘ in zusätzlichen Aktionszyklen widmen möchte. So wurde zwischen Januar 2023 (bis Juni 2023) ein zweiter Aktionszyklus durchgeführt und für das zweite Schulhalbjahr 2024 ist ein dritter Aktionszyklus angesetzt, in dem insbesondere die Aufgabenentwicklung und die Unterrichtsgestaltung in einem digital gestützten, aufgabenbasierten Grundschulenglischunterricht untersucht wurden bzw. werden. Die Aktionsforschung dient als ein „Instrument zur Erprobung und ggf. Weiterentwicklung wissenschaftlicher Erkenntnisse in der Praxis“ (Caspari, 2016, S. 72). Dabei ist ‚EMMa‘ bewusst als längerfristiger, zyklischer Prozess geplant, um theoretische Annahmen zur Veränderung des Englischunterrichts in der Unterrichtspraxis zu überprüfen und für diese nach erneuter Reflexion sowohl als revidierte Praxisvorschläge als auch Veränderungen der theoretischen Annahmen zu formulieren (Burns, 2010, S. 9). Das Aktionsforschungsprojekt ‚EMMa‘ bedient sich eines qualitativen Zugangs zur Analyse von Unterrichtsqualität im Fach Englisch. Wie dieser Beitrag zeigt, ist das Potenzial dieses Ansatzes insbesondere die Konzentration auf einen spezifischen Kontext (= Gestaltung eines digital gestützten aufgabenbasierten Englischunterrichts), anhand dessen auch andere Akteur:innen (z. B. Forscher:innen, Lehrer:innen) Schlussfolgerungen (z. B. hinsichtlich Medienauswahl, Aufgabenentwicklung, Forschungsdesign) für ihre jeweiligen Kontexte und Ziele ableiten können. Zudem werden im Aktionsforschungsprojekt ‚EMMa‘ umfangreiche Unterrichtsaufnahmen angefertigt, die „als Katalysator einer (fach)didaktisch konkretisierten Unterrichtsforschung“ (Reusser & Pauli, 2021, S. 193) an-

gesehen werden. Das Aktionsforschungsprojekt ‚EMMa‘ bietet gerade aufgrund seines triangulierenden Studiendesigns neue Denkansätze und Impulse, welche Bedeutung Aufgaben für die Gestaltung eines qualitativvollen digital gestützten Unterrichts in den Fremdsprachen haben können. Durch die Arbeit mit und in der Schulpraxis möchte ‚EMMa‘ sowohl auf theoretischer und konzeptueller wie auch unterrichtspraktischer Ebene wertvolle Hinweise für die zukünftige Gestaltung (z. B. Aufgabendesign, fachspezifischer und kompetenzförderlicher Medieneinsatz) von digital gestützten Lernsettings im Grundschulenglischunterricht geben und damit konkrete Angebote zur Weiterentwicklung von Unterrichtspraktiken schaffen. Die Ergebnisse (siehe Abschnitte 3 und 4.2) zeugen von einer Verwobenheit generischer und spezifischer Aspekte: Somit bekräftigt das Kapitel den weiteren Bedarf an fachspezifischer Forschung und darüber hinaus den Wunsch eines regen Austauschs und engen Zusammenwirkens von Bildungswissenschaften und Fachdidaktiken (siehe Lindl et al., in diesem Band).

5.2 Generische und fachspezifische Aspekte als Beiträge zur qualitativvollen Unterrichtsgestaltung

Durch die Anbahnung fremdsprachlicher Kompetenzen mit digitalen Medien und Materialien bringt sich das Fach Englisch in den derzeitigen bildungswissenschaftlichen Diskurs ein, indem es das fächer-, schulstufen- und schulartenübergreifende Ziel berücksichtigt, „den Lernenden eine mündige, souveräne und aktive Teilhabe an der digitalisierten Lebens- und Arbeitswelt [zu] ermöglichen“ (KMK, 2021, S. 6). Gleichwohl zeigt das Forschungsprojekt ‚EMMa‘ auf, wie wichtig Fachspezifik und -kultur sind, da im Schulfach Englisch letztlich immer das qualitativvolle Lehren und Lernen der englischen Sprache im Mittelpunkt stehen sollte. Im Forschungsprozess hat sich dabei vor allem gezeigt, dass Aufgaben auch in einem digital gestützten Lernsetting von essenzieller Bedeutung für fremdsprachliche Lehr- und Lernprozesse sowie die fremdsprachliche Kompetenzentwicklung der Schüler:innen sind.

Zudem konnte ein ‚Konzept für aufgabenbasiertes Lehren und Lernen im digital gestützten Fremdsprachenfrühbeginn‘ (Abb. 5) erstellt werden. Das Modell identifiziert Merkmale eines qualitativvollen digital gestützten Englischunterrichts, der sowohl fachspezifischen Kompetenzen (d. h. Englischkompetenzen) als auch fächerübergreifenden Kompetenzen (d. h. Medienkompetenzen) gerecht wird. Weiterhin bestärkt das Forschungsprojekt ‚EMMa‘, wie wichtig das dem MAIN-TEACH-Modell (Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) zugrunde liegende Verständnis des Wechselspiels zwischen unterrichtlichem Angebot auf Lehrkraftseite und dessen Nutzung auf Schüler:innenseite sowie reflexiv-zirkulären und kokonstruktiven Aushandlungsprozessen ist: Aufgaben nehmen darin eine zentrale Rolle ein und bilden den fremdsprachlichen, inhaltlichen und digital gestützten Rahmen, in dem Schüler:innen und Lehrer:innen kompetenz- und ergebnisorientiert arbeiten.

Literatur

- Alhinty, M. (2015). Young Language Learners' Collaborative Learning and Social Interaction as a Motivational Aspect of the iPad. *International Journal of Emerging Technologies in Learning*, 10(2), 24–29. <https://doi.org/10.3991/ijet.v10i2.4313>
- Altrichter, H., Posch, P., Spann, H. (2018). *Lehrerinnen und Lehrer erforschen ihren Unterricht: Unterrichtsentwicklung und Unterrichtsevaluation durch Aktionsforschung*. Bad Heilbrunn: Klinkhardt. <https://doi.org/10.36198/9783838547541>
- Aufenanger, S. (2017). Zum Stand der Forschung zum Tableteinsatz in Schule und Unterricht aus nationaler und internationaler Sicht. In J. Bastian & S. Aufenanger (Hrsg.), *Tablets in Schule und Unterricht. Forschungsmethoden und -perspektiven zum Einsatz digitaler Medien* (S. 119–138). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-13809-7_6
- BIG-Kreis (Hrsg.) (2015). *Der Lernstand im Englischunterricht am Ende von Klasse 4. Ergebnisse der BIG-Studie*. München: Domino.
- Brunsmeyer, S. (2023). Kommunikative Kompetenzen im digital-gestützten Grundschulenglischunterricht entwickeln. In H. Boettger & N. Schlueter, Norbert (Hrsg.), *Fortschritte im frühen Fremdsprachenlernen. Konferenzband zur 5. Tagung* (S. 125–133). Salzburg: Schriftbild Verlag.
- Brunsmeyer, S. (2024). Digital Silence? Fostering Oracy in the (Digital) EFL Classroom. [Unveröffentlichtes Manuskript]. Passau.
- Brunsmeyer, S. & Holberg, S. (2022). Making the match: English and media competences. Digitale Medien im Englischunterricht der Grundschule. *Grundschule Englisch*, 79, 2–7.
- Burns, A. (2010). *Doing Action Research in English Language Teaching. A Guide for Practitioners*. New York: Routledge. <https://doi.org/10.4324/9780203863466>
- Butler, Y. G. (2019). Gaming and young learners. In S. Garton, Sue & F. Copland (Hrsg.), *The Routledge Handbook of Teaching English to Young Learners* (S. 305–319). Oxon: Routledge. <https://doi.org/10.4324/9781315623672-20>
- Butler, Y. G. (2022). Language Education in the Era of Digital Technology. *JALT Journal*, 44(1), 137–152. <https://doi.org/10.37546/JALTJJ44.1-7>
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511733109>
- Caspari, D. (2016). Prototypische Designs. In D. Caspari, F. Klippel, M. K. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik. Ein Handbuch* (S. 67–78). Tübingen: Narr.
- Decristan, J., Hess, M., Holzberger, D. & Praetorius, A. K. (2020). Oberflächen- und Tiefenmerkmale: Eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. *Zeitschrift für Pädagogik*, 66. Beiheft, 102–116. <https://doi.org/10.3262/ZPB2001102>
- DESI-Konsortium (Hrsg.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Drechsel, B. & Schindler, A. K. (2019). Unterrichtsqualität. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 353–372). Berlin: Springer. https://doi.org/10.1007/978-3-662-55754-9_18
- Dreßler, C. (2018). *Nature and enactment of tasks for early English as a foreign language teaching: A collaborative research project with teachers*. Tübingen: Narr.
- Elsner, D. & Viebrock, B. (Hrsg.) (2014). *Triangulation in der Fremdsprachenforschung*. Frankfurt am Main: Peter Lang. <https://doi.org/10.3726/978-3-653-04899-5>

- European Commission (2012). *European Strategy for a Better Internet for Children*. Brussels.
- Flick, U. (2019). Gütekriterien qualitativer Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 473–488). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-21308-4_33
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D. & Friedman, T. (Hrsg.) (2018). *IEA International Computer and Information Literacy Study 2018. Assessment Framework*. Cham: Springer. <https://doi.org/10.1007/978-3-030-19389-8>
- Frisch, S. (2017). Eckpfeiler des kommunikativen Englischunterrichts in der Grundschule. *Fremdsprachen lehren und lernen*, 46(2), 55–69.
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L. & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Grünewald, A. (2012). Förderung interkultureller Kompetenz durch Lernaufgaben. *Fremdsprachen lehren und lernen*, 41(1), 54–71.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie* (S. 345–363). Weinheim: Beltz.
- Kersten, K., Bruhn, A.-C., Ponto, K., Böhnke, J. & Greve, W. (2018). Teacher Input Observation Scheme (TIOS). *Studies on Multilingualism in Language Education*, 4. Universität Hildesheim.
- Kolb, A. & Schocker, M. (2021). *Teaching English in the primary school: A task-based introduction for pre- and in-service teachers*. Hannover: Klett/Kallmeyer.
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N. & Blum, W. (2020). Competence as a continuum in the COACTIV study: The “cascade model”. *ZDM – Mathematics Education*, 52, 311–327. <https://doi.org/10.1007/s11858-020-01151-z>
- Kultusministerkonferenz (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz*. Berlin: Eigendruck.
- Kultusministerkonferenz (2021). *Lehren und Lernen in der digitalen Welt. Ergänzung zur Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“*. Berlin: Eigendruck.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Schöningh. <https://doi.org/10.36198/9783838538952>
- Laranjeiro, D. (2021). Development of Game-Based M-Learning Apps for Preschoolers. *Education Sciences*, 11(5). <https://doi.org/10.3390/educsci11050229>
- Mayring, P. & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 633–647). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-21308-4_42
- Meyer, M., Volkmann, L. & Grimm, N. (2022). *Teaching English*. Tübingen: Narr. <https://doi.org/10.24053/9783823393931>
- mpfs = Medienpädagogischer Forschungsverbund Südwest (2021). *KIM-Studie 2020. Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger*. Stuttgart. (https://www.mpfs.de/fileadmin/files/Studien/KIM/2020/KIM-Studie2020_WEB_final.pdf)
- Praetorius, A. K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>

- Pinter, A. (2015). Task-Based Learning with Children. In J. Bland (Hg.), *Teaching English to young learners: Critical issues in language teaching with 3–12 year olds* (S. 113–127). London/New York: Bloomsbury.
- Pinter, A. (2019). Agency and technology-mediated task repetition with young learners. Research and implications for primary classroom practice. *Language Teaching for Young Learners*, 1(2), 139–160. <https://doi.org/10.1075/ltyl.00010.pin>
- Reusser, K. & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49, 189–202. <https://doi.org/10.1007/s42010-021-00117-8>
- Schocker-von Ditfurth, M. & Müller-Hartmann, A. (2010). *Teaching English: Task-Supported language learning*. Paderborn: UTB. <https://doi.org/10.36198/9783838533360>
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- vbw = Vereinigung der Bayerischen Wirtschaft e.V. (Hrsg.) (2021). *Digitale Bildung an bayerischen Schulen vor und während der Corona-Pandemie*. München. https://www.vbw-bayern.de/Redaktion/Frei-zugaengliche-Medien/Abteilungen-GS/Bildung/2021/Downloads/Bi-174-001_Digitale-Bildung-an-bayerischen-Schulen-vor-und-waerend-der-Corona-Pandemie_17.03.2021.pdf
- Vodafone Stiftung GmbH (Hrsg.) (2020). *Schule auf Distanz. Perspektiven und Empfehlungen für den neuen Schulalltag. Eine repräsentative Befragung von Lehrkräften in Deutschland*. Düsseldorf. https://www.vodafone-stiftung.de/wp-content/uploads/2020/05/Vodafone-Stiftung-Deutschland_Studie_Schule_auf_Distanz.pdf
- Welling, S. (2017). Methods matter. Methodisch-methodologische Perspektiven für die Forschung zum Lernen und Lehren mit Tablets. In J. Bastian & S. Aufenanger (Hrsg.), *Tablets in Schule und Unterricht. Forschungsmethoden und -perspektiven zum Einsatz digitaler Medien* (S. 15–36). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-13809-7_2
- Whyte, S. & Cutrim Schmid, E. (2019). Classroom technology for young learners. In S. Garton & F. Copland (Hrsg.), *The Routledge Handbook of Teaching English to Young Learners* (S. 338–355). Oxon: Routledge. <https://doi.org/10.4324/9781315623672-22>
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49, 211–219. <https://doi.org/10.1007/s42010-021-00105-y>
- Willis, J. (2005). *A framework for task-based learning*. Harlow: Longman.
- Ziegler, N. (2016). Taking Technology to Task: Technology-Mediated TBLT, Performance, and Production. *Annual Review of Applied Linguistics*, 36, 136–163. <https://doi.org/10.1017/S0267190516000039>

Im Kapitel thematisierte Apps

- Tools for Schools Limited. *Book Creator* [Mobile app]. App Store. <https://bookcreator.com/>
- Duck Duck Moose LLC. *ChatterPix Kids* (Version 1.3.7) [Mobile app]. App Store. <https://apps.apple.com/de/app/chatterpix-kids/id734046126>
- Scratch Foundation. *Scratch Junior* [Mobile app]. App Store. <https://www.scratchjr.org/>

Dennis Tark

Der schulische Fremdsprachenunterricht Russisch: Unterrichtsqualität und Fachlichkeit

Abstract (der Herausgeber:innen)

In his article “Teaching Russian as a foreign language at school: Instructional quality and subject-specificity,” Dennis Tark focuses on the distinctive role of Russian as both a foreign and heritage language and, above all, the considerable heterogeneity of learners in this school subject. Due to the specific challenges and the dearth of research in Russian didactics in this area, the question of quality criteria of Russian lessons remains largely unanswered. In the light of the above, the article will proceed to examine the aforementioned considerations of effective Russian teaching. The synthesis framework, or MAIN-TEACH model, thus serves as a starting point, reference point and counterpoint for this development.

1 Russischunterricht in Deutschland

1.1 Institutionelle Rahmenbedingungen

Das Unterrichtsfach Russisch blickt in Deutschland auf eine lange Tradition zurück und gehörte, auch wenn die Anzahl der Lernenden seit einigen Jahren rückläufig ist, bis vor Kurzem zu den fünf Fremdsprachen, die am häufigsten an deutschen Schulen gelehrt werden. Im Schuljahr 2020/2021 belegten das Fach noch ca. 94.000 Schüler:innen (Destatis, 2022; vgl. auch Reimann, in diesem Band).¹ Dieses wird in der Regel als zweite oder dritte Fremdsprache ab der siebten Jahrgangsstufe oder höher im Sekundarbereich I und II gewählt; in Thüringen kann Russisch sogar schon in der Grundschule ab der dritten Klasse 3 und in bestimmten Einrichtungen beispielsweise an Waldorfschulen als erste Fremdsprache bzw. an Europaschulen als Bildungssprache ab der ersten Jahrgangsstufe gelernt werden.² Zwar bestehen derartige Möglichkeiten für Russisch als erste Fremdsprache an anderen Schulen nicht (Bergmann, 2016), allerdings bieten einige Bundes-

1 Der deutliche Rückgang der Lernendenzahlen im Vergleich zu früheren Jahren ist wohl auf die politischen Umstände, den besonderen Status der Sprache an Schulen in der DDR sowie auf deren niedriges gesellschaftliches Prestige, aber auch auf die geringen Lernerfolge zurückzuführen, die eine Vermeidung und Abwahl der Sprache begünstigen (z. B. Bergmann, 2017, S. 131f.; Bergmann & Böhmer, 2020, S. 71).

2 Der Unterricht an Europaschulen wird durchgehend bilingual gehalten: <https://www.berlin.de/sen/bildung/schule/besondere-schulangebote/staatliche-europaschule/>

länder Russisch als Herkunftssprache schon ab der ersten Klasse in der Primarstufe bzw. ab dem Übergang zur Sekundarstufe I an, sofern eine entsprechende Nachfrage (meist ab ca. zehn Schüler:innen) besteht (z.B. Bremen, Hamburg, Nordrhein-Westfalen, Niedersachsen, Rheinland-Pfalz, Sachsen, Saarland). Dieser herkunftssprachliche Unterricht wird im Sekundarbereich II jedoch in keinem Bundesland fortgeführt, sodass etwaige Schüler:innen spätestens dann in den regulären Fremdsprachenunterricht integriert werden, was nicht unproblematisch ist, da dies die sprachbezogene Leistungsheterogenität der Schülerschaft vergrößert und den Unterricht vor besondere Herausforderungen stellt (siehe hierzu auch Abschnitte 1.3 und 1.4). Einen zusammenfassenden Überblick über die vielfältigen Optionen, Russischunterricht an deutschen Schulen zu besuchen, bietet Abbildung 1.

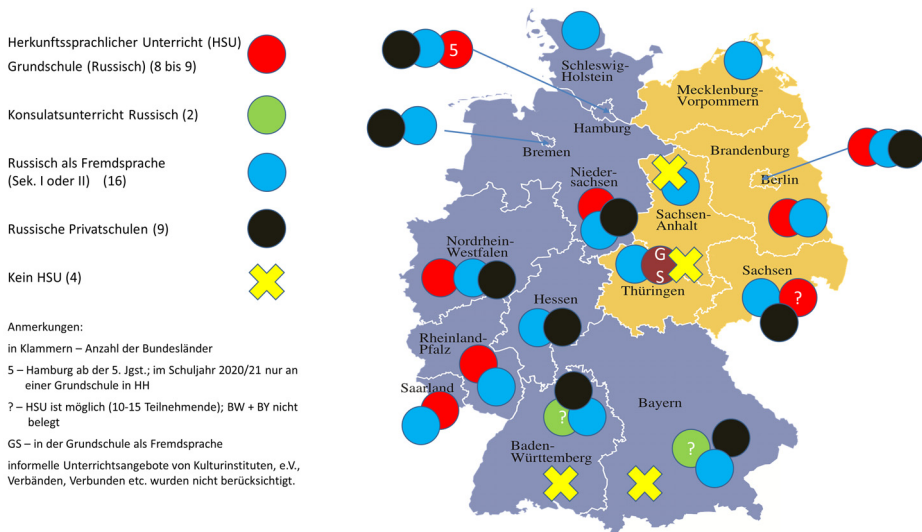


Abbildung 1: Russischunterricht im schulischen Kontext (Stand 2021)

In denjenigen Bundesländern, in denen die Wahl des Fachs Russisch als Herkunfts- oder Fremdsprache möglich ist, liegen hierfür eigene Lehr-, Rahmen- bzw. Bildungspläne vor. Bundesweit einheitliche Bildungsstandards existieren nicht, sodass die Voraussetzungen sowie Anforderungen, die an Schüler:innen gestellt werden, variieren und beispielsweise nach bestimmten Jahrgangsstufen erreichte Kompetenzniveaus nur eingeschränkt vergleichbar sind. Orientierungshilfen bieten hier lediglich die Bildungsstandards für die modernen Fremdsprachen Englisch und Französisch (Bergmann, 2014; Mengel & Eckert, 2006), welche allerdings die Besonderheiten des Russischunterrichts, unter anderem bezüglich der Gewichtung z. B. grammatischer Kenntnisse bereits auf den untersten Niveaustufen, nicht beachten. Zudem erschwert dieses Fehlen eines (länder-)gemeinsamen Referenzrahmens auch eine Konzeption übergreifender Qualitätsmerkmale und Zielvorstellungen von Russischunterricht (vgl. Abschnitt 2.3), zumal dieser

je nach Fokus auf Herkunfts- oder Fremdsprache ganz unterschiedliche Schwerpunkte setzen kann und muss.

