pedocs
Open Access Erziehungswissenschaften

DIPF

Lindl, Alfred; Lohe, Viviane; Kirchhoff, Petra

Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen – Grundlagen und Forschungsansätze

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, S. 9-37

Quellenangabe/ Reference:
Lindl, Alfred; Lohe, Viviane; Kirchhoff, Petra: Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen – Grundlagen und Forschungsansätze - In: 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, S. 9-37 - URN: urn:nbn:de:0111-pedocs-324200 - DOI: 10.25656/01:32420; 10.31244/9783830999201.01

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-324200
https://doi.org/10.25656/01:32420

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-License: 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

Mitglied der Leibniz-Gemeinschaft

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 qualitätsvollen 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 we-

Lindl, A., Lohe, V. & Kirchhoff, P. (2024). Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen – Grundlagen und Forschungsansätze. In V. Lohe, A. Lindl & P. Kirchhoff (Hrsg.), *Unterrichtsqualität in schulischen Fremdsprachen* (S. 9–37). https://doi.org/10.31244/9783830999201.01

nig 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 Fremdsprachen, 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 qualitätsvoll 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 disziplinare 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 qualitätsvoll 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 dieser 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 Angebot-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.

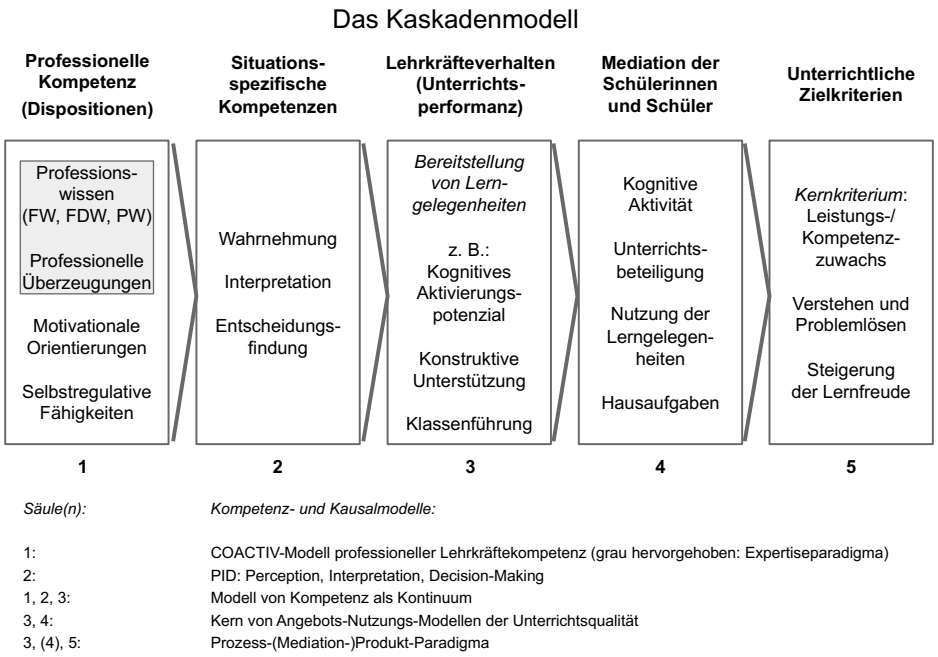

Abbildung 1: Das Kaskadenmodell (Krauss et al., 2020; übers. und zit. nach Lindl et al., 2023)

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 unmittelbarenintervention 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 mathematikspezifischer 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. (2020) identifizieren in ihrem (erweiterten) Syntheseframework, das auf einer Synthese und Weiterentwicklung von zwölf internationalen Ansätzen zur beobachtungsbasierten 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 (*M*ultilayered *a*nd *in*tegrated in conceptualizing the quality of *teach*ing; 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-Model 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).

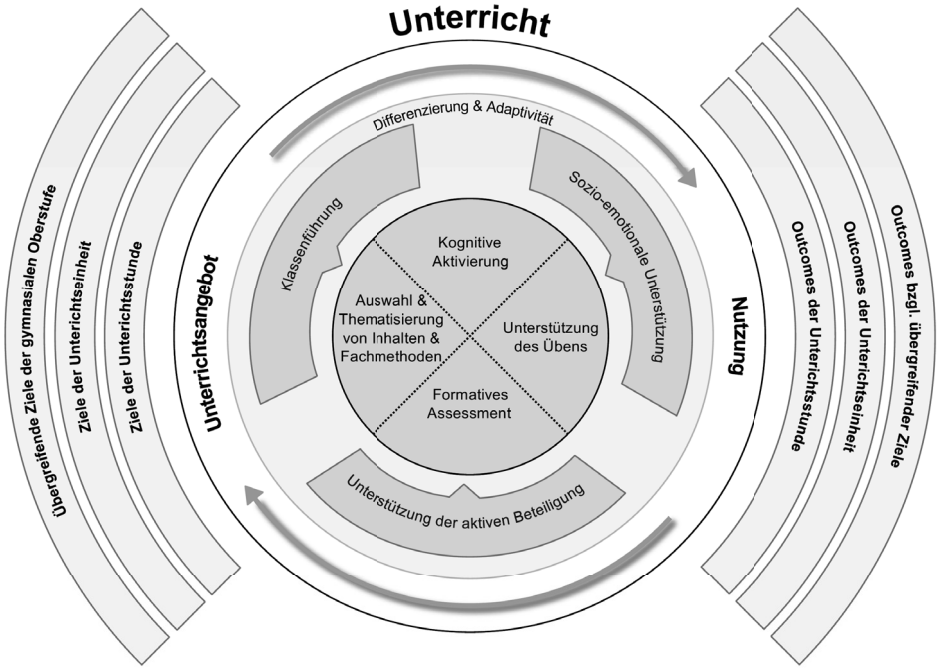

Abbildung 2: Adaptiertes MAIN-TEACH-Modell (Multi-layered and integrated in conceptualizing the quality of teaching; Charalambous & Praetorius, 2020; dt. Abb. nach Ufer & Praetorius, 2022)

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 Angebot-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 verschie-

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 Disziplinenspektrum 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 eigenen 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.