1.2 Herkunfts- versus Fremdsprache

Die Frage nach der Herkunftssprache von Lernenden ist nach Mehlhorn (2022) im Anschluss an die englischsprachige Literatur nicht zwingend an dem in dieser Sprache erreichten Kompetenzniveau, sondern an der jeweiligen biografischen und familialen Situation festzumachen. Dementsprechend definiert er als die

„Herkunftssprache die Erstsprache einer Person [...], die in einer Familie aufwächst, in der nicht (ausschließlich) die Sprache der umgebenden Mehrheitsgesellschaft verwendet wird. [...] Eine Person, die als Herkunftssprecher:in bezeichnet wird, wurde entweder schon im Aufnahmeland geboren oder ist mit den Eltern aus dem Herkunftsland noch vor dem Schuleintritt eingewandert und wächst somit seit der Kindheit in einer mehrsprachigen Lebenswelt auf. [...] In der Regel wird während der Schulzeit die Sprache der Mehrheitsgesellschaft zur dominanten Sprache“ (S. 1).

Damit erfolgt die Zuschreibung von Herkunftssprache eher über die Reihenfolge des Spracherwerbs als das darin erreichte Kompetenzniveau, sodass ein etwaiger Unterricht sogar dazu beitragen kann, eine zwischenzeitlich eventuell vernachlässigte Herkunftssprache nach längerem Nichtgebrauch (erst wieder) zu erlernen (Mehlhorn, 2022, S. 1f.). Ähnlich beschreibt dies auch Bergmann (2021), indem

„[...] das Konzept Herkunftssprache vage und nicht eindeutig abgrenzbar gegenüber dem Konzept Muttersprache [ist]. Es sagt auch nur in einem groben Rahmen etwas über die sprachlichen Kompetenzen von Herkunftssprechern aus, die sehr breit variieren können. Auf der individuellen Ebene ist die Herkunftssprache nicht zwangsläufig ein dominierender Identifikationsmarker [...]“ (S. 5).

Dies hat auch Auswirkungen auf den Unterricht von Russisch als Herkunftssprache und wurde bereits in mehreren Forschungsarbeiten als ein wesentlicher Unterschied zum Fremdsprachenunterricht herausgearbeitet (z.B. Anstatt, 2008, 2011; Bergmann, 2014, 2016, 2017; Lüttenberg, 2010; Olfert, 2019). Denn es gibt keine speziellen Voraussetzungen für den Besuch des herkunftssprachlichen Unterrichts (z.B. Einstufungsverfahren, Eingangsdiagnostik) oder das Erfordernis bestimmter Vorkenntnisse. So können der Gruppe der Herkunftssprecher:innen Personen mit teilweise sehr heterogenen Sprachniveaus zugeordnet werden (Bergmann, 2021, S. 9), die aus unterschiedlichen Lernbiografien bzw. Erwerbssituationen resultieren. Auf der einen Seite befinden sich hierrunter Schüler:innen, deren sprachliche Kompetenzen rudimentär und kaum von denjenigen, die Russisch ohne jegliche Vorkenntnisse als Fremdsprache lernen, zu unterscheiden sind (Olfert, 2019, S. 84). Sie werden nicht selten auch gerade wegen unterrichtlicher Rahmenbedingungen und ggf. schulorganisatorischer Notwendigkeiten

(z. B. kleine Lernendengruppen, Mangel an qualifiziertem Lehrpersonal) in Kurse von Herkunftssprecher:innen aufgenommen und dadurch gewissermaßen ‚ethnisiert‘, was nicht unproblematisch erscheint (Bergmann, 2021, S. 10). Andererseits gibt es Lernende, deren Sprachkompetenzen in der Herkunftssprache Russisch kaum von denjenigen in der Muttersprache bzw. L1 Russisch abweichen, wenn Deutsch auch weiterhin als primäre Sprache identifizierbar ist (Bergmann, 2021, S. 5). Mit dieser potenziell stark ausgeprägten sprachlichen Heterogenität der Gruppe von Herkunftssprecher:innen (z. B. Bergmann, 2021, S. 8), die bei der Unterrichtsgestaltung und individuellen Förderung der Schüler:innen berücksichtigt werden sollte, beschäftigt sich die bisherige didaktische Forschung vornehmlich und hebt die Bedeutung des Spracherhalts und der -attrition hervor (z. B. Anstatt, 2008, 2011; Lüttenberg, 2010; Olfert, 2019; siehe dazu auch Abschnitt 1.4). Damit gehen auch Fragen nach adäquaten Differenzierungsmöglichkeiten von Lehr- und Lernmaterialien sowie nach didaktischen Möglichkeiten einher, bei den Schülerinnen und Schülern die Motivation zur Verstetigung und Weiterentwicklung herkunftssprachlicher Kompetenzen aufrechtzuerhalten. Den hiervon abweichenden Herausforderungen und sprachlichen Besonderheiten, denen Lernende mit Deutsch als Erstsprache gegenüberstehen, wird hingegen weniger Aufmerksamkeit zuteil. Konkrete didaktische Ansätze und Innovationen zu Themen, Materialien und Methoden für diese ebenfalls sehr leistungsdifferente Lernendengruppe scheinen hier dem pädagogischen Geschick, dem Ermessen und der Verantwortung der individuellen Lehrkraft überlassen (Bergmann, 2016, S. 60).

1.3 Leistungsheterogenität der Schülerschaft

Als ein zentrales Charakteristikum schulischen Russischunterrichts sowohl als Herkunfts- als auch als Fremdsprache lässt sich also das breite Spektrum an sprachlicher Leistungsheterogenität zwischen Lernenden mit Migrationshintergrund und nahezu erstsprachlichen Kompetenzen sowie Anfängerinnen und Anfängern ohne jegliche Vorkenntnisse identifizieren (Anstatt, 2008, 2011; Bergmann, 2016). Diese gemischte Gruppe der Rezipientinnen und Rezipienten setzt sich – wie in Tabelle 1 dargestellt – im Wesentlichen aus vier Teilgruppen zusammen, ohne dass diese in sich homogen sind. Eine solche formale Klassifikation dient zu einem besseren Überblick über die vielfältigen Gruppen von Lernenden und zu deren analytischer Beschreibung; für eine Zuordnung von Einzelfällen wäre eine differenzierte Diagnostik erforderlich.

Die erste Gruppe in der äußeren linken Spalte von Tabelle 1 bilden die deutschsprachigen Schüler:innen, die Russisch in der Regel als zweite oder dritte Fremdsprache wählen. Zur zweiten Gruppe in der zweiten Spalte von Tabelle 1 gehören diejenigen, die Russisch als Herkunfts- oder Zweitsprache in ihrem familialen Umfeld oder Freundeskreis sprechen. Sofern sie die Sprache nur verstehen und nicht selbst sprechen können, gehören sie, wenn sie in Deutschland geboren sind, ebenfalls der ersten Gruppe an. Je nach Geburtsort, Kontaktzeitpunkt bzw.

-dauer kann es dabei auch vorkommen, dass keine der beiden Sprachen auf herkunftssprachlichem Niveau gesprochen wird.

Tabelle 1: Klassifizierung der Schülerschaft im Fach Russisch – Schüler:innen mit Deutsch als Muttersprache und mit Russisch als L1, L2, Lx

Schüler:innen mit Deutsch als Muttersprache	Herkunftssprecher:innen in Deutschland geboren	Herkunftssprecher:innen mit Migrationshintergrund	Andere Herkunftssprachen
– Deutsche Muttersprachler:innen ohne Vorkenntnisse	<ul style="list-style-type: none"> – L1: Deutsch (wird besser beherrscht) – L2: Russisch (unabhängig davon, ob beide Eltern Russisch als L1 sprechen oder nur ein Elternteil): <ul style="list-style-type: none"> • wird bestenfalls auf Alltagssprachlichem Niveau gesprochen bzw. sind nur mündliche Kompetenzen (gut) entwickelt • wird schlechtestenfalls nur (wenig) verstanden (i. d. R. sprechen Kinder ihre Herkunftssprache nicht, weil sie sich ihrer mangelnden Sprachkenntnisse bewusst sind (eigene Meinung über den Sprachstand meist unbegründet)) – Konsequenz: Russisch fossilisiert oder wird (fast) verlernt bzw. wurde evtl. nie richtig beherrscht 	<ul style="list-style-type: none"> – L1: Russisch, im Vorschulalter zugewandert – Konsequenz: Russisch fossilisiert (Tendenz rückläufig), Deutsch als L1 – L1: Russisch, Einschulung im Ausland und im Grundschulalter zugewandert, dann Grundschule in Deutschland – Konsequenzen: Russisch fossilisiert und Deutsch als L1 oder Russisch fossilisiert und Deutsch als lückenhafte L1 z. B. im Bereich des Wortschatzes (allgemein, Bildungssprache, Fachsprache) – L1: Russisch, zugewandert nach der Grundschule – Konsequenz: Russisch fossilisiert, DaZ (evtl. gleiche Probleme wie im zweiten obigen Fall) – L1: Gastarbeitende („expats“) – Konsequenz: Russisch fossilisiert 	<ul style="list-style-type: none"> – Russisch wird als zweite oder dritte Sprache gesprochen (Deutsch nicht mitgezählt) – Konsequenz: (evtl.) Interesse am Erhalt der russischen Sprache

Eine weitere Gruppe in der dritten Spalte von Tabelle 1 stellen Lernende dar, die Russisch als L1 sprechen, neu zugewandert oder geflüchtet sind und sich noch nicht so lange in Deutschland aufhalten, dass die L1 stagnieren oder gar rückläufig würde. In der letzten Gruppe in der ganz rechten Spalte von Tabelle 1 sind Schüler:innen zusammengefasst, die Russisch zwar auf einem nahezu erstsprachlichen Niveau beherrschen, es aber als dritte, vierte usw. Sprache sprechen. Die

geografische Herkunft dieser Gruppe kann sehr vielfältig sein und muss sich nicht auf ehemalige Sowjetrepubliken beschränken.

1.4 Sprachspezifische Herausforderungen und (Lösungs-)Ansätze einer akkumulierenden Didaktik

Anspruch und Ziel guten Russischunterrichts muss sein, bezüglich aller genannten Zielgruppen bzw. gemäß den Bedürfnissen jeder und jedes Einzelnen eine sprachunterstützende, -förderliche und -entwickelnde Funktion zu erfüllen (Bergmann, 2016; Mengel, 2006; Reich, 2014), auch wenn dies angesichts der außerordentlich großen Heterogenität der Schülerschaft besondere Herausforderungen an diese selbst wie auch an Lehrkräfte stellt. Wie schwierig dies ist, zeigen aktuelle Studien, die eine Sprachattrition in der Herkunftssprache (mittlere Gruppen in Tabelle 1) und mangelnde Kompetenzen in der Fremdsprache Russisch (erste Spalte in Tabelle 1) erkennen lassen, welche nicht nur auf eine unzureichende Differenzierung zurückzuführen sind (Anstatt, 2011; Bergmann, 2017; Olfert, 2019). Als weitere Gründe für eine Sprachattrition im herkunftssprachlichen Unterricht führt Anstatt (2008, S. 6) unter anderem an:

- lebensweltliche Qualität und Quantität des russischen Inputs,
- Vermischung der Sprachen in Gesprächen im familialen Umfeld,
- niedriges Sprachbewusstsein im Elternhaus bzw.
- Einfluss der Schule und der darin dominierenden Mehrheitssprache bei Grundschulkindern.

(Weiter-)Erwerb und Ausbildung der Herkunftssprache werden von der Umgebungssprache Deutsch beeinflusst, sodass die gesprochene Varietät der Herkunftssprache von dem zielsprachlichen Standard (stark) abweichen (Bergmann, 2021, S. 12, 14; Mehlhorn & Rutzen, 2020, S. 223) und ggf. eine Mischvarietät entstehen kann, die Lehrkräfte im herkunftssprachlichen Unterricht wertschätzend aufgreifen und zielführend weiterentwickeln müssen (Lüttenberg, 2010). Hierzu gehören eine Förderung von Sprachbewusstheit und -reflexivität, die sonst eher Ziele muttersprachlichen Unterrichts sind (Bergmann, 2021, S. 13), und eine Steigerung der Sensibilität gegenüber fehlerhaftem Sprachgebrauch. Ähnlich wie in der Muttersprache werden sprachliche Komplexität sowie korrekter schriftlicher und mündlicher Ausdruck angestrebt (Mehlhorn & Rutzen, 2020, S. 222), wobei unterschiedliche Vorerfahrungen und Sprachlernbiografien der Schüler:innen einbezogen werden sollten (Bergmann, 2021, S. 13). Analog zum Unterricht Deutsch als Zweitsprache (DaZ) können Sprachvergleiche zur Veranschaulichung von Gemeinsamkeiten und Unterschieden von Deutsch und Russisch als Lernstrategie genutzt werden, um „Transferbasen in anderen gelernten Sprachen aufzuzeigen sowie das metasprachliche Bewusstsein zu fördern“ (Bredthauer, 2018, S. 278). Ferner wird gezielt an der Registerentwicklung gearbeitet, wie sie auch ein Merkmal des erstsprachlichen Deutschunterrichts ist (Bergmann & Böhmer,

2020, S. 78; Bergmann 2014; Mehlhorn & Rutzen, 2020, S. 223). In dieser Hinsicht orientiert sich der herkunftssprachliche Unterricht zwar an den Niveaustufen des Gemeinsamen europäischen Referenzrahmens für Sprachen (Europarat/Rat für kulturelle Zusammenarbeit, 2001), greift jedoch – im Gegensatz zum Fremdsprachenunterricht – auch auf Modelle (z.B. Mündlichkeit und Schriftlichkeit nach Koch & Österreicher, 1994) sowie Konzepte (wie *language awareness*, Wildemann & Vach, 2020) zurück, die in der Deutsch- und DaZ-Didaktik Anwendung finden. Dies erlaubt „eine andere Perspektive auf das Verhältnis von Mündlichkeit und Schriftlichkeit und den Erwerb von Text(sorten)kompetenz im Sprachenlernen“ (Bergmann & Böhmer, 2020, S. 78). Ein erheblicher Unterschied zum Unterricht in der Fremdsprache besteht außerdem darin, dass derjenige in der Herkunftssprache darauf abzielt, Schüler:innen „ganzheitlich in ihrer sprachlichen, kognitiven, sozialen und emotionalen Entwicklung (zu) unterstütz(en)“ (Sächsisches Staatsministerium für Kultus, 2014, S. 4). Mittels dieses holistischen Ansatzes sollen die Kinder ihre Mehrfachzugehörigkeit und Identität (weiter-)entwickeln können (Rahmenplan Grundschule. Herkunftssprache, Sachsen, 2014, S. 4; Bergmann, 2017, S. 140). Weitere (didaktische) Konvergenzen und Disparitäten zwischen dem herkunfts- und mutter- bzw. fremdsprachlichen Unterricht wurden bislang noch nicht systematisch untersucht (Mehlhorn & Rutzen, 2020, S. 220; Mehlhorn, 2022, S. 2; Bergmann, 2021, S. 11), sind allerdings anzunehmen.

Demgegenüber verfolgt der Fremdsprachenunterricht Russisch den Erwerb fremdsprachlicher und interkultureller Kompetenzen gemäß den Niveaus des Gemeinsamen europäischen Referenzrahmens für Sprachen (Europarat/Rat für kulturelle Zusammenarbeit, 2001). Ein besonderer Wert wird dabei auf die Entwicklung mündlicher kommunikativer Fähigkeiten gelegt. Der bereitgestellte Input ist insofern künstlich konstruiert, als dass er keinen oder nur begrenzten Bezug zur alltäglichen Lebenswelt der Lernenden aufweist. Die Verwendung der Sprache erfüllt also keine realitätsgebundene kommunikative Funktion (Bergmann, 2016), wenngleich die Herkunftssprecher:innen in diesem Zusammenhang eine didaktisch reichhaltige Ressource darstellen könnten. Diese zu nutzen, ohne Schüler:innen ohne Migrationsgeschichte und Vorkenntnisse aufgrund der sprachlichen Überlegenheit jener zu demotivieren (Bergmann, 2016, S. 47), wäre umso wichtiger, da sprachliche Handlungen auf Russisch in solchen Kontexten bzw. Gruppenkonstellationen in hohem Maße authentisch sein könnten. Hierfür sind von der Fachdidaktik allerdings erst noch tragfähige Konzepte zu entwickeln, welche die heterogenen Gruppen von Russischlernenden (vgl. Tab. 1) entsprechend ihrer unterschiedlichen Voraussetzungen und individuellen Fähigkeiten adäquat berücksichtigen, motivieren und einbeziehen. Ansatzpunkte hierfür bieten verschiedene Bereiche angrenzender Wissenschaftsdisziplinen, von der allgemeinen Sprachdidaktik bis zur DaF/DaZ-Didaktik (vgl. Abb. 2), und können für sich wie auch akkumuliert wertvolle Impulse zur Weiterentwicklung der Didaktik und Methodik des Herkunfts- bzw. Fremdsprachenunterrichts Russisch geben.

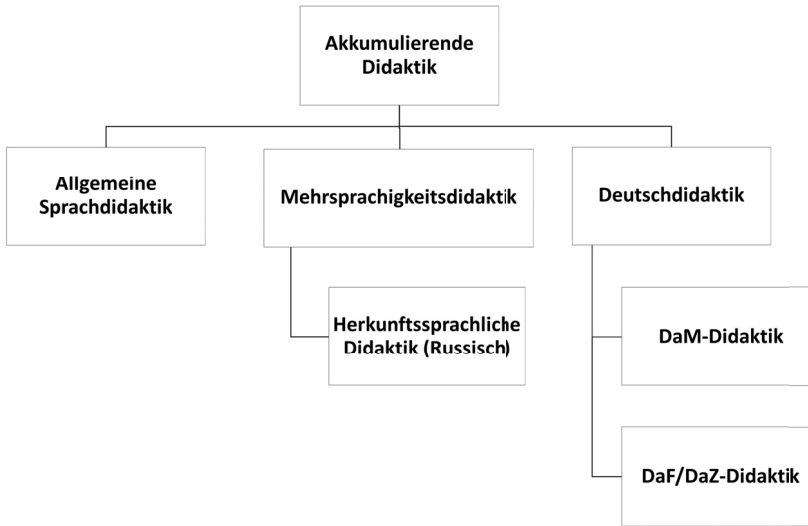


Abbildung 2: Vorschlag einer akkumulierenden Didaktik

Ein derart interdidaktisches Zusammenführen verschiedener Konzepte unter fachbezogener Perspektive ermöglicht die Genese theoretisch breit fundierter, ansprechender und differenzierter Lernangebote für heterogene Leistungsgruppen (Belyaev, 2023), methodisch-didaktische Lernräume, welche Schüler:innen unterstützen, ihre Erstsprachen sowie ihre Erfahrungen im Umgang mit anderen Sprachen zum Erwerb und zur Weiterentwicklung einer neuen Sprache (Russisch oder DaZ) zu nutzen. Ein solches integratives Konzept, das in Abbildung 2 als akkumulierende Didaktik vorgestellt wird und die Heterogenität der Schülerschaft produktiv nutzbar macht, kann zugleich die Unterschiede zwischen herkunfts- und fremdsprachlichem Unterricht nivellieren. Die Übergänge zwischen der einen Form, an der Herkunftssprecher:innen mit sehr unterschiedlichen Sprachniveaus partizipieren, sowie der anderen Ausprägung, in der Sprachneulinge mit Herkunftssprecherinnen und -sprechern unterschiedlicher Provenienz interagieren, sind fließend. Daher sollten diese passender als *Herkunftssprachenunterricht Russisch mit fremdsprachlichen Anteilen* bzw. als *Fremdsprachenunterricht Russisch mit herkunftssprachlichen Anteilen* bezeichnet werden. Die Qualitätskriterien, die an die Lehr- und Lernprozesse in diesen Formaten anzulegen sind, unterscheiden sich indes nicht (Mehlhorn, 2022, S. 3), sodass sich die nachfolgenden Ausführungen auf beide beziehen.

2 Unterrichtsqualität und das Schulfach Russisch

Im Gegensatz zur Bildungswissenschaft und zu anderen Fachdidaktiken (vgl. Lindl et al., in diesem Band) spielen Merkmale von Unterrichtsqualität im russischdidaktischen Diskurs bislang kaum eine Rolle. Die Forschung beschränkt sich auf einzelne, skizzenhafte theoretische Überlegungen (z.B. zu Kompetenzorientierung, Aktivierung von Schülerinnen und Schülern, Leistungsbewertung; Bergmann, 2016, S. 176; Bergmann, 2017, S. 133; Olfert, 2019, S. 135), die als Ansatzpunkte für (mögliche) Indikatoren qualitätvollen Russischunterrichts gelten könnten. Systematische (empirische) Betrachtungen oder Modellierungen stehen bisher aber aus, wenngleich deren Notwendigkeit erkannt wird (Bergmann, 2014).

Vor diesem defizitären Forschungshintergrund kann und will der vorliegende Beitrag nicht den Anspruch erheben, ein vollständiges und umfassendes Modell zu Qualitätsmerkmalen des Schulfachs Russisch als Herkunft- und Fremdsprache zu erstellen. Vielmehr soll es primär darum gehen, diese Forschungsaufgabe aus der eigenen fachbezogenen Perspektive anzudenken und unter Rekurs auf das Syntheseframework von Praetorius et al. (2020) bzw. dessen Weiterentwicklung im MAIN-TEACH-Modell (Ufer & Praetorius, 2022) erstmalig zu erschließen (eine ausführliche Erläuterung des Modells ist in der Einleitung zu diesem Band zu finden; vgl. Lindl et al.). Hierzu werden zunächst zentrale Desiderate kurz rekapituliert, die für eine vertiefte Auseinandersetzung mit dem Thema Unterrichtsqualität im Fach Russisch von Bedeutung sind. Daraufhin werden die Dimensionen und Subdimensionen der genannten Modelle aus fachspezifischer Perspektive kritisch reflektiert und vor dieser Folie potenzielle Merkmale guten Russischunterrichts entwickelt und erwogen.

2.1 Desiderate der russischdidaktischen Unterrichtsforschung

Mit einer Beschäftigung mit Aspekten von Unterrichtsqualität sind aus fachlicher Sicht unter anderem folgende Themen eng verknüpft (vgl. auch die exemplarischen Fragestellungen im Einleitungskapitel von Lindl et al., in diesem Band) und bedürfen einer genaueren Betrachtung sowie systematischen Untersuchung, die in der Regel noch aussteht:

- Auswahl und Sequenzierung von Inhalten des Herkunft- und Fremdsprachenunterrichts Russisch vom Primar- bis zum Sekundarbereich II (Mehlhorn, 2022, S. 2; Mehlhorn & Rutzen, 2020, S. 220);
- Alltagsbezug, Motivation und Prestige: Eine maßgebliche Rolle für die Gestaltung des Angebots seitens der Lehrkraft und dessen Nutzung spielt der Faktor gesellschaftliches Prestige und aus der Sicht der Herkunftssprecher:innen zusätzlich familieninternes Prestige des Russischen, zumal dieses nicht zu ‚prestigetragtigen‘ Sprachen wie Englisch, Französisch, Spanisch oder Latein zählt (Anstatt, 2011; Bergmann, 2014, 2017; Bredthauer & Engfer, 2018; Dirim, 2016, 2017). Selbst diejenigen Familien, in denen Russisch als Familiensprache ge-

sprochen wird, präferieren für Kinder prestigeträchtigere Sprachen und weisen familienintern wenig Bewusstsein für Russisch auf (Bergmann, 2014; Lange & Pohlmann-Rother, 2020; Lüttenberg, 2010; Mehlhorn & Bremer, 2018;). Dabei spielt nicht nur Motivation eine Rolle, sondern auch die Frage der Selbstpositionierung, das Gefühl der Mehrfachzugehörigkeit und die eigene Identität (Bergmann, 2014; Dirim, 2017; Helmke, 2022). Demnach unterscheidet sich der herkunfts- oder fremdsprachliche Russisch- vom herkömmlichen Fremdsprachenunterricht auch durch den persönlichen Bezug von Herkunftssprecherinnen und -sprechern (Mehlhorn, 2022, S. 3), was jedoch nicht unbedingt mit einer erhöhten Lernmotivation einhergehen muss (Anstatt, 2008; Bergmann, 2016, 2017).

- Umgang mit der Heterogenität der Schülerschaft: umfassende Analyse der individuellen Lernbedürfnisse und -ziele der Schüler:innen in den heterogenen Lerngruppen;
- didaktische Unterrichtskonzepte für heterogene Lernendengruppen: Welche spezifischen Methoden, Strategien, Materialien und Prozesse werden eingesetzt, um sowohl Herkunftssprecher:innen als auch Lernende ohne Vorkenntnisse in einer Klasse individuell und in Gruppen angemessen zu fördern? Wo liegen didaktische und methodische Grenzen? Welche Unterrichtskonzepte und -prozesse erweisen sich als besonders effektiv und in welcher Hinsicht?
- Vergleich zwischen Russisch als Herkunft- und Fremdsprache: kriterienbasierte Untersuchung der Unterschiede in Konzeption und Durchführung zwischen beiden Lehrformen (auch an bilingualen Schulen sowie als erste, zweite etc. Fremdsprache);
- Evaluierung der Rolle der deutschsprachigen Umgebung: Wie kann die deutschsprachige Umgebung ideal in den Russischunterricht einbezogen und als Ressource genutzt werden, um Lernenden ein besseres Verständnis und eine effektivere Anwendung der russischen Sprache zu vermitteln?
- Mediation und Translanguaging: Wie können diese effektiv in (sprachlich heterogenen) Gruppen für Lernzwecke eingesetzt werden? (Wie) Lassen sie sich mit entdeckendem Lernen kombinieren?
- langfristige Lernergebnisse und -effekte des Russischunterrichts in heterogenen Lerngruppen: Wie wirkt sich dieser auf die Entwicklung von Herkunftssprecherinnen und -sprechern sowie Lernende ohne Vorkenntnisse sowohl in Bezug auf ihre sprachlichen Fähigkeiten (Zertifikatsprüfung) als auch auf ihr kulturelles Verständnis aus?

Auch wenn diese Forschungsdesiderate das empirisch abgesicherte, verfügbare Wissen über Qualitätsmerkmale von Russischunterricht einschränken, sollen hierzu im Folgenden einige Vorschläge auf theoretischer Basis erarbeitet und diskutiert werden (Abschnitt 2.3). Diese schließen an aktuelle bildungswissenschaftliche Modelle zur Beschreibung von Unterrichtsqualität an (Abschnitt 2.2) und greifen als konzeptuellen Orientierungspunkt auf dasjenige zurück, dessen inhaltliche Kategorien auf die Gegebenheiten des Russischunterrichts in Deutschland möglichst übertragbar bzw. daran anpassungsfähig sind.

2.2 Impulse aus der Bildungswissenschaft

In der Forschung zu Unterrichtsqualität (z. B. Helmke & Schrader, 2008; Praetorius et al., 2020; Wiprächtiger-Geppert et al., 2021; s. auch Einleitung zu diesem Band) werden unterschiedliche Modelle diskutiert und Synopsen erstellt. Diese Modelle beschäftigen sich einerseits nur mit fächerübergreifenden Merkmalen oder Bereichen wie beispielsweise Helmke (2022), der insgesamt fünf fachübergreifende Qualitätsbereiche nach einer kurzen Vorstellung anderer Modelle definiert und diese Bereiche dann ausführlich erörtert (Helmke, 2022). Andererseits stellen manche Autorinnen und Autoren die klassischen drei Basisdimensionen ausführlicheren moderneren Modellen gegenüber, die auch fachspezifische Aspekte beinhalten. So weisen z. B. Praetorius et al. (2020a; 2020b) auf die blinden Flecken des Modells der drei Basisdimensionen hin und stellen als Alternative das auf den Vorschlag von Praetorius und Charalambous (2018) aufbauende Syntheseframework vor. Dieses umfasst generische und fachspezifische (Sub)Dimensionen und Indikatoren, die sich auf eine prototypische Lehr-Lernsequenz beziehen und fachbezogen anpassen lassen, um eine umfassende Beschreibung von Unterrichtsqualität zu erlauben. Das Modell gewährleistet, dass entscheidende Aspekte wie Differenzierung und Adaptivität, Themenauswahl etc. angemessen berücksichtigt werden können. Damit unterstützt es nicht nur Lehrkräfte bei der Planung und Verbesserung ihres Unterrichts, sondern kann auch einer effektiveren Gestaltung des Russischunterrichts dienen. Für eine detaillierte Darstellung des Modells sei an dieser Stelle auf den Beitrag von Praetorius et al. (2020b) und auf den Einleitungsbeitrag in diesem Band (Lindl et al.) verwiesen.

Das Syntheseframework, das unter anderem im MAIN-TEACH-Modell aufgeht (Ufer & Praetorius, 2022), bietet sich als Ausgangsmodell vor allem an, weil es trotz eines generischen Rahmenkonzepts die Besonderheiten eines Fachs berücksichtigen kann und interdisziplinäre Vergleiche beispielsweise zwischen einzelnen (Fremd-)Sprachen in der Schule ermöglicht (Praetorius et al., 2020, S. 432). Dies ist sowohl für die fachspezifische wie auch interdisziplinäre didaktische Forschung nicht unbedeutend. Somit scheint es aktuell das vorteilhafteste Referenzmodell zu sein (siehe ausführlicher in der Einleitung zu diesem Band).

2.3 Überlegungen zu Merkmalen guten Russischunterrichts

Angesichts der Besonderheiten des Schulfachs Russisch sind einige (fachspezifische) Anpassungen auf Ebene der Subdimensionen und Indikatoren des Syntheseframeworks oder MAIN-TEACH-Modells zu erwägen, die bei Überlegungen zu Merkmalen guten Russischunterrichts als Ausgangspunkte dienen können. So knüpfen der Ausbau (fremd-)sprachlicher Kompetenzen und die Entwicklung schriftsprachlicher Fähigkeiten nicht nur bzw. nicht bei allen Lernenden an vorliterale Erfahrungen an, wohl aber an Erfahrungen durchgehender, fächerübergreifender, unterrichtsbezogener sowie außerunterrichtlicher Sprachbildung (Anstatt,

2011). Eine stärkere Berücksichtigung der lebensweltlichen Perspektive und der subjektiven Nützlichkeitswahrnehmung des Russischen jeder Schülerin und jedes Schülers stellt somit ein gewinnbringendes Merkmal qualitativ anspruchsvollen Russischunterrichts dar. Als solches gilt für Herkunftssprecher:innen zudem die Ausrichtung auf die Entwicklung schriftsprachlicher Kompetenzen, Aneignung von höheren sprachlichen Registern sowie vertiefende Auseinandersetzungen mit der russischsprachigen Literatur und Kultur. Bei Nicht-Muttersprachlerinnen und -sprachlern stehen anwendungsorientierte mündliche Kompetenzen und der Erwerb bzw. Ausbau von pragmatischen Kompetenzen in Anknüpfung an die Teilkompetenzen Lesen und Schreiben für ein angemessenes zielsprachliches Handeln im Vordergrund. Sprachlicher In- und Output sind dabei jedoch quantitativ und qualitativ anders als in der zielsprachlichen Umgebung. Dies sollte beim Lehren und Lernen besonders beachtet werden, da Russisch im Unterricht gleichzeitig als Lerngegenstand und als Medium fungiert (vgl. für Englisch Wilden, 2021).

Eine Rücksichtnahme auf die deutschsprachige Umgebung und Schaffung einer entsprechenden Unterrichtskultur, die sich unter anderem mit der deutsch-russischen Mischvarietät (Lüttenberg, 2010) auseinandersetzt und sprachvergleichend vorgeht, stellt ein weiteres Qualitätsmerkmal dar. Darin sollen Mediation und Translanguaging einen eigenen Platz einnehmen und lernförderlich genutzt werden. Von großer Relevanz erscheint in diesem Kontext, dass die Fachdidaktik Russisch diesbezüglich neue Ansätze anderer Sprachdidaktiken (v.a. der DaZ-Didaktik) rezipiert und von interdidaktischen Synergien profitiert. Sie sollte Erkenntnisse der Deutsch-, Fremdsprachen- oder Herkunftssprachendidaktik nutzen, um erforderliche Konzepte für einen sprachlich heterogenen und mehrsprachigen Fremdsprachenunterricht in Russisch zu entwickeln (Dirim & Mecheril, 2018; Reich, 2014). Hilfreich hierfür wären insbesondere Lehrkonzepte für Russisch als L3 und Lx sowie kurze Lehrbücher, die durch ihr übersichtliches Format einem Motivationsverlust seitens der Schüler:innen vorbeugen.

Auch die Integration der Grammatik in den Unterrichtsverlauf und die Art ihrer Vermittlung bedürfen innovativer, nach einschlägigen Qualitätskriterien ausgerichteter Ansätze, deren Definition und Identifikation auf Basis von Lehrere Erfahrungen mit herkunftssprachlichen Lernenden von der Grundschule bis zum Sekundarbereich II ebenfalls noch eine offene Forschungsaufgabe darstellt (Bergmann, 2021, S. 11). Denn es ist anzunehmen, dass diese nicht nur vom jeweiligen Unterrichtsformat und der Gruppenkonstellation abhängig sind, sondern sich auch nach Jahrgangs- und Altersstufen unterscheiden. Angesichts der deutschen Sprachumgebung – sei es mit muttersprachlich deutschen Schülerinnen und Schülern, innerhalb der Schule selbst oder in der weiteren sprachlichen Umgebung –, wären in Bezug auf Aspekte von Unterrichtsqualität auch Vergleiche mit dem Deutschunterricht interessant, um Transfermöglichkeiten und (interdidaktische) Optimierungspotenziale zu erkennen.

Eine klare Ausrichtung auf bestimmte Zielkriterien wie die Vorbereitung auf eine Zertifikatsprüfung wird schließlich auch noch als wesentlicher Aspekt qualitätsvollen Russischunterrichts angesehen, der gerade für die Wahl und das (Wei-

ter-)Lernen der Sprache wichtig ist. Dies gilt insbesondere für die Gruppe der Herkunftssprecher:innen, jedoch nicht ausschließlich. Laut Mengel (2006) liegt nämlich die Motivation, das Fach Russisch in der Schule zu belegen, unter anderem in der Aussicht begründet, am Ende eine Sprachprüfung zu absolvieren. Diese war bis dato an dem russischen Zertifizierungssystem orientiert, welches nur pro forma an die europäischen Standards bzw. Niveaustufen nach dem Gemeinsamen europäischen Referenzrahmen für Sprachen (Europarat/Rat für kulturelle Zusammenarbeit, 2001) angepasst wurde.

Zusammenfassend lassen sich also folgende Vorschläge für Qualitätsmerkmale für das Schulfach Russisch formulieren, deren Zusammensetzung und Bedeutung sich je nachdem, ob sie sich auf Herkunfts- (HSU) oder Fremdsprachenunterricht (FSU) beziehen, unterscheiden können:

- Berücksichtigung (pädagogisch, didaktisch und methodisch) der Heterogenität der Lernendengruppe (Russisch als Erst-, Herkunfts-, Fremdsprache, Lx),
- Rücksichtnahme auf die deutschsprachige Umgebung und die Anwendung eines sprachkontrastiven und reflexiven Ansatzes,
- angemessene Förderung der Registerentwicklung bei Herkunftssprecherinnen und -sprechern sowie effektiver Umgang mit Mischvarietät,
- kleinschrittiges Vorgehen (FSU) vs. weniger kleinschrittiges Vorgehen (HSU; FSU in Bezug auf Herkunftssprecher:innen),
- Einbezug literarischer Texte unter Berücksichtigung kultureller Kontexte (im Klassenzimmer) (HSU),
- Schulung von Sprachbewusstheit und Vermittlung von Lernstrategien (HSU),
- Anpassung von mündlicher und schriftlicher Kommunikation an die Bedürfnisse und sprachlichen Fähigkeiten der Lernendengruppen,
- Berücksichtigung der individuellen Erfahrungen und Perspektiven der Lernenden aus dem persönlichen und lebensweltlichen Kontext,
- Förderung der Lernmotivation der Schüler:innen,
- angemessene (d. h. auch möglichst authentische), differenzierte Aufgaben für alle Lernendengruppen (konzept- und sprachdatengeleitet; adäquates ziel sprachliches Handeln in Anknüpfung an Lesen und Schreiben),
- effektive Integration der Grammatikvermittlung im Kontext des Kompetenzerwerbs,
- Förderung von Kommunikation zwischen Lernendengruppen im Unterricht (FSU),
- Sprache als Unterrichtsgegenstand und als Medium (HSU),
- individuelle Sprachförderung (ausgehend vom sprachlernbiografischen Hintergrund),
- Mediation und Translanguaging,
- angemessene Lehr-Lernkonzepte für Russisch als L3 und Lx sowie verkürzte Lehrbücher mit aktuellen Themen für Nicht-Erstsprachler:innen.

Bezugnehmend auf das Syntheseframework und das MAIN-TEACH-Modell von Praetorius et al. (2020) können somit generische wie auch fachspezifische Prä-

zisierung auf subdimensionaler und Indikatorebene vorgenommen werden. Die fachspezifischen Überlegungen, Ergänzungen und Anpassungen betreffen alle darin formulierten Dimensionen von Unterrichtsqualität, insbesondere jedoch die Dimensionen *Auswahl und Thematisierung von Inhalten*, *kognitive Aktivierung*, *Unterstützung des Lernens aller Schüler:innen* und *sozio-emotionale Unterstützung* (sowohl für Schüler:innen wie auch für Lehrpersonen; Praetorius et al., 2020).

3 Fazit und Ausblick

Da Qualitätsmerkmale von Russischunterricht in deutscher Sprachumgebung bislang noch nicht im Fokus der forschenden Fachdidaktik standen, ging dieser Beitrag zunächst von bildungswissenschaftlichen Modellen und Konzeptualisierungen aus, um auf diesen theoretischen Annahmen fußend fachspezifische Ergänzungen und Erweiterungen zu identifizieren, zu reflektieren und diskursiv auszuhandeln. Die hier erstmalig vorgeschlagenen Aspekte sollen Ansätze und Impulse für weiterführende Überlegungen, Modellierungen und empirische Untersuchungen bieten und damit einerseits den fachspezifischen, andererseits den interdisziplinären und bildungswissenschaftlichen Diskurs voranbringen. Sie erheben deshalb keinerlei Anspruch auf Vollständigkeit, Endgültigkeit oder Wirksamkeit, welche hinsichtlich der verschiedenen Zielsetzungen von Russischunterricht erst noch empirisch zu überprüfen ist. Nach erfolgreicher Revision und Validierung sollte ein Modell zu Qualitätsmerkmalen für Russischunterricht in der Schule als überschaubarer Referenzrahmen für eine reflektierte Praxis nutzbar sein, beispielsweise indem darin nur nachweislich wirksame Aspekte aufgenommen werden. Diese können Lehrkräften als nützliche Orientierungshilfen für didaktische Entscheidungen bei der Unterrichtsplanung, -reflexion und -evaluation, d.h. bei der Unterrichtssteuerung in einer infernalisch heterogenen Lernenden-Gruppe, dienen (Helmke, 2007; 2022).

Abgesehen von ihren Wirkungen hinsichtlich unterrichtlicher Zielkriterien sollte in einem weiteren Schritt auch untersucht werden, wie sich qualitätsvolle Lehr-Lernprozesse im Russischunterricht erzeugen, beeinflussen und gewährleisten lassen. Diesbezüglich schreiben gängige unterrichtliche Wirksamkeitsannahmen wie im Kaskadenmodell (Lindl et al., in diesem Band) der professionellen Kompetenz von Lehrkräften eine entscheidende Rolle zu. Diese kommt bei der Planung, Durchführung und Weiterentwicklung des Russischunterrichts zum Tragen und ist insbesondere in Kontakt mit heterogenen Lernendengruppen und herkunftssprachlichen Schülerinnen und Schülern erforderlich (Bergmann, 2016). Denn gerade deutschsprachige Lehrkräfte stehen hier vor der großen Herausforderung, den Lernendengruppen – allen voran den Herkunftssprecherinnen und -sprechern – mit ihren eigenen notwendigerweise sehr hohen Sprachkompetenzen und einem differenzierten nicht nur konzept-, sondern auch sprachdatengeleiteten Lernangebot, d.h. methodisch-didaktischer Versiertheit in verschiedenen Themenbereichen, gerecht zu werden. Sie müssen ihren Unterricht adressatenori-

entiert ausrichten und diverse sprachliche Kompetenzen der heterogenen Klasse angemessen fördern (Bergmann, 2016). Welche Aspekte professioneller Kompetenz von Russischlehrkräften hierfür relevant sind und sich wie auf eine qualitätsvolle Gestaltung von Lehr- und Lernprozessen auswirken, ist bislang allerdings ungeklärt.