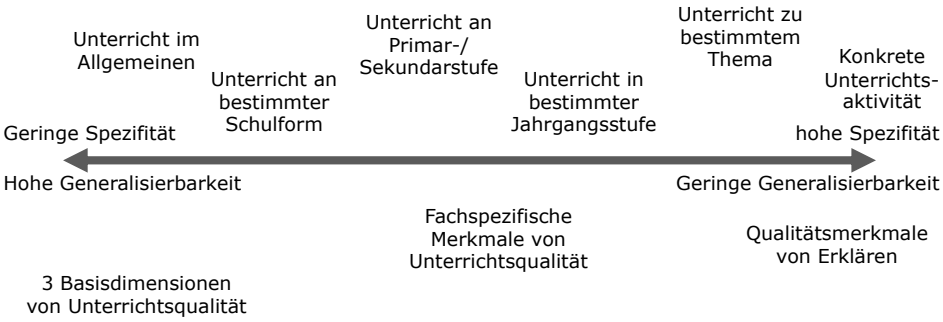

Abbildung 3: Spezifität als Kontinuum (mit exemplarisch benannten Abstufungen)

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-

ralisierbarkeit) 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

	Konzeptualisierung	
Operationalisierung	generisch	spezifisch
generisch	+ gute Vergleichbarkeit + hohe Verallgemeinerbarkeit + dekontextualisierte Theoriebildung – kontextlose Erfassung kaum umsetzbar – reduzierte inhaltliche Aussagekraft – evtl. eingeschränkte Validität	(nicht möglich)
spezifisch	+ Vergleichbarkeit gegeben + Verallgemeinerungen möglich + für Theoriebildung abstraktionsfähig + i. d. R. gegenstandsbezogen angemessene Erfassung – Überprüfung der konstrukttreuen Konkretisierung erforderlich – Aussagegehalt von umgesetzter Operationalisierung abhängig	+ 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	– 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	– Involviertheit beeinflusst Wahrnehmung – soziale/subjektive Erwünschtheit (eigener blinder Fleck) – verzerrte Wahrnehmung sogar von Oberflächenstrukturen (wie z. B. Redezeit)
Schüler:innen	– Adressat:innen des Unterrichts (z. B. Verständlichkeit, sozioemotionale Unterstützung) – Kenntnis der Kontextinformationen – variabler, längerer Beurteilungszeitraum möglich – ökonomisch durchführbar und hohe Zuverlässigkeit auf Klassenebene	– keine fachdidaktische oder pädagogische Expertise (z. B. kognitive Aktivierung) – Involviertheit in unterrichtliches Geschehen – (im Primarbereich: Verständnisschwierigkeiten und Fehlinterpretationen)
Beobachter:innen.	– nicht selbst involviert, kriterienfokussiert, trainiert, distanziertere Wahrnehmung – Vergleichsmöglichkeiten und Maßstäbe aufgrund Erfahrung – bei Videoverfahren: Standardisierung, Wiederholbarkeit, hohe Reliabilität und Validität	– 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 videographierten 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 et al. (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 Lehreräußerungen" (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, α = .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 Englischfertigkeiten 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. Hofrichter et al., in diesem Band; Guttke & Porsch, in diesem Band; Helsper, in Vorbereitung). 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 qualitätsvollen 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 qualitätsvollen Unterricht ausmacht, zu finden. Für einen ersten Überblick seien hier exemplarisch einige weiterführende Denkanstöße 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 qualitätsvollen 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 Kirchhoff** 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 lehrwerkbasierten 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 fu-

ßen die beiden 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 Kapitel „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 Kapitel „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-wissenssoziologische Professionsforschung. Perspektiven aus Früh- und Schulpädagogik, Fachdidaktik und Sozialer Arbeit* (S. 155–174). Bad Heilbrunn: Verlag 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

Brunsmeier, 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., 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

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., Lenske, 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 Ankopplung 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.

Guttke, 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). FALKO-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 Lehrerkompetenzen. 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.bildungsserver.de/onlineressource.html?onlineressourcen_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). *FALKO – Fachspezifische Lehrerkompetenzen. 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/l0.1007/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., Gutsmiedl, 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. & Röhrl, 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* – FALKO-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.), *FALKO – 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., Gutsmiedl, 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. https://doi.org/10.1515/dwir-2023-0062

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. Rolff, 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., Lauermann, 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., … Thim-Mabrey, C. (2021). FALKE: Experiences From Transdisciplinary Educational Research by Fourteen Disciplines. *Frontiers in Education, 5*, 579982. https://doi.org/10.3389/feduc.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