Literatur

- Anstatt, T. (2008). Russisch in Deutschland: Entwicklungsperspektiven. *Bulletin der deutschen Slavistik*, 14, 67–74.
- Anstatt, T. (2011). Sprachattrition. Abbau der Erstsprache bei russisch-deutschen Jugendlichen. *Wiener Slavistischer Almanach*, 67, 7–31.
- Belyaev, D. (2023). „Bist du shy/Shay شاي?“ – Wenn sprachliche Heterogenität den Rahmen bricht. https://www.ger.ovgu.de/iger_media/Fachgebiete/Fachdidaktik+Deutsch/Tagungsprogramm.pdf
- Bergmann, A. (Hrsg.) (2014). *Fachdidaktik Russisch. Eine Einführung*. Tübingen: Narr Francke Attempto.
- Bergmann, A. (Hrsg.) (2016). *Kompetenzorientierung und Schüleraktivierung im Russischunterricht*. Frankfurt: Peter Lang. <https://doi.org/10.3726/978-3-653-04271-9>
- Bergmann, A. (2017). Russischunterricht im deutschen Bildungssystem: Traditionen, gesellschaftliche Anforderungen und bildungspolitische Perspektiven. In C. Yildiz, N. Topaj, R. Thomas & I. Gülzow (Hrsg.), *Zukunft der Mehrsprachigkeit im deutschen Bildungssystem: Russisch und Türkisch im Fokus* (S. 129–153). Frankfurt: Peter Lang.
- Bergmann, A. (2021). Zur Einführung: Russisch als Herkunftssprache in Deutschland unterrichten. *Slavic Language Education*, 1, 5–18.
- Bergmann, A. & Böhmer, J. (2020). Perspektiven auf Heterogenität: Lehrpläne für Russisch als Fremdsprache und Russisch als Herkunftssprache. *Die Neueren Sprachen. Jahrbuch des Gesamtverbands moderne Fremdsprachen*, 8/9, 68–81.
- Bredthauer, S. (2018). Mehrsprachigkeitsdidaktik an deutschen Schulen – eine Zwischenbilanz. *DDS – Die Deutsche Schule*, 110(3), 275–286. <https://doi.org/10.31244/dds.2018.03.08>
- Bredthauer, S. & Engfer, H. (2018). *Natürlich ist Mehrsprachigkeit toll! Aber was hat das mit meinem Unterricht zu tun?* <https://kups.ub.uni-koeln.de/8092/>
- Busch, B. (2021). *Mehrsprachigkeit*. Wien: Facultas. <https://doi.org/10.36198/97838388556529>
- Destatis = Statistisches Bundesamt (2024). *Zahl der Woche Nr. 13 vom 29. März 2022: 94 000 Schülerinnen und Schüler lernten im Schuljahr 2020/2021 Russisch als Fremdsprache*. https://www.destatis.de/DE/Presse/Pressemitteilungen/Zahl-der-Woche/2022/PD22_13_p002.html
- Dirim, I. (2016). „Ich wollte nie, dass die anderen merken, dass wir zu Hause Arabisch sprechen“. Perspektiven einer linguizismuskritischen pädagogischen Professionalität von Lehrerinnen und Lehrern. In M. Hummrich, I. Dirim & C. Freitag (Hrsg.), *Kulturen der Bildung* (S. 191–207). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-10005-6_16
- Dirim, I. (2017). *Linguizismus und linguizismuskritische pädagogische Professionalität*. https://wp.sung.sk/wp-content/uploads/2020/07/SZfG_2017_1_7.pdf

- Dirim, I. & Mecheril, P. (2018). *Heterogenität, Sprache(n) und Bildung. Eine differenz- und diskriminierungstheoretische Einführung*. Bad Heilbrunn: Julius Klinkhardt. <https://doi.org/10.36198/9783838544434>
- Europarat/Rat für kulturelle Zusammenarbeit (Hrsg.) (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Berlin u. a.: Langenscheidt.
- Helmke, A. (2007). *Was wissen wir über guten Unterricht? Wissenschaftliche Erkenntnisse zur Unterrichtsforschung und Konsequenzen für die Unterrichtsentwicklung*. https://www.bildung.koeln.de/imperia/md/content/selbst_schule/downloads/andreas_helmke_pdf
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung. Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Hannover: Klett Kallmeyer.
- Helmke, A. & Schrader, F.-W. (2008). Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen. *SEMINAR – Lehrerbildung und Schule*, 3, 17–47. [http://www.unterrichtsdiagnostik.info/media/files/publ/Helmke%20Schrader%20\(2008\)_Merkmale%20der%20Unterrichtsqualit%C3%A4t%20-%20Potenzial%20Reichweite%20und%20Grenzen%20SEMINAR%203-2008%20S.17-47.pdf](http://www.unterrichtsdiagnostik.info/media/files/publ/Helmke%20Schrader%20(2008)_Merkmale%20der%20Unterrichtsqualit%C3%A4t%20-%20Potenzial%20Reichweite%20und%20Grenzen%20SEMINAR%203-2008%20S.17-47.pdf)
- Lange, S. D. & Pohlmann-Rother, S. (2020). Überzeugungen von Grundschullehrkräften zum Umgang mit nicht-deutschen Erstsprachen im Unterricht. *Zeitschrift für Bildungsforschung*, 10(1), 43–60. <https://doi.org/10.1007/s35834-020-00265-4>
- Lüttenberg, D. (2010). Mehrsprachigkeit, Familiensprache, Herkunftssprache. Begriffsvielfalt und Perspektiven für die Sprachdidaktik. *Wirkendes Wort*, 2, 299–315.
- Mehlhorn, G. (2022). Unterricht in der Herkunftssprache – Zum Forschungsstand. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 27(2), 1–16.
- Mehlhorn, G. & Brehmer, B. (Hrsg.) (2018). *Potenziale von Herkunftssprachen. Sprachliche und außersprachliche Einflussfaktoren*. Tübingen: Stauffenburg.
- Mehlhorn, G. & Rutzen, K. M. (2020). Didaktische Prinzipien für den Unterricht mit Herkunftssprachen- und Fremdsprachenlernenden. In I. Gogolin, A. Hansen, S. McMonagle & D. Rauch (Hrsg.), *Handbuch Mehrsprachigkeit und Bildung* (S. 219–225). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-20285-9_32
- Mengel, S. (2006). Zertifikatsprüfungen für Russisch – Notwendigkeit und Spezifik (mit H. Eckert). In D. Rytel-Kuc & J. Tambor (Hrsg.), *Sprachpolitik und Zertifizierung* (S. 159–166). Katowice/Leipzig: Wydawnictwo Gnome.
- Olfert, H. (2019). *Spracherhalt und Sprachverlust bei Jugendlichen. Eine Analyse begünstigender und hemmender Faktoren für Spracherhalt im Kontext von Migration*. Tübingen: Narr Francke Attempto.
- Praetorius, A.-K., Herrmann, E. G., Zülsdorf-Kersting, M., Heinz, B. & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Rogh, W. & Kleickmann, T. (2020). Blinde Flecken der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft*, 48(3), 303–318. <https://doi.org/10.1007/s42010-020-00072-w>
- Reich, H. (2014). *Über die Zukunft des herkunftssprachlichen Unterrichts. Überarbeitete Fassung eines Vortrags bei der GEW Rheinland-Pfalz in Mainz am 31.01.2012*. https://www.uni-due.de/imperia/md/content/prodaz/reich_hsu_prodaz.pdf
- Sächsisches Staatsministerium für Kultus (2014). *Rahmenplan Grundschule – Herkunftssprache*. https://migration.bildung.sachsen.de/download/21_02_11_Rahmenplan_GS_Herkunftssprache_2014.pdf

- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Wildemann, A. & Vach, K. (2020). *Deutsch unterrichten in der Grundschule. Kompetenzen fördern, Lernumgebungen gestalten*. Hannover: Kallmeyer/Klett.
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>
- Wiprächtiger-Geppert, M., Stahns, R. & Riegler, S. (2021). Fachspezifität von Unterrichtsqualität in der Deutschdidaktik. *Unterrichtswissenschaft*, 49(2), 203–209. <https://doi.org/10.1007/s42010-021-00109-8>

Wer Ohren hat zu hören

Qualitätskriterien für guten Hörverstehensunterricht aus der Perspektive von Englischfachleitungen

Abstract (der Herausgeber:innen)

In his article, “Who has ears to hear – Quality criteria for good listening comprehension lessons from the perspective of experts in teaching English as a foreign language,” Jens-Folkert Folkerts addresses the question of what actually constitutes good listening comprehension lessons. To this end, he first describes the key processes that occur during listening and characterises the distinctive features of trans- and interactional listening. The article’s core is a qualitative interview study in which seven distinguished experts in teaching English as a foreign language were asked about their understanding of the quality criteria for good listening comprehension lessons. The study and its results, which are presented here, form part of an overarching design-based research project that aims to develop a comprehensive criteria grid for competence-oriented listening comprehension lessons.

1 Einleitung

„Deciens repetita placebit“ – „Zum zehnten Mal wiederholt,
wird sie gefallen.“ (Horaz, *Ars Poetica*, 365)

Die Wiederholung – wenn auch nicht zehnmals – ist eines der markantesten Merkmale des Hörverstehensunterrichts in Deutschland. Es ist jedoch sowohl unklar, ob diese gängige Praxis überhaupt guter Unterricht ist, als auch, was denn guten Hörverstehensunterricht ausmacht. Im vorliegenden Beitrag werden Ergebnisse einer qualitativen Vorstudie dargestellt, in der sieben Englischfachleitungen für Gymnasien und Gesamtschulen im Rahmen eines Leitfadenterviews zu ihrem Verständnis hinsichtlich Qualitätskriterien für guten Hörverstehensunterricht befragt wurden.

Mündlichkeitskompetenz, d. h. die kommunikativen Kompetenzen Sprechen und Hören (Matz, Rumlich & Rogge, 2023; Reckermann et al., 2024), hat in den vergangenen Jahrzehnten zunehmend an Bedeutung für den englischen Fremdsprachenunterricht im deutschen Schulkontext gewonnen, was sich beispielsweise in den bildungspolitischen Vorgaben ausdrückt (z. B. KMK, 2023). Die Vorgaben zur Stärkung der Mündlichkeit bedingen nicht nur, dass Lehrkräfte diese Kompe-

tenzen fokussieren und gezielt unterrichten, sondern sie müssen darüber hinaus auch die Fähigkeiten und Fertigkeiten der Lernenden diagnostizieren und messen. Dies wird für Hörverstehen summativ beispielsweise standardorientiert im Rahmen von VERA8 (VERgleichsArbeiten in Jahrgang 8) durchgeführt. Auch die zentralen Prüfungen (ZP10 – Zentrale Prüfungen am Ende von Jahrgang 10 – und Zentralabitur) der deutschen Bundesländer beinhalten zunehmend eine isolierte Überprüfung der Hörverstehenskompetenz. Wenig überraschend ist somit, dass die fremdsprachliche Forschung zum Bereich Hörverstehen zu einem Großteil auf die valide Überprüfung dieser Kompetenz ausgerichtet ist (vgl. insbesondere Buck, 2001; Rossa & Nold, 2007; Rossa, 2012; Field, 2019).

Bei einem Blick auf die internationale Forschungslage zur Förderung des fremdsprachlichen Hörverstehens fällt auf, dass vielversprechende theoretisch-konzeptionelle Modelle, insbesondere zu prozessorientiertem Hörverstehensunterricht (u. a. Field, 2009) und zur Schulung metakognitiver Hörverstehensstrategien (u. a. Goh & Vandergrift, 2022) vorliegen. Da es bisher jedoch sehr wenige empirische Hinweise zum fremdsprachlichen Hörverstehensunterricht in Deutschland gibt, ist unklar, ob diese Modelle dort auch (umfangreich) eingesetzt werden. Ein erster Versuch sich dieser Informationslücke zu nähern, wird von Folkerts & Matz (2024) sowie von Mertes (2019) unternommen. Der Beitrag von Folkerts & Matz (2024) erhebt in einer *small scale*-Studie die Perspektive der Lehrenden und Lernenden, die Dissertation von Mertes (2019) fokussiert sich auf die subjektiven Theorien von Lehrkräften zum Hörverstehensunterricht. Mit dem vorliegenden Beitrag wird die Perspektive von Fachleitungen für Englisch an Gymnasien und Gesamtschulen in Hinblick auf guten Hörverstehensunterricht näher erschlossen. Ferner werden die Aussagen der Fachleitungen auf das in diesem Band diskutierte MAIN-TEACH-Modell (=Multi-layered and integrated in conceptualizing the quality of teaching, Lindl et al., in diesem Band; Praetorius & Gräsel, 2021; Ufer & Praetorius, 2022) bezogen und auf Implikationen für fachspezifische Ausprägungen hin untersucht.

Im Folgenden wird zunächst der theoretische Rahmen abgesteckt, indem eine Konzeptionalisierung der Elemente des Hörverstehensunterrichts vorgestellt wird. Ausgehend von den Befunden der Studie, welche zum Ziel hat, spezifische Qualitätsmerkmale von Hörverstehensunterricht aufzudecken, werden die theoretisch dargestellten Konzepte weiterentwickelt. Folgende Forschungsfrage wird in diesem Artikel untersucht:

Was sind aus Sicht von Expert:innen für Unterrichtsqualität¹ spezifische Merkmale von und Bewertungskriterien für Hörverstehensunterricht?

1 In diesem Beitrag werden die befragten Englisch-Fachleitungen als Expert:innen für Unterrichtsqualität bezeichnet. Für eine ausführliche Begründung dieser Entscheidung, vgl. Abschnitt 3.1.

2 Theoretische Grundlagen

Zu Beginn des 21. Jahrhunderts wurde von Vandergrift angemerkt, dass das Hörverstehen einerseits das Kernelement des Fremdsprachenlernens sei, es sich dabei aber andererseits um die am wenigsten verstandene und am wenigsten erforschte Fertigkeit handle (Vandergrift, 2007, S. 191). Etwa 15 Jahre später stellt Aryadoust (2022, S. 69) fest: „In the past few decades, research has illuminated many issues in second language (L2) listening assessment. The L2 listening constructs have improved in coverage, clarity, and relevance and operationalization methods have gradually grown outside of the boundaries of traditional testing.“ In dieser Aussage zeigt sich, dass die Hörverstehensforschung in den letzten Jahren viele relevante Erkenntnisse gewinnen konnte. Allerdings spiegelt sich darin auch sehr genau der Schwerpunkt der Erkenntnisse wider: Während über die fremdsprachliche Hörverstehenskompetenz als solche zunächst wenig empirisch validierte Informationen vorlagen, hat sich die isolierte Überprüfung dieser fremdsprachlichen Kompetenz stark weiterentwickelt. Getrieben von dem Fokus auf die validitätsorientierte Testung von fremdsprachlichem Hörverstehen konnte im Zuge dessen auch das Verständnis über die zugrundeliegenden kognitiven Prozesse erweitert werden, welche die Grundlage für valide Inferenzen über die Leistung der Prüflinge bilden. Inwiefern diese Erkenntnisse darüber hinaus auch Einfluss auf die Art und Weise, wie Hörverstehenskompetenz im Unterricht angebahnt wird, gefunden haben, steht im Zentrum dieses Kapitels.

2.1 Hörverstehen – eine mündliche Kompetenz

Bis Mitte des 20. Jahrhunderts wurde Hörverstehen als passive Fähigkeit aufgefasst, die keine gezielte Betrachtung im Rahmen des Fremdsprachenunterrichts benötigt. Martínez-Flor und Usó-Juan (2006, S. 30) halten diesbezüglich fest: „Up to the end of the 1960s, the status of listening comprehension in language learning and teaching was one of neglect“. Obgleich in den späten 70er Jahren die Auffassung zunahm, dass Hörende während des Hörens sehr wohl aktiv sind, herrschte dennoch ein Verständnis von Hörverstehen vor, dass es sich dabei eher um eine einzelne Fertigkeit im Rahmen kommunikativer Kompetenz handle als um eine komplexe Kompetenz der Mündlichkeit. Dies verdeutlicht beispielsweise die veraltete und widerlegte Auffassung, dass allein die reine Konfrontation mit immer anderen Ausgangstexten dazu führe, dass sich die Hörverstehenskompetenz der Lernenden weiter verbessere. Es hat sich jedoch gezeigt, dass Hörverstehen eine komplexe, mündliche Kompetenz ist, die eine gezielte und umfassende unterrichtliche Anbahnung erfordert (z. B. Field, 2009; Burns & Siegel, 2018; Mertes, 2019; Goh & Vandergrift, 2022). Dies soll im Rahmen der folgenden Unterkapitel dargelegt werden.

2.1.1 Der Hörverstehensprozess

Über die kognitiven Aktivitäten, die während des Hörverstehensprozesses von den Hörenden ausgeführt werden, ist – unter anderem durch die Erkenntnisse der Kognitionspsychologie – bereits vieles bekannt (Rossa, 2012). Insgesamt lassen sich vier Ebenen von Verarbeitungsprozessen voneinander unterscheiden (Rost, 2016): neurologische, linguistische, semantische und pragmatische. Es ist dabei hervorzuheben, dass alle vier Ebenen miteinander interagieren und im Verlauf des Hörverstehensprozesses aufeinander einwirken.

Die neurologischen Verarbeitungsprozesse betreffen die physische Verarbeitung von Schallwellen durch das Gehirn. Hierbei sind unterschiedliche Grade von Bewusstsein über das Hören und die Aufmerksamkeit der Hörenden entscheidende Elemente. Zu beachten ist die Unterscheidung von *hearing* und *listening*, die deutlich macht, dass bloßes Hören nicht ausreicht, um weitere kognitive Verarbeitungsprozesse anzuregen. Es benötigt immer auch die Intention der Hörenden, eine mehr oder weniger bewusste Entscheidung, sich auf das Gehörte einzulassen (Kriegel, 2013; Rost, 2016).

Die linguistischen Verarbeitungsprozesse, die sich insbesondere auf die Verarbeitungsebenen *perception* und *parsing* (z. B. Goh, 2000) beziehen und auch als *lower-level processes* bezeichnet werden, basieren auf der Annahme, dass Hörende Sprachdaten in der phonologischen Schleife des Gedächtnisses aufzeichnen (Anderson & Lynch, 1988, S. 9) und retrospektiv dekodieren. Sie nutzen ihr linguistisches Wissen, das sowohl phonetisches und prosodisches als auch Wort-, grammatikalisches und Syntaxwissen umfasst, um die in der gesprochenen Sprache verschlüsselten Informationen zu entnehmen. Gerade beim fremdsprachlichen Hörverstehen sind die Sprachdatenverarbeitungsprozesse sehr fehleranfällig, da sie auf linguistischem Wissen basieren, das gerade bei unerfahrenen Lernenden noch sehr limitiert ist (Field, 2009; Rossa, 2012). Erschwerend kommt auch der Unterschied der Sprachen und deren bedeutungstragender Phoneme hinzu, da die Differenzierung dieser Phoneme für die Lernenden, insbesondere ältere Lernende, eine große Herausforderung darstellt (Rossa, 2012).

Die semantischen Verarbeitungsprozesse, *higher-level processes*, beziehen sich insbesondere auf die Verarbeitungsebene *utilization* und umfassen die Bereiche Verstehen, Inferieren und Lernen (Rost, 2016). Dabei interagieren die semantischen Verarbeitungsprozesse mit dem Gedächtnis (auf den Ebenen der phonologischen Schleife, dem Arbeitsgedächtnis und dem Langzeitgedächtnis). Rossa und Meißner (2017) bezeichnen die semantische Verarbeitung folgerichtig als „konzeptgeleitete [...], vom Welt- und Erfahrungswissen ausgehend [...] (*top-down*) Verarbeitung des Gehörten“ (S. 85, kursive Hervorhebungen im Original), bei der es für die Hörer:innen darauf ankomme, die verarbeiteten Elemente im Rahmen des Diskurses einzuordnen, an dem sie gerade teilnehmen (Rossa, 2012, S. 12; Field, 2009, S. 209–214). Verstehen meint hier, dass Bedeutung auf der Grundlage konstruiert wird, dass neue Informationen und bereits Bekanntes zueinander in Bezug gebracht werden. Dieser Prozess ist auch als interpretatorischer Akt

der Hörenden aufzufassen, der es erfordert, dass verschiedene logische Schlussfolgerungen auf Basis des Verstandenen und des vorhandenen Wissens gezogen werden (z. B. Goh, 2000). Da die Wissensbestände, die zur Interpretation des Gesagten herangezogen werden, höchst individuell sind, hat Douglas' (1997, S. 22) Aussage „No two listeners hear the same message“ weiter Bestand.

Die pragmatische Ebene der Verarbeitungsprozesse bezieht sich auf das Erschließen der Intentionen von Sprecher:innen und das Ziehen von Inferenzen auf der Grundlage sozialer und kommunikativer Konventionen. Es ist in diesem Kontext nützlich, einzuordnen, welche Rolle die Hörenden einnehmen. Diese Rolle kann innerhalb des Kontinuums ‚mitwirkend – nicht mitwirkend‘ wechseln. Rost (2016, S. 68) beschreibt fünf verschiedene Ebenen innerhalb des Kontinuums: *Teilnehmer:in*, *Adressat:in*, *Hörer:in*, *Overhearer* und *Richter:in*. Auffallend ist, dass im Rahmen des Hörverstehensunterrichts häufig ein Fokus auf die beiden letztgenannten Rollen gelegt wird, die jedoch nicht Teil des Diskurses sind, weshalb eine genauere Betrachtung dieses Kontinuums (von interaktional zu transaktional) für das Verständnis des Hörverstehensunterrichts sinnvoll erscheint.

2.1.2 Transaktionales und interaktionales Hörverstehen

Eine Begründung für die Vernachlässigung der Berücksichtigung von diversen anderen pragmatischen Rollen von Hörer:innen im Kontext des fremdsprachlichen Unterrichts könnte der Fokus auf den Einsatz von transaktionalem Hörverstehen sein. Beim transaktionalen Hören findet der Informationstransfer nur in eine Richtung statt: vom Text zur hörenden Person. Der Schwerpunkt dieses Modus des Hörens liegt somit eindeutig auf der Informationsentnahme und eignet sich vor allem für eine isolierte Testung des Hörverstehens (Folkerts, eingereicht). Durch die vielerorts gängige Prüfungspraxis des Hörverstehens haben auch viele damit verbundene Prinzipien als Folge Einfluss auf die Unterrichtsgestaltung erhalten. Eines dieser Prinzipien ist, dass ein Hörtext zunächst im Hinblick auf ein Allgemeinverständnis vorgespielt wird und dann ein zweites Mal zur Entnahme von spezifischen Informationen (*Global-to-Detail-Approach*). Um feststellen zu können, ob die Lernenden die richtigen Informationen entnommen haben, werden insbesondere geschlossene und halb-offene Aufgabenformate eingesetzt (Field, 2009; Field, 2019). Ein Unterricht, der allein auf diese Form der Überprüfung des Hörverstehens abzielt (Field bezeichnet dies als den *Comprehension Approach*; Field, 2009, S. 26–34), trägt jedoch nur wenig zum tatsächlichen Erlernen dieser komplexen Kompetenz bei, da nur ein limitierter Bereich des Hörverstehensprozesses in den Blick genommen wird. Field stellt dem *Comprehension Approach* einen *Process Approach* entgegen, der gezielt darauf ausgerichtet ist, insbesondere die Lücken der fremdsprachlichen Lernenden im Verarbeitungsprozess der Sprachdaten in den Blick zu nehmen und zu trainieren. Dies kann z. B. mit sogenannten *micro listening* Aufgaben erfolgen, deren Wirksamkeit bereits in internationalen Studien festgestellt wurde (z. B. Rezaei & Hashim, 2013). In

Deutschland sind erst in den vergangenen Jahren erste Praxisartikel zur Möglichkeit des Einsatzes von *micro listening* erschienen (z. B. Rossa & Helsper, 2021; Rossa & Matz, 2023; Folkerts, 2023b), weshalb anzunehmen ist, dass diese Form des Hörverstehensunterrichts (noch) seltene Praxis ist. Darüber hinaus ist auch der Einsatz von metakognitiven Strategien für das Hörverstehen (z. B. Goh & Vandergrift, 2022) als besonders lernwirksam hervorzuheben. Wie von Goh (2000) gezeigt, bietet es sich zur Einübung verschiedener metakognitiver Strategien für das Hörverstehen an, diese den einzelnen Prozessebenen (*perception*, *parsing* und *utilization*) zuzuordnen, um gezielt entweder häufige oder individuelle Schwierigkeiten mit dem Hören anzugehen.

Obgleich der transaktionale Hörverstehensunterricht sicherlich die am häufigsten auftretende Form zu sein scheint, ist es dennoch lohnend, einen Exkurs zur Bedeutung des interaktionalen Hörens zu machen. In der Regel wird im kommunikativen Englischunterricht vorausgesetzt, dass die Lernenden dem *Classroom Discourse* folgen können und dass die kommunikative Bedeutungsaushandlung in Gruppen- oder Partnerarbeitsphasen funktioniert – an den Stellen, an denen dies nicht der Fall wäre, müsste es ja auffallen. Insofern erscheint es auf den ersten Blick nicht erforderlich, interaktionales Hörverstehen spezifisch in den Blick zu nehmen (obwohl dies durchaus im Rahmen von Kernlehrplänen gefordert wird). Zu einer ganzheitlichen Ausbildung der Hörverstehenskompetenz sollte dieser Modus des Hörverstehens jedoch stärker in den Fokus der Forschung rücken, da die kognitiven Herausforderungen in interaktionalen Situationen weit über transaktionales Hörverstehen hinausgehen. So ist beispielsweise die pragmatische Notwendigkeit, innerhalb des Diskurses angemessen auf das Gesagte zu reagieren, eine Herausforderung, auf die die Lernenden gezielt vorbereitet werden sollten. Auch wenn im Rahmen dieses Beitrags nicht näher auf diesen Bereich eingegangen werden kann, so sei dennoch ein Verweis auf dessen Bedeutung erlaubt.

2.2 Konzeptionalisierung des Hörverstehensunterrichts

Im Folgenden wird auf das transaktionale Hörverstehen fokussiert, weshalb zunächst die Elemente des Hörverstehensunterrichts skizziert werden, die auf die Ausbildung dieses Modus des Hörverstehens abzielen (vgl. Abb. 1). Für einen Unterricht, der interaktionales Hörverstehen fördern soll, wäre eine andere Konzeption notwendig (vgl. für Grundlagen einer solchen Konzeption Goh & Vandergrift, 2022, S. 37–56; aber auch Folkerts & Matz, 2024). Es ist anzunehmen, dass sich die spezifische Ausgestaltung der einzelnen Elemente auf die Qualität des transaktionalen Hörverstehensunterrichts auswirkt. Insofern dient die in Abbildung 1 entworfene Konzeption der Elemente des Hörverstehensunterrichts als theoretische Grundlage, die für die Auswertung der Interviews herangezogen wird. Die von den Expert:innen benannten Kriterien für guten Hörverstehensunterricht werden auf die hier skizzierten Elemente bezogen und dienen insofern der weiteren Ausdifferenzierung dieser Konzeption.

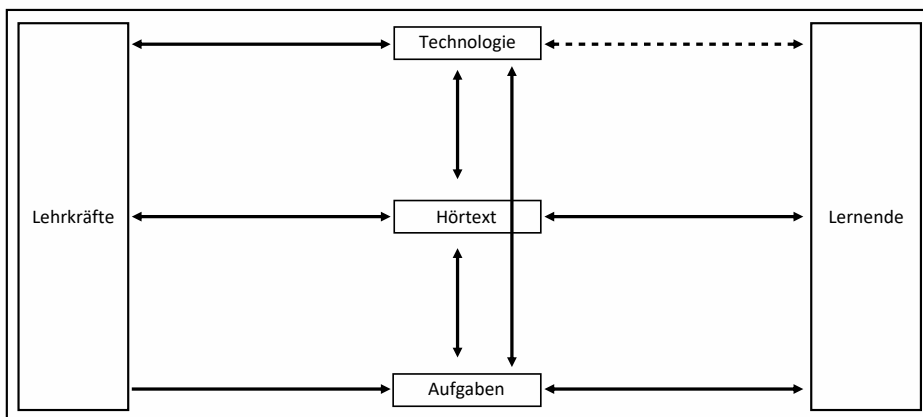


Abbildung 1: Konzeptionalisierung der Elemente des Hörverstehensunterrichts

Wie aus Abbildung 1 ersichtlich wird, steht der Hörtext im Zentrum des transaktionalen Hörverstehensunterrichts², verfügt über spezifische Eigenschaften und wird sowohl von den Lehrkräften als auch von den Lernenden, meist zu unterschiedlichen Zeitpunkten, rezipiert. Neben dem inhaltlichen Gehalt des Ausgangstextes, der nicht zu vernachlässigen ist, können außerdem die Aspekte in Tabelle 1 eine Rolle spielen:

Tabelle 1: Aspekte eines Hörverstehenstexts (vgl. Field 2019, S. 35–70)

Eigenschaften	Sprachliche Gestaltung
Anzahl der Sprecher:innen	Authentizität/natürliche gesprochene Sprache
Länge des Textes	<i>Native Speaker, Varieties</i> etc.
Hintergrundgeräusche	Vokabular
Adressat:innenorientierung	Grammatikalische Elemente

Ein weiteres Element des Hörverstehensunterrichts, das in Bezug zu dem Hörtext wie auch zu den Lernenden steht, sind die zum jeweiligen Hörtext passenden Aufgaben. Die Zielsetzungen der Aufgaben können stark divergieren und wirken sich letztendlich auch auf die Unterrichtsgestaltung aus. Während z. B. Aufgaben,

2 Im Sinne von Siegels (2018, S. 198–199) „one-way listening“ basiert der hier verwendete Begriff des transaktionalen Hörverstehensunterrichts vor allem auf der Idee, dass die Hörenden nicht mehr Teil des ursprünglichen Bedeutungsaushandlungsprozesses des Hörtextes sind. Beispielsweise kann der Inhalt einer im Unterricht gehörten Podcastepisode nicht mehr von den Lernenden nachträglich verändert werden, da sie mit den Urhebern des Textes nicht direkt in Interaktion treten können. Die Transaktion der Inhalte des Podcasts gehen dementsprechend nur in eine Richtung (von den Hörtextproduzenten zu den Hörenden), während in interaktionalen Bedeutungsaushandlungsprozessen beide Transaktionsrichtungen möglich sind, da die Hörenden eigene Hörtextproduzenten werden können. Vgl. auch die Ausführungen in Abschnitt 2.1.2.

die Elementen des *Comprehension Approaches* folgen, stärker darauf ausgerichtet sind, festzustellen, ob die Lernenden globale oder spezifische Informationen richtig verstanden haben, zielen Aufgaben des *Process Approaches* eher darauf ab, zu diagnostizieren, an welchen Stellen Verständnisschwierigkeiten vorliegen (vgl. Field, 2003, S. 326).

Auch die Technologie ist als ein Element für den transaktionalen Hörverstehensunterricht unerlässlich und hängt mit den anderen bisher benannten Elementen zusammen. Zunächst muss der Hörtext den Lernenden zugänglich gemacht werden. Dabei hat sich in der Praxis die Vorgehensweise etabliert, den Hörverstehenstext zentral durch die Lehrkraft vorzuspielen – häufig auch mehrmals. Da die Lernenden in diesem Szenario nur begrenzten Einfluss auf das Element der Technologie haben, ist es in Abbildung 1 gestrichelt dargestellt. Der zunehmende Einfluss der Digitalität führt dazu, dass gerade dieses Element des Hörverstehensunterrichts derzeit stark im Wandel ist. Damit gehen verschiedene Möglichkeiten (z.B. die individuelle Steuerung des Hörtextes durch die Lernenden), aber auch Herausforderungen einher (z.B. eine funktionale Aufbereitung des Hörtextes durch die Lehrkraft).

3 Studiendesign

Die vorliegende Studie und ihre Ergebnisse sind Teil eines übergeordneten designbasierten Forschungsprojekts, welches das Ziel verfolgt, ein umfassendes Kriterienraster für kompetenzorientierten Hörverstehensunterricht zu entwickeln. Es handelt sich um eine empirische Studie, bei der verschiedene Perspektiven von identifizierten Stakeholdern (Schüler:innen und Englischlehrer:innen) miteinander trianguliert werden. Als Produkt wird in einem kollaborativen Ansatz mit Lehrkräften einer Kooperationsschule eine Unterrichtsreihe für die Jahrgangsstufe 9 zur Förderung von Hörverstehen entwickelt. Die Unterrichtsreihe basiert auf den zuvor heuristisch entwickelten und theoriebasierten, vorläufigen Designprinzipien. Im weiteren Verlauf des übergeordneten Projekts werden diese in iterativen Schleifen verfeinert und sollen abschließend in ein Kriterienraster überführt werden. Als zusätzlicher Validierungsprozess wird dieses aber auch im Rahmen einer Kooperation mit Fachleiter:innen von zwei Bezirksregierungen in NRW (im Folgenden bezeichnet als BezRegA und BezRegB) kritisch diskutiert. Die in der vorliegenden Studie befragten Fachleitungen sind Teil dieser Kooperation, die jedoch eine größere Gruppe von Englischfachleiter:innen umfasst. Es handelt sich bei dem im Folgenden dargestellten Ausschnitt des Forschungsprojekts um den explorativen Teil der Vorstudie, bei der die Einschätzungen der Fachleitungen zu Kriterien guten Hörverstehensunterrichts in leitfadengestützten Interviews erhoben wurden.

3.1 Auswahl der Interviewteiler:innen

Die Interviewteiler:innen sind Fachleitungen für das Fach Englisch an Gymnasien und Gesamtschulen der BezRegA ($n = 5$) und BezRegB ($n = 2$). Von den sieben Fachleitungen sind sechs weiblich, eine männlich. Sie weisen insgesamt eine breite Erfahrungsspannweite auf. Zwei Fachleiterinnen waren erst seit kurzem im Amt, hatten aber bereits mehrjährige Erfahrung in der Unterrichtsbewertung im Rahmen ihrer Funktion als Ausbildungsbeauftragte ihrer jeweiligen Schulen. Drei Fachleitungen waren seit etwa zehn Jahren in dieser Funktion, zwei weitere sind seit mehr als 15 Jahren als Englisch-Fachleitungen tätig. Darüber hinaus sind vier auch als Fachberater:innen der Fachdezernate Englisch tätig. Drei der Interviewten sind seit mehreren Jahren in den Kompetenzteams der jeweiligen Bezirksregierungen und geben regelmäßig Fortbildungen für Englischlehrkräfte. Der Kontakt zu den Fachleitungen ist unter anderem über zwei Workshops entstanden, die von dem Autor in Dienstbesprechungen der Fachleitungen für Englisch an Gymnasien und Gesamtschulen durchgeführt wurden. Zum Zeitpunkt der Interviews ist anzumerken, dass drei der interviewten Teiler:innen aus BezRegA bereits an einem Workshop des Autors teilgenommen hatten, in dem ein Fokus auf interaktionales und extensives Hörverstehen gelegt wurde. Die Dienstbesprechung, in deren Rahmen der Workshop angeboten wurde, legte einen Fokus auf das Hörverstehen, sodass alle fünf Fachleitungen sich mit neueren wissenschaftlichen Erkenntnissen zum Thema auseinandergesetzt haben. Einzelne Aussagen der Interviewten replizieren ggf. Inhalte der Dienstbesprechung, auch wenn diese mehrere Wochen zuvor stattgefunden hatte. Hierauf wurde bei der Auswertung der Daten Rücksicht genommen. Die Fachleitungen aus BezRegB hatten zum Zeitpunkt des Interviews noch keinen Workshop oder eine dienstliche Fortbildung zu Hörverstehen besucht, weshalb die Aussagen dieser Teilerinnen nicht beeinflusst wurden.

Die Bezeichnung von Interviewteiler:innen als Expert:innen ist nicht ganz unproblematisch, da behauptet werden könnte, dass „jeder Mensch Experte sein [könnte]“ (Misoch, 2019, S. 120). Folglich ist es notwendig, enger abzugrenzen und zu begründen, weshalb es sich nach Ansicht des Autors bei den interviewten Fachleitungen um Expert:innen handelt. Als ein entscheidendes Kriterium für Expert:innen nennt Riemer (2022, in Anlehnung an Meuser & Nagel, 2009), dass es sich bei den befragten Personen um Menschen handelt, „die für den gewählten Untersuchungsgegenstand als besonders kompetent gelten und soziale Repräsentanten des entsprechenden Handlungsfelds mit Gestaltungs- und Entscheidungsfunktionen sind“ (S. 174). Die Interviewteiler:innen haben im Rahmen der Übernahme der Fachleitungsfunktion in einem Prüfverfahren nachgewiesen, dass sie Unterrichtsqualität aus der Praxisperspektive heraus fundiert beurteilen können und gelten dementsprechend innerhalb des Schulsystems als kompetent für den untersuchten Gegenstand. Darüber hinaus trifft auch das zweite von Riemer (2022) genannte Kriterium auf die befragten Fachleiter:innen zu, da sie bei der Bewertung der Qualität der beobachteten Unterrichtsstunden

einen deutlichen Gestaltungs- und Entscheidungsspielraum haben und nicht nur ausführende Organe sind.

Außerdem nennt Misoch (2019, S. 119–121) weitere Kriterien für die Bestimmung von Expert:innen. Zunächst verweist sie auf das Kriterium, dass Expert:innen über ein *Sonderwissen* (Hitzler, 1994; Misoch, 2019) verfügen. Dieses wurde „zumeist durch lange (Aus)Bildungswege erworben, die oftmals mit der Verleihung von Zertifikaten und anderen Formen der objektiven Kompetenzbescheinigung einhergehen“ (Misoch, 2019, S. 119). Beides trifft auf die befragten Fachleitungen zu, da sie nach dem Studium sowohl vor als auch nach Übernahme der Fachleitungsfunktion mehrere Fort- und Weiterbildungen besucht haben, die sie auf die Fachleistungstätigkeit vorbereitet und während der Ausübung dieser Funktion unterstützt haben. Darüber hinaus geht die Übernahme der Fachleitungsfunktion auch mit einer Verleihung einer Urkunde einher, die als objektive Form der Kompetenzbeschreibung anerkannt ist. Misoch (2019) argumentiert weiterhin, dass das Sonderwissen von Expert:innen auch „durch spezielle Tätigkeiten/Funktionen z. B. innerhalb von Organisationen“ (S. 121) entsteht. Innerhalb der Organisation ‚Schulsystem‘ nehmen Fachleitungen eine bedeutende, systembildende und -wahrende Funktion ein, die durch die spezifische Tätigkeitsbeschreibung dafür sorgt, dass Fachleitungen eben jenes Sonderwissen erwerben, das sie von vielen anderen Lehrkräften unterscheidet. Die Fachleitungen bewerten und beurteilen bei den beobachteten Unterrichtsstunden, basierend auf dem von ihnen erworbenen Sonderwissen, ob es sich anhand der von ihnen gesetzten Kriterien sowohl um guten als auch effektiven Unterricht handelt. Dabei muss bedacht werden, dass die beobachteten Stunden nicht immer auch „realistischem“ Unterricht entsprechen (siehe hierzu den Verweis von Lindl et al., in diesem Band, zu „Showstunden“). Ferner wird den Fachleitungen auferlegt, sich regelmäßig fortzubilden³, um so systemisch sicherzustellen, dass die von den Fachleitungen angelegten Kriterien idealerweise dem aktuellen Stand der Wissenschaft entsprechen und nicht allein aus subjektiven Theorien gehandelt wird.

Außerdem ist hervorzuheben, dass sich Expert:inneninterviews in Form und Zielsetzung von anderen Interviewtypen unterscheiden. Die Setzung des Interviewers, dass es sich bei den befragten Personen um Expert:innen handelt, wirkte sich auf die Art der Interviewführung aus, insbesondere auf das Verständnis der Rollenbeschreibung bzw. des Status von interviewter Person und Interviewer (Misoch, 2019, S. 121–123). Ferner liegt das Erkenntnisinteresse des Interviews explizit auf der Erfassung des Sonderwissens und nicht z. B. auf biografischen Erhebungen. Die befragten Fachleitungen wurden im Sinne von Liebold & Trinczek (2009) als „Funktionseliten innerhalb eines organisatorischen und institutionellen

3 Vgl. hierzu z. B., die Aufgabenbeschreibung von Fachleitungen der Bezirksregierung Düsseldorf, aus der hervorgeht, dass Fachleitungen fähig sein sollen, „Lerngegenstände und Vermittlungsprozesse am aktuellen Stand der Bildungswissenschaften und Fachwissenschaften[,] deren Didaktiken [und] den Ergebnissen der Unterrichts- und Bildungsforschung auszurichten, um reflexiv gesteuerte Handlungskompetenzen zu entwickeln“ (Bezirksregierung Düsseldorf, 2014, S. 6).

Kontextes“ (S. 34) aufgefasst, weshalb die von ihnen vertretenen „institutionalisierte[n] Wissensbestände und Problemlösungen“ (Misoch, 2019, S. 120) im Vordergrund der Studie stehen.

3.2 Erhebung und Aufbereitung der Daten

Die vorliegende Studie verwendet ein qualitatives Forschungsdesign. Gegenstand des Untersuchungsinteresses sind die Auffassungen und Bewertungskriterien der Fachleitungen zum Thema ‚Was ist guter Hörverstehensunterricht?‘, die hier mit einem Fokus auf das transaktionale Hörverstehen dargestellt werden. Die Erhebung, Aufbereitung und Auswertung der Daten basiert auf verschiedenen methodologischen Entscheidungen, die zur Erhöhung der intersubjektiven Nachvollziehbarkeit im Folgenden dargestellt werden.

3.2.1 Leitfadententwicklung und Pilotierung

Bei der Erstellung des Leitfadens für das Interview wurde nach der SPSS-Methode vorgegangen (Helfferich, 2011; Helfferich, 2019, S. 676–678; SPSS steht für Sammeln, Prüfen, Sortieren und Subsummieren der Fragen). Zunächst sammelte der Autor im Rahmen eines offenen Brainstormings alle Fragen zum gesamten Bereich des Untersuchungsinteresses, die anschließend auf ihre Eignung für die Beantwortung der Forschungsfrage geprüft wurden. Nach einer eingehenden Sortierung wurden die Fragen in eine Tabelle subsumiert, welche die Spalten a) Überkategorie, b) Leitfrage/Stimulus/Erzählaufforderung, c) (Nach-)Fragen mit obligatorischer Formulierung und d) inhaltliche Aspekte (ggf. Nachfragen bei fehlender Benennung) umfasst.

Um die Praktikabilität des Leitfadens zu gewährleisten, wurde dieser zweimal mit zwei Lehrkräften, die als Fachberaterin bzw. Fachleiter tätig und nicht Teil der Stichprobe sind, pilotiert. In beiden Pilotierungen konnte die Praktikabilität des Leitfadens sichergestellt werden, weshalb keine wesentlichen Änderungen notwendig waren; lediglich einzelne Erzählaufforderungen wurden sprachlich angepasst. Der Aufbau des Leitfadens gliedert sich in folgende Überkategorien

- a) Einstieg/Allgemeine Einstellungen (Erzählimpuls, z.B. „Was hat Sie dazu bewogen, Fachleiter:in für Englisch zu werden?“)
- b) Abgrenzung spezifischer Kriterien für die Qualität von HV-Stunden (Erzählimpuls, z.B. „Was sind typische Herausforderungen von Hörverstehensstunden?“)
- c) Stundenplanung (Erzählimpuls, z.B. „Sie lesen bei einer UPP den Stundenentwurf zu einer Hörverstehensstunde. Woran erkennen Sie dabei, dass es sich um eine gelungene Stunde handeln könnte?“)
- d) Hörverstehensstunden (Spezifika von HV-Stunden)/Ausgewählte Aspekte von HV-Stunden (Ausgangstext, Aufgabenstellung, HV-Prozess) (Erzählimpuls,

z. B. „Nach welchen Kriterien bewerten Sie die Textauswahl der Referendar:innen in Hörverstehensstunden?“; „Woran machen Sie am Ende einer beobachteten Hörverstehensstunde fest, dass ein tatsächlicher Lernzuwachs im Hinblick auf das Hörverstehen stattgefunden hat?“)

- e) Aufgreifen weiterer themenrelevanter Aspekte (Erzählimpuls, z. B. „Welche Implikationen haben die Digitalisierung des Fremdsprachenunterrichts und die hohe Verfügbarkeit von authentischem Material für eine gute Hörverstehensstunde?“)

Zum Abschluss des Interviews wurde den Befragten mit einer vollständig offenen Frage die Möglichkeit gegeben, Dinge im Hinblick auf Kriterien einer guten Hörverstehensstunde anzusprechen, die nicht durch die Befragung erfasst wurden.

3.2.2 Durchführung der Interviews und Aufbereitung der Daten

Der Autor überließ den Interviewten die Auswahl des Settings. Das Interview fand bei der befragten Person zuhause statt, drei Interviews digital über Zoom, zwei in der Schule der Interviewten und eines an einem ruhigen öffentlichen Ort. Bei allen Interviews erfolgte eine zweifache Aufzeichnung der Tonspur; Videodaten wurden nicht erfasst. Nach Beendigung der Interviews wurde die Aufnahme mit der jeweils besten Audioqualität für die Transkription verwendet, die andere direkt gelöscht.

Um der von Misoch (2019) vorgebrachten Kritik an vielen Expert:inneninterviews, die auf einer selektiven Transkription beruhen, zu entgehen, wurde ein vollständiges Transkript mittels Standardorthografie erstellt. Dies wurde mit der Software *f4* umgesetzt. Hierzu wurde zunächst ein vorläufiges Transkript durch die automatisierte Spracherkennung angefertigt, das in der Folge vom Autor und studentischen Hilfskräften überarbeitet wurde. Bei allen Transkripten erfolgte eine abschließende Überprüfung durch den Autor sowie eine Sichtung von Auszügen (in etwa 15%) im Rahmen eines Forschungskolloquiums, um sicherzustellen, dass die Transkripte gängigen Gütekriterien (Misoch, 2019) entsprechen und mit den Tonaufnahmen übereinstimmen. Bei der Transkription wurde auf die Einhaltung gängiger Transkriptionsregeln (in Anlehnung an Kuckartz, 2016, S. 167, und Misoch, 2019, S. 274–275) geachtet.

3.3 Auswertung der Daten

Die Auswertung der Expert:inneninterviews orientierte sich an den Leitlinien der strukturierenden qualitativen Inhaltsanalyse nach Kuckartz (2016). Ein Kategoriensystem, das sich in verschiedene deduktive und induktive Kategorien unterteilt, wurde entwickelt. Zunächst wurden die Hauptkategorien festgelegt, die de-

duktiv durch die Einteilung des Leitfadens in Überkategorien vorgegeben waren (vgl. Tab. 2).

Im nächsten Schritt wurden die Textstellen gemäß den Hauptkategorien kodiert und strukturiert zusammengestellt, um anschließend Subkategorien zu bilden, die sich einerseits induktiv aus dem Datenmaterial ergaben, aber andererseits (vor allem in der Bezeichnung) auch die theoretischen Vorannahmen des Autors widerspiegeln (vgl. Tab. 2).

Tabelle 2: Kategoriensystem für die Auswertung der Interviews mit den Fachleiter:innen

Hauptkategorie 1: Allgemeine Aussagen zu gutem Englischunterricht	Hauptkategorie 2: Spezifika von Hörverstehensunterricht	Hauptkategorie 3: Bewertungskriterien für Hörverstehensunterricht
Kommunikative Aktivierung	Hörtexte	Kognitive Aktivierung
Schüler:innenzentrierung	Aufgabenstellungen	Diagnostik
Lerngegenstände	Technologie	Gegenstandsanalyse und Inhalte
Kompetenzentwicklung		Aktive Beteiligung
		Stundenplanung
		Erlernbarkeit von Hörverstehen

Das fertige Kategoriensystem wurde zwei Peers (d. h. Post-Docs der Fachdidaktik Englisch) vorgestellt, kritisch diskutiert und geringfügig angepasst. Zwei Subkategorien wurden verworfen, da sie nicht präzise genug für die Fragestellung waren. Zuletzt wurden eine ausdifferenzierte Kodierung des gesamten Textkorpus durchgeführt sowie eine thematische Zusammenfassung angefertigt.

4 Darstellung der Ergebnisse

Im Folgenden werden zunächst allgemeine (fremdsprachendidaktische) Aspekte von und Bewertungskriterien für Unterrichtsqualität (in Form von ‚was zeichnet guten Englischunterricht aus‘) dargestellt, um im Anschluss daran spezifische Kriterien für fremdsprachlichen Hörverstehensunterricht zu identifizieren. Bei den thematisch orientierten Zusammenfassungen ist deutlich geworden, dass eine Aufteilung in allgemeine Aussagen zu gutem Unterricht und zu expliziten Bewertungskriterien nicht immer trennscharf möglich ist. Dies spiegelt sich auch in Aussagen einiger Fachleitungen wider, die im Anschluss an die Interviews darauf hingewiesen haben, dass es für sie gut gewesen sei, explizit über ihre Bewertungskriterien nachdenken und sprechen zu müssen, da dies im Alltag typischerweise nicht vorkomme.

4.1 Allgemeine Aussagen und explizite Bewertungskriterien von Expert:innen im Hinblick auf guten Englischunterricht

Im ersten Teil des Interviews, in dem es darum ging, Referendar:innen ein Rezept mit Grundzutaten für eine gelungene Englischstunde an die Hand zu geben, wurde ersichtlich, dass nahezu alle Beteiligten eine solche Verallgemeinerung für guten Unterricht ablehnten oder sich zumindest damit unwohl fühlten. Exemplarisch steht hierfür die Aussage von Fachleitung E: „bei der Frage sträubt sich bei mir einiges, weil die Referendare und Referendarinnen hätten ja immer gerne das Rezept und zwar das mit Erfolgsgarantie. Das gibt es nicht.“ (FL E, 17).⁴ Trotz dieses ersten inneren Widerstrebens antworteten alle Fachleitungen dennoch ausführlich auf diese Frage. Die Antworten weisen mehrere Übereinstimmungen auf, von denen eine hervorsteht (Subkategorie Schüler:innenzentrierung). Für alle Fachleitungen zeichnet sich guter Englischunterricht dadurch aus, dass er vor allem auf die Lerngruppe abgestimmt ist. Auf den Punkt bringt dies Fachleitung E:

„Ähm, zwei Dinge, wenn ich jetzt mal so, also ganz grob, ähm, die Schülerinnen und Schüler, die Schülerinnen und Schüler, Schülerinnen und Schüler, Nummer 1, Nummer 2, Nummer 3, so das sind schon drei Sachen, nein, eigentlich die Lerngruppe. Die Lerngruppe angucken, ein Gefühl für die Lerngruppen bekommen, ähm, auf verschiedenen Ebenen, ähm, wie ticken die so? Ähm, aber auch, was können die? Was brauchen die? Wie kann ich die jetzt weiterbringen?“ (FL E, 17)

Darüber hinaus war für eine Mehrzahl der Fachleitungen (fünf von sieben) wichtig, dass der Unterricht vom Ende her gedacht wird bzw. eine Zielbestimmung erfolgt (Subkategorie Kompetenzentwicklung). Die Fachleitungen legen Wert darauf, dass sich die Referendar:innen im Hinblick auf den Kompetenzzuwachs der Lernenden verschiedene Gedanken machen. Leitfragen dabei sind z. B. „was möchte ich denn erreichen? Wozu möchte ich die Schülerinnen und Schüler begeistern? Und was sollen sie fachlich gelernt haben? Was sollen sie an Kompetenzen hinterher zeigen können? Und dahingehend dann sozusagen zu überlegen, welche Materialien, welche Fragestellungen bringe ich mit? Und wie eröffne ich den Unterricht?“ (FL B, 17). Weitere Überschneidungen ergaben sich bei den Aussagen im Bereich Lebensweltbezug. Dieser wird von den Interviewteilnehmer:innen an verschiedenen Stellen als sehr bedeutsam für guten Unterricht hervorgehoben (vgl. z. B. FL A, 25 & 28; FL B, 19 & 49; FL C, 21; FL E, 26 & 29; FL F, 137; FL G, 63). Für die Fachleitungen war es darüber hinaus in diesem Zusammenhang wichtig, dass das ausgewählte Material für die Schüler:innen sowohl ansprechend als auch authentisch ist. Es ist hierbei anzumerken, dass die Fachleitungen in diesem Zusammenhang authentisches Material mit nicht-didaktisiertem Material gleichsetzen, welches einen höheren Mehrwert für die Unterrichtsgestaltung aufweist (z. B. FL E, 21 & 61). Neben diesen Überschneidungen gab es aber auch

4 Im Folgenden werden Verweise auf die Interviewdaten immer angegeben mit: Zitierte Fachleitung, Absatz im Transkript.

einige individuelle Schwerpunktsetzungen der Fachleitungen, beispielsweise „Einsprachigkeit, und zwar in allen Schulformen, sprachliche Förderung, gezielte, in allen Stunden passgenaue, angemessene, situationsangemessene, kompetente, sensible Fehlerkorrektur, aber nicht keine Fehlerkorrektur“ (FL F, 15; ähnliche Ergebnisse bei Hofrichter et al., in diesem Band) oder „deutlich stärker Vokabular fokussieren“ (FL A, 13).

Insbesondere bei der Bewertung eines hypothetischen Unterrichtsentwurfs zum Hörverstehen, der von Referendar:innen im Rahmen von unterrichtspraktischen Prüfungen vorgelegt werden muss, verwiesen die Fachleitungen auf allgemeine Bewertungskriterien für eine gelungene Stunde. Hierbei achten die Fachleitungen (über die bereits zuvor genannten Aspekte hinaus) darauf, dass folgende Elemente im Rahmen des Entwurfs besprochen werden, von denen dann auch zu erwarten ist, dass sie in gelungener Form im Rahmen des Unterrichts umgesetzt werden:

- Sachanalyse des Unterrichtsgegenstands (z. B. FL D, 35; FL E, 49)
- Diagnostizierung der Lerngruppe (z. B. FL D, 35)
- Auf die Lerngruppe abgestimmte Differenzierung (z. B. FL C, 35)
- Überprüfbarkeit des Lernzuwachses (z. B. FL F, 64; FL G, 49)
- Strukturierung der Unterrichtsstunde (z. B. FL B, 37)
- Einbettung der Unterrichtsstunde in die Unterrichtsreihe (z. B. FL A, 33; FL E, 46)

Die Fachleitungen waren sich insgesamt bewusst, dass sie in vielen Fällen Showstunden sehen (z. B. FL F, 119), die nicht immer unbedingt mit der Unterrichtsrealität übereinstimmen. Ein Diskurs, der diesbezüglich von mehreren Fachleitungen geführt wurde, drehte sich um die Frage, ob denn zwangsläufig alle drei Anforderungsbereiche (AFB I-III⁵) im Rahmen der Unterrichtsbesuche abgedeckt werden müssten, damit es auch gute Stunden seien. Hierbei gab es leichte Unterschiede in den Auffassungen. Konsens herrschte aber insofern, als dass sich die Fachleitungen dafür aussprachen, dass die Voraussetzung für eine gute Stunde nicht in der Abdeckung aller drei AFBs in einer 45- oder 60-minütigen Stunde liegen könne (z. B. FL A, 58; FL C, 59; FL E, 73; FL F, 111). Demnach solle zwar ein Schwerpunkt auf dem AFB III liegen (z. B. FL B, 67), aber um diesen innerhalb der begrenzten Zeit eines Unterrichtsbesuchs zu erreichen, sei nicht immer zwingend eine Abdeckung von AFB I (oder II) notwendig (z. B. FL A, 58).

Ein anderer Aspekt, der von den Fachleitungen angesprochen wurde, war die Setzung eines Kompetenzschwerpunkts für die Unterrichtsstunde. Obwohl es

5 Die Unterscheidung der drei Anforderungsbereiche ist nicht immer ganz trennscharf möglich, da es oft fließende Übergänge gibt. Ein Modell, das sich nach Ansicht des Autors zur Beschreibung der Anforderungsbereiche eignet, ist das *Model of Learning Objectives*, das auf die überarbeitete Version der Lernzieltaxonomie von Bloom zurückgeht: Anderson & Krathwohl (2001). AFB I entspricht hierbei den *lower order thinking skills* ‚remember‘ und ‚understand‘; AFB II entspricht ‚apply‘ und ‚analyze‘; AFB III spiegelt sich in den *higher order thinking skills* ‚evaluate‘ und ‚create‘ wider.

selbstverständlich sei, dass immer mehrere Kompetenzen zum Tragen kämen, zeichne sich eine gelungene Stunde dadurch aus, dass eine Kompetenz schwerpunktmäßig in den Blick genommen werde (z.B. FL F, 119; FL A, 68). Für den Schwerpunkt Hörverstehen sehen die Fachleitungen zwei Möglichkeiten, dies umzusetzen. Die eine wäre die inhaltliche Orientierung, sodass anhand der Art und Weise wie die Lernenden mit dem Verstandenen weiterarbeiten, erkannt werden kann, dass sie ihr Hörverstehen zielgerichtet eingesetzt haben. Eine andere Möglichkeit bestehe in der gezielten Fokussierung der Kompetenz als solcher und beispielsweise der daran anschließenden Reflexion der Lernenden über einen Kompetenzzuwachs (z.B. im Hinblick auf metakognitive Strategien):

„Wenn die Stunde eher diesen, ähm, etwas methodischeren Zugriff hatte, dann würde man das sozusagen, ähm, im Hinblick auf eine Reflexion mit den Schülerinnen und Schülern erkennen können. Welche Schwierigkeiten hattet ihr? Und was ist euch gut gelungen? Und warum?“ (FL B, 65).

4.2 Aussagen und explizite Bewertungskriterien von Expert:innen im Hinblick auf guten Hörverstehensunterricht

Der Schwerpunkt der Interviews lag auf der Isolierung von spezifischen Kriterien für die Bewertung von Englischunterricht, der auf die Förderung von Hörverstehen abzielt. Dabei ist den Expert:innen an einzelnen Stellen aufgefallen, dass sie genau solche isolierten, spezifischen Kriterien nur bedingt formulieren konnten, weil die von ihnen gewählten Kriterien auch auf die Gestaltung von Unterrichtsstunden mit anderen Kompetenzschwerpunkten übertragbar waren. Prototypisch hierfür steht folgende Aussage: „Momentan würde ich tatsächlich sagen, [...] Hörverstehen ist in vielerlei [Hinsicht] auch nicht anders als andere, ähm, Inter-, ähm, Dekodierung und Interaktion von, [und] mit Texten.“ (FL E, 104). Dennoch lassen sich Aspekte identifizieren, die sich von ihrer Grundannahme mit anderen Kompetenzbereichen decken, dennoch spezifisch für Hörverstehen sind.

Besonders gut lässt sich dies mit dem Bereich „Auswahl des Ausgangstextes“ illustrieren. Für nahezu alle Fachleitungen ist hier das übergeordnete Kriterium relevant, dass seitens der Lehrkraft eine Sachanalyse stattgefunden hat und dass der Text passend für die jeweilige Lerngruppe ist. Im Hinblick auf die konkrete Umsetzung dieser Passung treten dann jedoch die spezifischeren Bewertungskriterien zum Vorschein, die folgende Aspekte umfassen (alphabetische Auflistung):

- Akustische Verständlichkeit (auch im Hinblick auf *Varieties*, Sprecher:innenanzahl, Sprechtempo) (z.B. FL B, 49; FL D, 39; FL G, 67)
- Authentizität (z.B. FL A, 42; FL C, 51; FL E, 51; FL B, 49)
- Inhaltliche Relevanz (u. a. für die Lernenden, das thematische Dach, die curricularen Vorgaben) (z.B. FL A, 38; FL E 59; FL F, 98; FL G, 69)
- Länge des Ausgangstextes (z.B. FL C, 47; FL F, 40; FL G 67)
- Lebensweltbezug (z.B. FL D, 47; FL F 97)

- Mehrwert durch das Hören im Vergleich zum Lesen (FL A, 38)
- Vokabular (FL C, 43).

Bis auf die Aspekte akustische Verständlichkeit und Mehrwert durch das Hören im Vergleich zum Lesen, handelt es sich um Kriterien, die auch für die Auswahl von Lesetexten herangezogen werden könnten. Gleichzeitig werden sie aber mit hörverstehensspezifischen Attributen versehen (z.B. *Varieties* oder der Einsatz von Podcasts im Bereich Authentizität).

Ein anderes Element, das sich in ähnlicher Weise auf andere Kompetenzbereiche übertragen lässt, ist der Bereich der Aufgaben. Hierbei wird von den Fachleitungen vielfach auf die Sequenzierung der Aufgaben in *Pre-/While-/und Post-Listening* verwiesen, die sie generell als positiv bewerten (z.B. FL A, 51). Im Hinblick auf die Formate der Aufgaben zeigen sich die Fachleitungen in vielen Fällen kritisch gegenüber üblichen geschlossenen und halb-offenen Aufgabenformaten im Rahmen von Unterrichtssituationen. Nicht nur seien diese besonders schwierig selbst zu konstruieren, sodass sie wirklich sinnvoll für den Lernprozess seien (z.B. FL B, 61; FL G, 75; „Da gehört ganz bestimmt nicht dazu, dass die Schüler viele Häkchen machen, ähm, das, ähm, ich glaube, das haben wir auch an anderer Stelle schon gesehen, ähm, dass viele Häkchen machen in der Regel wenig hilft“ FL A, 17). Auch gibt es die Auffassung, dass diese Aufgabenformate sehr stark den Rezeptionsprozess lenken und damit auch die Lernenden in gewisser Weise einschränken:

„Ähm, das, also wäre dann ein Kriterium, ähm, die Aufgabenstellung muss schülerorientiert sein oder auch eher selbst, ähm, selbstentdeckend oder, ja einen persönlichen Zugang, ähm, zum Text ermöglichen und nicht den Lehrer, also meine Lehrerfragen [gemeint sind geschlossene Aufgabenformate] beantworten unbedingt.“ (FL D, 55).

Gerade vor dem Hintergrund der Förderung der Hörverstehenskompetenz, bei der die geschlossenen und halb-offenen Aufgabenstellungen oft auf der Ebene des AFB I verbleiben, wurde darauf verwiesen, dass höhere AFBs andere Aufgabenformate erfordern. Für Fachleitung G ist es dementsprechend beispielsweise auch von Bedeutung, dass mit den Informationen sinnvoll weitergearbeitet wird:

„Also, wenn ich das, wenn ich eine Aufgabe konzipiere, mit Multiple-Choice, kann ich ja durchaus dann, wenn ich dann die Anforderungen, die richtigen Anforderungen vorgebe, kann ich ja auch sehen, inwiefern die Schülerinnen und Schüler die Aufgaben dann auch korrekt bearbeitet haben. Das kann ich natürlich sehen. Aber es ist natürlich, ist ja die Frage, wie man jetzt auch die die Weiterarbeit auch anlegt. Also, wenn ich jetzt das Hörverstehen nutze als Vorbereitung, um danach aufs Sprechen überzuleiten, dann kann ich ja anhand des Outputs – des Sprachprodukts – auch sehen, ob die wesentlichen Inhalte, die zu identifizieren waren im Rahmen des Hörverstehenteils, ob die dann auch angewendet werden können.“ (FL G, 77)

Insgesamt werden aber auch die Vorteile geschlossener und halb-offener Aufgabenformate gesehen, z. B. im Hinblick auf die Aktivierung schwächerer Schüler:innen, die von stärker gelenkten Aufgaben profitieren können (z. B. FL C, 63; FL G, 81). Darüber hinaus sind sich die Fachleitungen auch dessen bewusst, dass diese Aufgabenformate die Schüler:innen auf die Überprüfung dieser Kompetenz (z. B. im Rahmen von VERA8 oder der ZP10) vorbereiten. Ebenfalls im Kontext der Aufgabenstellung wird von den Fachleitungen auf das vielfach in der Praxis auffindbare Prinzip vom Global- zum Detailverstehen eingegangen. Während dies für einige Fachleitungen als gängiger Standard gesehen wird (z. B. FL E, 61), gibt es auch kritische Stimmen, die dies im Lernkontext einer Unterrichtsstunde als unnatürliche Bevormundung der Lernenden sehen:

„Global to Detail wäre so ein Approach oder erstmal Globalverstehen und dann Detail und meine Erfahrung ist aber was anderes. Also ich finde das immer total gängelnd. Wenn ich einen offenen Impuls setze, ‚was habt ihr verstanden‘, dann nennen die Schüler mir teilweise minikleine Details, die nur der Schüler verstanden hat, und andere nennen mir ganz grobe, ähm, Eckpunkte. Ja es geht so ungefähr um, und das ist die Spanne. Und da denke ich dann immer für den Unterricht, ähm, ich will die ja nicht abwürgen. Also die sollen mir ruhig Details nennen. Ich kann das hinterher mit den Schülern sortieren, was ein Detail ist und was ein Grundverständnis war.“ (FL D, 53).

Ein letzter Aspekt, der für Hörverstehensstunden spezifisch ist, betrifft die Ebene der Technologie. Grundsätzlich unterscheiden die Fachleitungen zwischen der zentralen Bereitstellung des Hörverstehentexts durch die Lehrkraft oder aber einer dezentralen Verfügbarkeit, bei der die Lernenden individuell den Hörverstehentext steuern (z. B. FL B, 73 & 75; FL C, 81; FL D, 77; FL E, 85 & 87). Auch hierbei ist – bei allen Chancen und Risiken, die sich durch die Digitalität bieten – für die Fachleitungen das entscheidende Gütekriterium die Passung des technologischen Elements sowohl zur Lerngruppe als auch zum Lernziel der jeweiligen Unterrichtsstunde. Im Hinblick auf die Chancen der Digitalität wird von den Fachleitungen vor allem das Potenzial für differenziertes und individualisiertes Hörverstehen benannt (z. B. durch die individuelle Steuerung des Hörverstehentexts über ein digitales Endgerät; z. B. FL C, 70–71; FL G, 89) sowie die Möglichkeit, sehr leicht an authentisches Ausgangsmaterial zu kommen. Gleichzeitig liegt hierin auch aus Sicht der Fachleitungen ein sehr großes Risiko. Zum einen können sich sowohl die Lehrkräfte als auch die Schüler:innen in den Weiten des hörbaren Materials verlieren (z. B. FL D, 75; FL F, 127). Zum anderen erfordert gerade die Differenzierung eine besonders gründliche didaktische Aufbereitung (z. B. FL D, 77; FL E 87). Insgesamt bewerten die Fachleitungen den (zielführenden) Einsatz digitaler Medien im Kontext von Hörverstehensunterricht aber als sehr positiv.

5 Interpretation der Ergebnisse

5.1 Implikationen der Ergebnisse im Hinblick auf das MAIN-TEACH Modell

Die von den Fachleitungen genannten Kriterien für guten Englischunterricht spiegeln in vielfacher Weise die im MAIN-TEACH-Modell (vgl. insbes. Lindl et al., in diesem Band) abgebildeten Dimensionen von Unterrichtsqualität wider. Das MAIN-TEACH-Modell weist drei miteinander interagierende Ebenen auf. Das Fundament des Modells, die unterste Ebene, wird von den Aspekten Differenzierung und Adaptivität gebildet. Diese sind eng mit der darüberliegenden, mittleren Ebene verknüpft, die allgemeine Aspekte der Unterrichtsqualität umfasst (Klassenführung, sozio-emotionale Unterstützung, Unterstützung der aktiven Beteiligung). Die oberste Ebene des Kerns besteht aus vier weiteren Dimensionen von Unterrichtsqualität: kognitive Aktivierung, Unterstützung des Übens, Formatives Assessment und Auswahl und Thematisierung von Inhalten und Fachmethoden. Gerahmt wird der Kern des Modells durch äußere Schalen, die sowohl die Zielsetzungen als auch die *Outcomes* (der Kompetenzerwerb der Schüler:innen) widerspiegeln. Im Folgenden werden die Aussagen der Fachleitungen vor allem auf die drei beschriebenen Ebenen bezogen und interpretiert.

Der Befund der Interviews, dass für die Fachleitungen die Passung zwischen Lerngruppe und den weiteren, spezifischen didaktischen Entscheidungen die entscheidendste Rolle zu spielen scheint, ist an dieser Stelle als bedeutsam hervorzuheben. In der angesprochenen Passung spiegelt sich die unterste Ebene des MAIN-TEACH-Modells wider und bestätigt dessen Annahme, dass Differenzierung und Adaptivität das Fundament für alle weiteren, darüberliegenden Ebenen bilden. Im Hinblick auf die Fachspezifik ist anzumerken, dass abhängig von der individuellen Intensität der Auseinandersetzung der Lernenden mit der englischen Sprache in ihrer Freizeit einerseits sehr unterschiedliche Lernvoraussetzungen herrschen, andererseits aber auch reichhaltige Möglichkeiten für Lebensweltbezüge gegeben sind. Dass die Berücksichtigung des Lebensweltbezugs grundsätzlich ein entscheidendes Kriterium für Unterrichtsqualität ist, wird nicht nur von den Fachleitungen an vielen Stellen hervorgehoben, sondern geht auch aus dem MAIN-TEACH Modell hervor. Der Aspekt des Lebensweltbezugs findet sich sowohl in der obersten Ebene bei der „Auswahl und Thematisierung von (lebensweltbezogenen) Inhalten und Fachmethoden“ wieder als auch in der „kognitiven Aktivierung“. Ergänzend könnte man hier auch noch die fachspezifische Dimension der kommunikativen Aktivierung (Wilden, 2021) nennen, die deutlich leichter zu erreichen ist, wenn die Lernenden über Gegenstände und Inhalte sprechen, die sie selbst betreffen. Aufgrund der Bedeutung, welche die Fachleitungen den Lebensweltbezügen des Fachs Englisch zuschreiben, wäre zu überlegen, ob dieser Aspekt nicht vielmehr eine tragende Rolle im Modell einnehmen sollte. Für das Fach Englisch und insbesondere für den Hörverstehensunterricht, der für die Schüler:innen durch den eigenen auditiv geprägten Medienkonsum per se

einen Lebensweltbezug haben kann, könnte sich eine „Aufwertung“ als Basisebene für alle weiteren Überlegungen der Unterrichtsgestaltung als gewinnbringend erweisen.

Wenn es um die Bewertung von qualitativ hochwertigem Unterricht geht, ist hervorzuheben, dass die Fachleitungen in vielerlei Hinsicht die Kernfacetten des MAIN-TEACH Modells ansprechen (vgl. bspw. die Liste von Bewertungskriterien für Unterrichtsentwürfe in Abschnitt 4.1). Auch anhand dieser Kriterien wird deutlich, dass die Dimensionen Differenzierung und Adaptivität sowie Auswahl und Thematisierung von Inhalten und Fachmethoden für die Fachleitungen im Vordergrund stehen. Einschränkend ist anzumerken, dass es sich hierbei insofern um ein Artefakt der Interviewführung handeln kann, als allgemeinpädagogische Elemente von Unterrichtsqualität (wie z. B. Klassenführung) nicht Teil der Befragung waren, sondern es genuin um Fachspezifik ging. Es ist dementsprechend auffallend, dass die mittlere Ebene von Unterrichtsqualität nur selten zur Sprache kommt. Allein die im Modell auf der mittleren Ebene angesiedelte *Unterstützung der aktiven Beteiligung*, die sich beispielsweise durch ein entsprechendes Setting und den Fokus auf kommunikative Aktivierung erreichen lässt, wird von allen Fachleitungen als ein Element guten Englischunterrichts benannt. Zuletzt ist anhand der von den Fachleitungen aufgeführten Bewertungskriterien festzustellen, dass sie prinzipiell einen zielorientierten Unterricht erwarten, was zur Kernannahme des MAIN-TEACH Modells passt und durch die äußeren Schalen symbolisiert wird.

Ein Diskurs, der in allen Interviews geführt wird, bezieht sich auf die Dimension der kognitiven Aktivierung, die in der Thematisierung der drei AFBs in den Interviews zum Ausdruck kommt. Die Herausforderung, dass die kognitiven Verarbeitungsprozesse beim Hören nicht direkt beobachtet werden können, wurde von allen Fachleitungen geteilt. Dennoch sehen sie Möglichkeiten, die kognitive Aktiviertheit durch gezielte Unterrichtsgestaltung sichtbar(er) oder greifbar zu machen, zum Beispiel indem mit entsprechenden Anschlussaufgaben die verarbeiteten Informationen weiter genutzt werden. Die Fachleitungen heben hervor, dass ein Hörverstehensunterricht, der sich allein auf die bisher üblichen Aufgabenformate (d. h. vor allem halb-offene und geschlossene) stützt, nur selten über den AFB I hinausgeht und dementsprechend kaum Aussagen über tiefergehende Verarbeitungsebenen zulässt. Um im Rahmen von Hörverstehensunterricht eine höhere kognitive Aktiviertheit anzustreben – eine Zielsetzung, die nach Meinung der Fachleitungen sehr lohnenswert sei –, sind zwei mögliche Wege denkbar (vgl. Abschnitt 4.1). Zum einen eine inhaltsorientierte kognitive Aktivierung, die eine tiefergehende Auseinandersetzung mit den gehörten Inhalten anstrebt (z. B. durch weiterführende, kommunikative Aufgaben) oder aber eine metakognitiv-orientierte Ausrichtung, bei der es darum geht, die eigene Hörverstehenskompetenz und damit verbundene und angewendete Strategien zu reflektieren.

Eine im Hinblick auf Unterrichtsorganisation wichtige Diskussion, die von den Fachleitungen in diesem Kontext angesprochen wird, ist die Frage des Zeitmanagements, da sich nicht alle Ziele in 45 oder 60 Minuten erreichen lassen.

Dies spiegelt die bei Lindl et al. (in diesem Band) geführte Diskussion wider, dass verschiedene Lernziele miteinander konkurrieren oder einfach auch längere Unterrichtseinheiten erfordern können (vgl. Abschnitt 2.3). Gerade für einen kommunikativen Englischunterricht, der sich an komplexen Kompetenzaufgaben orientiert, ist die Erfassung von Unterrichtsqualität innerhalb von einer allzu kurzen Unterrichtseinheit dementsprechend herausfordernd. Zwar ist die zeitliche Dimension im MAIN-TEACH-Modell durch die äußeren Schalen angedeutet, jedoch wird nicht deutlich genug, wie stark die Zielsetzungen und der tatsächliche Lernzuwachs einer Stunde mit der zeitlichen Dimension verknüpft sind, sodass beispielsweise bestimmte Zielsetzungen (z. B. im Hinblick auf AFB I-III und kognitive Aktiviertheit) aufgrund mangelnder Zeit nicht zu erreichen sind.

5.2 Implikationen der Ergebnisse für die Konzeptionalisierung von Elementen des Hörverstehensunterrichts und Ausblick

Die Konzeption der Elemente des Hörverstehensunterrichts (vgl. Abschnitt 2.2) wurde entsprechend des Befunds erweitert, dass eines der bedeutsamsten Kriterien für guten Hörverstehensunterricht die Passung zur Lerngruppe ist. In dieser von den Fachleitungen geforderten Passung spiegeln sich zwei Dimensionen des MAIN-TEACH-Modells, die hier zum Tragen kommen. Zum einen erfordert eine zielgerichtete Passung zur Lerngruppe eine entsprechende Diagnostik mit einhergehender Rückmeldung an die Schüler:innen (Dimension *formatives Assessment*). Zum anderen muss aber auch auf Basis dieser Diagnostik eine Anpassung des Unterrichts erfolgen, die das Lernen aller Schüler:innen im Blick hat (Dimension *Unterstützung des Lernens aller Schüler:innen*). Dieser Forderung der Fachleitungen wurde insofern Rechnung getragen, als dass die Bereiche des formativen Assessments und der Unterstützung des Lernens aller Schüler:innen als ein neues Element hinzugefügt wurden, das sowohl ein Feedback an die Lernenden einschließt als auch die didaktischen Entscheidungen der Lehrkraft beeinflusst. Diese ziehen dann wiederum auch die entsprechenden Anpassungen bei den Elementen Technologie, Hörtext und Aufgaben nach sich, da sie auf die jeweilige Lerngruppe abzustimmen sind (vgl. Abb. 2). Es muss hier angemerkt werden, dass insbesondere die Dimension des formativen Assessments (im Sinne von Diagnostik für lernförderliches Feedback) für den Bereich Hörverstehen noch weiter zu spezifizieren ist. Da eine reine Diagnostik anhand bisheriger, standard- und validitätsorientierter Testung nur einen Teil der kognitiven Prozesse abdeckt (insbesondere *lower order thinking processes* im Bereich AFB I), wäre eine Ergänzung um weitere Elemente notwendig. Aus wissenschaftlicher Perspektive wären hier a) eine Diagnose der basaleren, sprachdatenverarbeitenden Prozesse (Diagnostik mit *micro listening*-Aufgabenformaten wie bspw. von Rossa und Helsper, 2021, vorgeschlagen) und b) auch die Diagnose der verwendeten metakognitiven Strategien für das Hörverstehen (bspw. anhand des von Vandergrift et al., 2006, entwickelten MALQ-Fragebogens) als zielführende Alternativen zu nennen, die

das diagnostische Repertoire von Lehrkräften erweitern sollten. Gleichzeitig bietet gerade die bisherige standardorientierte Diagnostik wenig Anregung für Lehrkräfte, ihren Unterricht anzupassen und stärker auf Hörverstehen auszurichten, da eben vor allem *lower order thinking processes* diagnostiziert werden, die nicht zwingend eine ausführlichere unterrichtliche Auseinandersetzung rechtfertigen (vgl. Abschnitt 4.1). Bei einem immer stärker auditiv-ausgeprägten Medienkonsum Jugendlicher (wie ihn z.B. die Ergebnisse der JIM-Studien nahelegen, u.a. MPFS, 2022) ist diese Praxis jedoch als kritisch zu bewerten.

Bei der Beurteilung der Unterrichtsstunden, die die Fachleitungen sehen, spielt die im MAIN-TEACH-Modell auf der innersten Ebene zu findende Auswahl und Thematisierung von Inhalten und Fachmethoden insofern eine zentrale Rolle, als dass die Fachleitungen eine angemessene Sachanalyse und Begründung der didaktischen Entscheidungen im Hinblick auf Hörverstehen erwarten. Hierbei handelt es sich um Entscheidungen, welche die Lehrkraft in der Regel vor der Durchführung des Unterrichts im Hinblick auf die konkrete Ausgestaltung der Elemente Hörtext, Aufgaben und Technologie trifft. Dieser Aspekt wurde der Konzeption der Elemente des Hörverstehens hinzugefügt und wird in Abbildung 2 nun als ‚materielle Dimension‘ bezeichnet. Informationen über das Kompetenzniveau des Hörverstehens der Lerngruppe, die beispielsweise auf Basis eines umfassenden formativen Assessments generiert wurden, sollten sich entsprechend auf die Gestaltung der materiellen Dimension auswirken. Dies wird in den Interviews an einigen Stellen deutlich, beispielsweise wenn auf das Leitmotiv des Einsatzes von authentischem und lebensweltbezogenem Material verwiesen wird. Dieses müsse entsprechend der Fähigkeiten und Fertigkeiten der Lerngruppe ggf. eingeschränkt werden bzw. alternativ müssten andere Elemente des Unterrichts angepasst werden (z.B. die Aufgabenstellungen oder technologischen Möglichkeiten, den Hörtext individuell zu steuern; vgl. Abb. 2), um auch bei schwächeren Lernenden natürliche gesprochene Sprache im Rahmen von Hörverstehensunterricht einsetzen zu können. Welche konkreten Anpassungen der materiellen Dimension zielführend sind, wird im Rahmen des übergeordneten Forschungsprojekts erkundet, um weitergehende Aussagen treffen zu können. Eng verknüpft mit der Anpassung der materiellen Dimension an die Lerngruppe sind auch die Dimensionen kognitive Aktivierung und Unterstützung des Übens. Für den Hörverstehensunterricht spezifizierten die Fachleitungen einerseits die Bedeutung der AFBs (vgl. Abschnitt 4.1), andererseits auch die Notwendigkeit des Einsatzes geeigneter Aufgabenformate, die über eine Überprüfung des Verständnisses hinausgehen und höhere AFBs abdecken (vgl. Abschnitt 5.1).

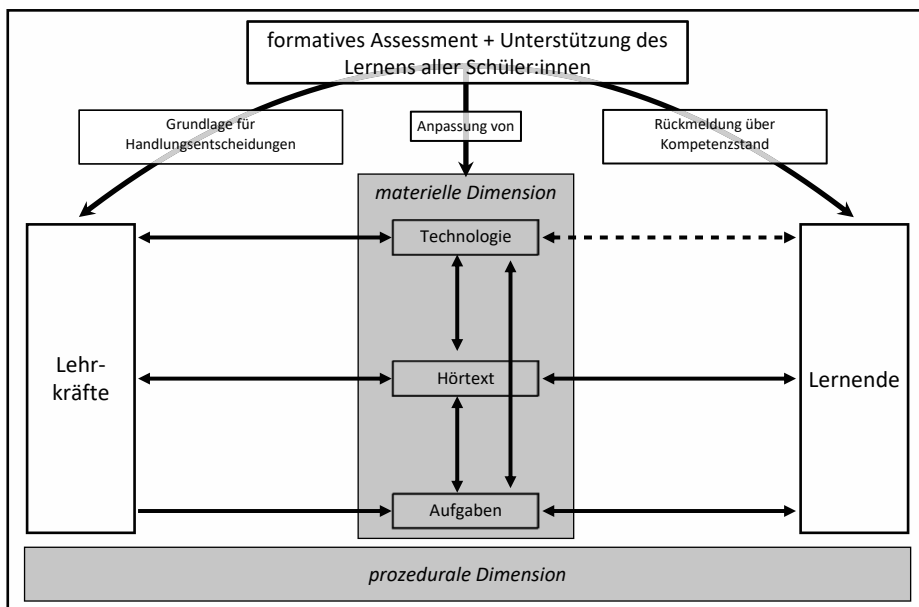


Abbildung 2: Angepasste Konzeptionalisierung der Elemente des Hörverstehensunterrichts

Darüber hinaus wurde auf den Faktor Zeit hingewiesen, der für Fachleitungen gerade im Kontext von Unterrichtsbesuchen eine bedeutende Rolle spielt. Eine Argumentationslinie der Expert:innen bezieht sich auf die Fähigkeit der Unterrichtenden, die fremdsprachendidaktischen Herausforderungen an den Unterricht mit der limitierten Unterrichtszeit zu vereinen. Aus diesem Grund betonen die Fachleitungen, dass es wichtig sei, sich in einzelnen Unterrichtsstunden im Kern auf einen beobachtbaren Kompetenzzuwachs zu konzentrieren. Die Unterrichtsreihe sollte jedoch eine sinnvolle Einbettung der Einzelstunde in einen komplexeren Lernzusammenhang gewährleisten. Da ein funktionales Erlernen des komplexen Lerngegenstands Sprache vor allem im Kontext zielführend ist, stellt sich hier eine fremdsprachenspezifische Herausforderung an Lehrkräfte, entweder innerhalb des Kontextes kleinere Lerneinheiten zu identifizieren oder aber Transparenz über das zu Erlernende über die Einzelstunden hinaus zu schaffen. Die Bedeutung der zeitlichen Rahmung zeigt sich bei der angepassten Konzeptionalisierung der Elemente des Hörverstehensunterrichts in Abbildung 2 anhand der ‚prozeduralen Dimension‘. Diese umfasst nach Ansicht des Autors auch die Zielorientierung des Unterrichts, wie sie im MAIN-TEACH Modell durch die äußeren Schalen verdeutlicht wird.

6 Limitationen der Studie

Im Folgenden werden gängige Gütekriterien qualitativer Forschung reflektiert, wie sie beispielsweise Caspari (2016, S. 16–18) für die Fremdsprachenforschung spezifiziert. Die intersubjektive Nachvollziehbarkeit (Steinke, 1999) wird insofern angestrebt, als dass die wesentlichen Entscheidungen und Schritte eines Forschenden möglichst transparent und nachvollziehbar dargestellt wurden, was jedoch nur bis zu einem gewissen Grad möglich ist. Ferner sollte die Verlässlichkeit von Daten und Vorgehensweisen insbesondere in puncto Transkription durch die Einhaltung zuvor aufgestellter Regeln sichergestellt werden. Eine Co-Kodierung hat nicht stattgefunden, allerdings wurde auf die strikte Befolgung des Kodiermanuals, welches in einem kritischen Peer-Review Prozess entstanden ist, geachtet. Durch die regelgeleitete Auswertung anhand der strukturierenden qualitativen Inhaltsanalyse wird suggeriert, dass möglichst offen mit den Daten umgegangen wird, allerdings ist dies nur in einem eingeschränkten Ausmaß tatsächlich der Fall, wie sich dies z.B. bei der Bezeichnung der induktiven Kategorien ablesen lässt. Durch die Vorüberlegungen des Autors und auch das spezifische Forschungsinteresse, das mit konkreten Fragestellungen an die Daten herantritt, wird deutlich, dass es keine vollständige Offenheit gegenüber den Daten geben kann. Um dennoch im Kontext des Forschungsvorhabens im höchsten Maße offen an die Auswertung der Daten heranzugehen, wurde ein hohes Maß an Selbstreflexivität angestrebt, das insbesondere eine Flexibilität mit unerwarteten Aussagen erfordert.

Zuletzt muss angemerkt werden, dass sich die vorliegenden Ergebnisse nur zu einem sehr geringen Maße generalisieren lassen. Es handelt sich um die Auffassungen einer kleinen Stichprobe von Expert:innen, die durch den explorativen Charakter vor allem dazu beitragen kann, Hypothesen für eine quantitativ orientierte Überprüfung zu generieren.

7 Zusammenfassung und Ausblick

Die der Studie zugrundeliegende Forschungsfrage lautete: Was sind aus Sicht von Expert:innen für Unterrichtsqualität spezifische Merkmale von und Bewertungskriterien für Hörverstehensunterricht? Auch die befragten Expert:innen taten sich schwer damit, spezifische Merkmale zu identifizieren. Die (zunächst trivial klingende) Schlussfolgerung einer Fachleitung, dass der Hörverstehensunterricht eigentlich auch nichts anderes sei als die Dekodierung von und Interaktion mit Texten, ist eine Erkenntnis mit weitreichenden Folgen für die Gestaltung des Hörverstehensunterrichts, der diese Prämisse nur unzureichend berücksichtigt. Der bisherige Schwerpunkt des Hörverstehensunterrichts lag tendenziell auf einer oberflächlicheren inhaltlichen Auseinandersetzung, die hinsichtlich der verwendeten Aufgabenstellungen nur bedingt über die Wiedergabe des Hörtextinhalts (AFB I) hinausgeht. Diese Art des Hörverstehensunterrichts wird hervorgerufen

und weiter gefördert durch die isolierte Überprüfung des Hörverstehens im Rahmen von standardorientierten Tests, die vor allem diese kognitiven Bereiche abdeckt. Diese Form der Prüfungspraxis ist jedoch als fehlgeleitet zu bezeichnen, da sie eine Unterrichtspraxis nach sich zieht, die der Herausforderung, Hörtexte zu dekodieren und mit ihnen in Interaktion zu treten, nicht umfassend genug gerecht wird und die Lernenden auch nicht dabei unterstützt, genau dies zu erlernen. Damit die Schüler:innen eine für die Lebenswelt angemessene Hörverstehenskompetenz entwickeln können, sollten sie – insbesondere in Lernsituationen – auch über die Ebene der reinen Informationsentnahme hinausgelangen. Hierzu muss auch im Hörverstehensunterricht *higher order thinking* angebahnt und gezielt in den Blick genommen werden. Gerade im Hinblick auf den Lebensweltbezug des Unterrichts, der von den Fachleitungen eingefordert wird, besteht hier Nachholbedarf. Ein Großteil der lebensweltlichen Texte, mit denen sich Lernende in ihrer Freizeit auseinandersetzen, wird gehört und nicht gelesen (MPFS, 2022). Schüler:innen sollten dementsprechend nicht nur wie bisher in die Lage versetzt werden, den Inhalt akustischer Texte (Hörtext oder audiovisueller Text) wiedergeben zu können (AFB I) und unhinterfragt als solches stehen zu lassen. Vielmehr sollten sie dazu angeregt werden, diesen zu analysieren (insbesondere im Hinblick auf die Frage, welche Mittel der Hörtext nutzt, um Hörer:innen zu lenken, AFB II), aber auch zu evaluieren und zu beurteilen (AFB III).

Die von den Fachleitungen benannten Bewertungskriterien wurden auf grundlegende Prinzipien reduziert, die zu einer Anpassung der Konzeptionalisierung der Elemente des Hörverstehensunterrichts geführt haben (siehe Abb. 2). Die darin aufgeführten Elemente lassen sich wiederum je nach Zielsetzung des Unterrichts anpassen und können vor allem durch eine Angleichung an das Kompetenzniveau der Lerngruppe (das heißt durch langfristige Berücksichtigung der prozeduralen Ebene) zur Erhöhung der Unterrichtsqualität beitragen. Hierbei spielt insbesondere die Auswahl der Aufgaben eine bedeutende Rolle. Diese sollten darauf abzielen, spezifische Ebenen des Hörverstehensprozesses in den Blick zu nehmen, um diese gezielt mit den Lernenden zu üben. Während sich intensive Aufgabenformen (wie z. B. *micro listening* Aufgaben, vgl. Rossa & Helsper, 2022; Rossa & Matz, 2023) vor allem für die Automatisierung von Dekodierungsprozessen eignen und in jüngeren Jahrgängen verstärkt eingesetzt werden sollten, eignet sich der Einsatz von extensiven und interaktionalen Aufgabenformaten für eine vertiefte inhaltliche Auseinandersetzung mit dem Gehörten, da diese auch *higher order thinking* in Lernsituationen anregen (vgl. hierzu z. B. den unterrichtspraktischen Vorschlag von Folkerts, 2023a). Ferner sollte eine gezielte Auseinandersetzung mit metakognitiven Strategien für das fremdsprachliche Hörverstehen früh angebahnt werden (für Möglichkeiten, dies umzusetzen vgl. z. B. Schlieckmann & Folkerts, 2023) und sich als langfristige Zielsetzung bis in die gymnasiale Oberstufe ziehen. Daran anknüpfend strebt das übergeordnete Forschungsprojekt der hier dargestellten Studie an, die identifizierten Zusammenhänge – wie sie in dem vorgeschlagenen Konzept für die Elemente des Hörverstehensunterrichts visuali-

siert wurden (vgl. Abb. 2) – für die Gestaltung von kompetenzorientiertem Hörverstehensunterricht weiter auszudifferenzieren.

Literatur

- Anderson, A. & Lynch T. (1988). *Listening*. Oxford: OUP. (Nachdruck 2004).
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J. & Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman. Verfügbar unter <https://www.celt.iastate.edu/wp-content/uploads/2015/09/Revised-BloomsHandout-1.pdf>
- Aryadoust, V. (2022). The Known and Unknown About the Nature and Assessment of L2 Listening. *International Journal of Listening*, 36(2), 69–79. <https://doi.org/10.1080/10904018.2022.2042951>
- Bezirksregierung Düsseldorf (2014). *Aufgaben von Fachleiterinnen und Fachleitern an einem Zentrum für schulpraktische Lehrerbildung in Nordrhein-Westfalen – eine Information*. Verfügbar unter https://www.brd.nrw.de/system/files/migrated_documents/media/document/2018-06/aufgabenprofil-fachleitungen.pdf
- Buck, G. (2001). *Assessing Listening*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511732959>
- Burns, A. & Siegel, J. (2018). *International Perspectives on Teaching the Four Skills in ELT: Listening, Speaking, Reading, Writing*. Cham: Springer. <https://doi.org/10.1007/978-3-319-63444-9>
- Caspari, D. (2016). Grundfragen fremdsprachendidaktischer Forschung. In D. Caspari, F. Klippel, M. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch* (S. 364–369). Tübingen: Narr Francke Attempto.
- Douglas, D. (1997). *Testing speaking ability in academic contexts: Theoretical considerations*. Princeton: Educational Testing Service.
- Field, J. (2003). Promoting Perception: Lexical Segmentation in L2 Listening. *ELT Journal*, 57(4), 325–334. <https://doi.org/10.1093/elt/57.4.325>
- Field, J. (2009). *Listening in the Language Classroom*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511575945>
- Field, J. (2019). *Rethinking the Second Language Listening Test: From Theory to Practice*. Sheffield/Bristol: Equinox.
- Folkerts, J.-F. (2023a). Are greenwashing campaigns a crime? Einen true crime-Podcast hören und in einer simulierten Gerichtsverhandlung diskutieren. *Der fremdsprachliche Unterricht Englisch*, 182, 39–45.
- Folkerts, J.-F. (2023b). Listening to Voices of Brooklyn. Mit TikTok Videos einen Stadtteil New Yorks erkunden. *Der fremdsprachliche Unterricht Englisch*, 183, 18–25.
- Folkerts, J.-F. (eingereicht). Listen Carefully, Please! Assessing Listening Comprehension. In F. Matz & D. Rumlich (Hrsg.), *Examinations in Foreign Language Teaching: Theory and Practice*. Frankfurt: Peter Lang.
- Folkerts, J.-F. & Matz, F. (2024). The Challenge of Learning to Listen. Insights into a Design-Based Research Study in German EFL Secondary Education. In J. Reckermann, P. Siepman & F. Matz (Hrsg.), *Oracy in English Language Education: Insights from Practice-Oriented Research* (S. 125–145). Cham: Springer. https://doi.org/10.1007/978-3-031-59321-5_8

- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28, 55–75. [https://doi.org/10.1016/S0346-251X\(99\)00060-3](https://doi.org/10.1016/S0346-251X(99)00060-3)
- Goh, C. C. M. & Vandergrift, L. (2022). *Teaching and Learning Second Language Listening: Metacognition in Action* (2. Aufl.). New York: Routledge. <https://doi.org/10.4324/9780429287749>
- Helfferrich, C. (2011). *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews* (4. Aufl.). Wiesbaden: VS. <https://doi.org/10.1007/978-3-531-92076-4>
- Helfferrich, C. (2019). Leitfaden- und Experteninterviews. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 669–686). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-21308-4_44
- Hitzler, R. (1994). Wissen und Wesen des Experten. In R. Hitzler, A. Honer & C. Maeder (Hrsg.), *Expertenwissen: Die institutionalisierte Kompetenz zur Konstruktion von Wirklichkeit* (S. 13–30). Wiesbaden: Vieweg Teubner. https://doi.org/10.1007/978-3-322-90633-5_1
- KMK (=Ständige Kultusministerkonferenz der Länder) (2023). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Ersten Schulabschluss und den Mittleren Schulabschluss*. In der Fassung vom 22.06.2023. Verfügbar unter https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2023/2023_06_22-Bista-ESA-MSA-ErsteFremdsprache.pdf
- Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (3. Aufl.). Weinheim/Basel: Beltz.
- Kriegel, U. (2013). The Phenomenal Intentionality Research Program. In U. Kriegel (Hrsg.), *Phenomenal Intentionality* (S. 1–26). New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199764297.001.0001>
- Liebold, R. & Trinczek, R. (2009). Experteninterview. In S. Kühl, P. Strodtzholz & A. Tafertshofer (Hrsg.), *Handbuch Methoden der Organisationsforschung* (S. 32–56). Wiesbaden: VS. https://doi.org/10.1007/978-3-531-91570-8_3
- Martínez Flor, A. & Usó Juan, E. (2006). *Current Trends in the Development and Teaching of the Four Language Skills*. Berlin/New York: De Gruyter. <https://doi.org/10.1515/9783110197778.3.139>
- Matz, F., Rumlich, D. & Rogge, M. (2023). *Die mündliche Prüfung. Eine Einführung*. Tübingen: Narr. <https://doi.org/10.24053/9783823393085>
- Mertes, M. (2019). *Hörverstehen im Englischunterricht. Bestandsaufnahme und Entwicklungsperspektiven*. Bochum/Freiburg: projektverlag.
- Meuser, M. & Nagel, U. (2009). Das Experteninterview – konzeptionelle Grundlagen und methodische Anlage. In S. Pickel, G. Pickel, H.-J. Lauth & D. Jahn (Hrsg.), *Methoden der vergleichenden Politik- und Sozialwissenschaft. Neue Entwicklungen und Anwendungen* (S. 465–479). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91826-6_23
- Misoch, S. (2019). *Qualitative Interviews*. Berlin/Boston: De Gruyter. <https://doi.org/10.1515/9783110545982>
- Medienpädagogischer Forschungsverbund Südwest und der Südwestrundfunk (MPFS) (2022). *JIM-Studie 2022, Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. Verfügbar unter https://www.mpfs.de/fileadmin/files/Studien/JIM/2022/JIM_2022_Web_final.pdf
- Praetorius, A. K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>

- Reckermann, J., Siepmann, P. & Matz, F. (Hrsg.) (2024). *Oracy in English Language Education: Insights from Practice-Oriented Research*. Cham: Springer. <https://doi.org/10.1007/978-3-031-59321-5>
- Rezaei, A. & Hashim, F. (2013). Impact of Awareness Raising about Listening Micro-skills on the Listening Comprehension Enhancement. An Exploration of the Listening Micro-skills in EFL Classes. *Australian Journal of Teacher Education*, 38(8), 1–15. <https://doi.org/10.14221/ajte.2013v38n8.4>
- Riemer, C. (2022). Befragung. In D. Caspari, F. Klippel, M. K. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik. Ein Handbuch* (2., vollst. überarb. und erw. Aufl., S. 162–180). Tübingen: Narr.
- Rossa, H. (2012). *Mentale Prozesse beim Hörverstehen in der Fremdsprache*. Frankfurt: Peter Lang. <https://doi.org/10.3726/978-3-653-01393-1>.
- Rossa, H. & Helsper, D. (2021). Intercultural Space Invaders. *Der fremdsprachliche Unterricht Englisch*, 172, 23–30.
- Rossa, H. & Matz, F. (2023). Learning to listen, listening to learn. Hörverstehen gezielt fördern. *Der fremdsprachliche Unterricht Englisch*, 182, 2–6.
- Rossa, H. & Meißner, F.-J. (2017). Hörverstehen. In B. Tesch, X. v. Hammerstein, P. Stanat & H. Rossa (Hrsg.), *Bildungsstandards Aktuell: Englisch/Französisch in der Sekundarstufe II* (S. 84–99). Braunschweig: Bildungshaus Schulbuchverlage.
- Rossa, H. & Nold, G. (2007). Hörverstehen. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen* (S. 178–196). Weinheim/Basel: Beltz.
- Rost, M. (2016). *Teaching and Researching Listening* (3. Aufl.). New York: Routledge. <https://doi.org/10.4324/9781315732862>
- Schlieckmann, R. & Folkerts, J.-F. (2023). What's my superpower? Mit einem Hörbuch Hörverstehensstrategien entwickeln. *Der fremdsprachliche Unterricht Englisch*, 182, 9–15.
- Siegel, J. (2018). Learning Listening. In A. Burns & J. C. Richards (Hrsg.), *The Cambridge Guide to Learning English as a Second Language* (S. 195–203). Cambridge University Press. <https://doi.org/10.1017/9781009024761.028>
- Steinke, Ines (1999). *Kriterien qualitativer Forschung: Ansätze zur Bewertung qualitativ-empirischer Sozialforschung*. Weinheim: Juventa.
- Ufer, S. & Praetorius, A.-K. (2022). Unterrichtsqualität im Mathematikunterricht der gymnasialen Oberstufe. In T. Rolfes, S. Rach, S. Ufer & A. Heinze (Hrsg.), *Das Fach Mathematik in der gymnasialen Oberstufe* (S. 287–315). Münster: Waxmann.
- Vandergrift, L. (2007). Recent Developments in Second and Foreign Language Listening Comprehension Research. *Language Teaching*, 40(3), 191–210. <https://doi.org/10.1017/S0261444807004338>
- Vandergrift, L., Goh, C., Mareschal, C & Tafaghodtari, M. (2006). The Metacognitive Awareness Listening Questionnaire: Development and Validation. *Language Learning*, 53(3), 431–462. <https://doi.org/10.1111/j.1467-9922.2006.00373.x>
- Wilden, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49, 211–219. <https://doi.org/10.1007/s42010-021-00105-y>
- Wilkinson, A. (1970). The Concept of Oracy. *The English Journal*, 59, 71–77. <https://doi.org/10.2307/811736>

Autor:innen des Bandes

Dr. Sonja Brunsmeier ist Fachdidaktikerin für Englisch. Zwei ihrer Arbeitsschwerpunkte sind interkulturelle kommunikative Kompetenz im Englischunterricht der Grundschule und die universitäre Ausbildung von Englischlehrkräften für die Primar- und Sekundarstufe. Sie verfügt über langjährige Erfahrung als Professorin, Englischlehrerin an Primar- und Sekundarschulen sowie Schulleiterin.

Jens-Folkert Folkerts arbeitete nach seinem Referendariat als Lehrer für die Fächer Englisch und Deutsch an einer Gesamtschule. Seit Februar 2021 ist er als abgeordneter Lehrer am Lehrstuhl für English Language Education an der Universität Münster tätig. Im Rahmen seines Dissertationsprojekts erforscht er Designprinzipien für kompetenzorientierten Hörverstehensunterricht im Fach Englisch.

Dr. Manuela Franke ist seit 2018 wissenschaftliche Mitarbeiterin am Lehrstuhl für Didaktik der romanischen Sprachen, Kulturen und Literaturen an der Universität Potsdam. Ihre Forschungsschwerpunkte sind unter anderem Literaturdidaktik, Lehrwerknutzungsforschung und digitale Leseforschung.

Joel Guttke ist seit 2021 wissenschaftlicher Mitarbeiter in der Didaktik des Englischen an der Universität Duisburg-Essen. Seine Forschungsschwerpunkte sind Unterrichtsqualität im Fremdsprachenunterricht mit besonderem Fokus auf kognitive Aktivierung, Professionalität von Fremdsprachenlehrkräften in der Primarstufe sowie die Test- und Fragebogenkonstruktion.

Dr. Maria Hofrichter ist wissenschaftliche Mitarbeiterin in der Forschungsgruppe FALKO-PV (*Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung*) am Lehrstuhl für Educational Data Science der Universität Regensburg. Ihre Forschungsschwerpunkte liegen in den Bereichen domänenspezifisches Professionswissen von Englischlehrkräften und fachspezifische Kriterien von Unterrichtsqualität im Fach Englisch.

Dr. Armin Jentsch ist Postdoctoral Fellow am Department of Teacher Education and School Research der University of Oslo. Seine aktuellen Forschungsschwerpunkte sind die Messung, Modellierung, Wirkung und Kontextabhängigkeit von Unterrichtsqualität.

Prof. Dr. Petra Kirchoff leitet seit 2023 den Lehrstuhl für Didaktik des Englischen an der Universität Augsburg. Zuvor war sie an der LMU München als Akademische Rätin und den Universitäten Regensburg und Erfurt als Professorin tätig. Ihre Forschungsschwerpunkte sind das Professionswissen von Englischlehrkräften sowie die englischspezifische Unterrichtsqualität und Forschung zum literarischen Kanon im Englischunterricht.

Annika Krämer ist Studienreferendarin im Vorbereitungsdienst für Lehramt Gymnasium mit den Unterrichtsfächern Latein, Altgriechisch und Ethik. Ihr Forschungsschwerpunkt liegt auf fachspezifischen Kriterien von Unterrichtsqualität in den alten Sprachen und der empirischen Forschung in der Lateindidaktik.

Dr. Alfred Lindl leitet die interdisziplinäre Forschungsgruppe FALKO-PV (*Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung*) an der Universität Regensburg. Seine Forschungsschwerpunkte liegen in den Bereichen professionelle Kompetenz von Latein- und Mathematiklehrkräften, fachspezifische Unterrichtsqualität, empirische Fachdidaktik und Methoden transdisziplinärer Bildungsforschung.

Dr. Viviane Lohe ist seit September 2020 akademische Rätin in der Abteilung der Sprachlehr- und -lernforschung (Englisch) an der Universität Erfurt. Ihre Forschungs- und Arbeitsschwerpunkte liegen im Bereich der Entwicklung von Sprachbewusstheit, der Mehrsprachigkeitsdidaktik, des frühen Fremdsprachenlernens und der Unterrichtsqualität in schulischen Fremdsprachen.

Prof. Dr. Raphaela Porsch ist seit 2019 Inhaberin des Lehrstuhls für Erziehungswissenschaft mit dem Schwerpunkt Allgemeine Didaktik an der Otto-von-Guericke-Universität Magdeburg. Ihre Forschungsschwerpunkte liegen im Bereich Schul- und Unterrichtsforschung, Allgemeine Didaktik, Lehrkräftebildung, Professionsforschung und Fremdsprachenforschung.

Mara Rader ist wissenschaftliche Mitarbeiterin in der Forschungsgruppe FALKO-PV (*Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung*) sowie am Lehrstuhl für Didaktik der deutschen Sprache und Literatur der Universität Regensburg. Ihre Forschungsschwerpunkte liegen in den Bereichen Professionswissen von Deutschlehrkräften und fachspezifische Kriterien von Unterrichtsqualität.

Prof. Dr. Daniel Reimann hat seit 2022 den Lehrstuhl für Didaktik der romanischen Sprachen und Literaturen an der Humboldt-Universität zu Berlin inne. Seine Forschungsschwerpunkte sind Mehrsprachigkeitsdidaktik (unter Berücksichtigung der Herkunfts- und Alten Sprachen), kompetenzorientierter Fremdsprachenunterricht, theoretische Modellierung kulturellen Lernens sowie Historiografie des Fremdsprachenunterrichts, der -didaktik und der -lehrkräftebildung mit Fokus auf romanische Sprachen.

Dr. Dennis Tark ist seit 2021 wissenschaftlicher Mitarbeiter im Fach Deutsch am Institut für Grundschulpädagogik der Universität Rostock. Seine Forschungsschwerpunkte sind L1- und L2-Erwerb, Mehrsprachigkeit(sdidaktik), ästhetische Bildung und Persönlichkeitsentfaltung im Unterricht, Didaktik des HSU, Sprachideologie und Spracheinstellungen, außerschulische Lernorte, Lehrkräfteprofessionalisierung und KI.