

Hofrichter, Maria; Lindl, Alfred; Rader, Mara; Krämer, Annika
**Was macht guten Fremdsprachenunterricht aus? Theoretische Überlegungen
und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern
Englisch und Latein**

*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. 109-137*



Quellenangabe/ Reference:

Hofrichter, Maria; Lindl, Alfred; Rader, Mara; Krämer, Annika: Was macht guten
Fremdsprachenunterricht aus? Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären
Onlinebefragung in den Fächern Englisch und Latein - 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. 109-137 -
URN: urn:nbn:de:01111-pedocs-324249 - DOI: 10.25656/01:32424; 10.31244/9783830999201.05

<https://nbn-resolving.org/urn:nbn:de:01111-pedocs-324249>

<https://doi.org/10.25656/01:32424>

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

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, 2018; Klieme et al., 2006). Gerade dies stellen aktuelle Forschungsansätze jedoch nicht zuletzt wegen einer – insbesondere interdisziplinär – inkonsistenten empi-

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* (S. 109–138). <https://doi.org/10.31244/9783830999201.05>

rischen 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 systematischen Review den Einflussfaktoren, Erfassungsmethoden sowie Effekten kognitiver Aktivierung. Weitere Übersichten zu Aspekten von Unterrichtsquali-

tät 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); 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)³ 	<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>
<p>(zentrale) Fragestellung und Ziele</p> <p><i>Welcher Fragestellung und Zielsetzung widmete sich die Studie?</i></p>	<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 Immersionsprogrammen 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
<p>Fokussierte Konstrukte</p> <p><i>Was wurde untersucht?</i></p>	<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 (Klitene, 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 - Videographie 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) - Videographie 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"> - Mehrebenenstrukturgleichungsanalysen 	<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 abschnitten 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. 	<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"> - 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"> - 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

<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 Zeitnutzung, 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	Themenv verwandte 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 qualitätvollen 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 (56 %)	93 (46 %)
I	Angemessene Auswahl, Thematisierung und Sequenzierung von Inhalten und Fachmethoden	28 (16 %)	32 (16 %)
I.1	Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten	3 (2 %)	5 (2 %)
I.2	Motivierung von Inhalten	18 (10 %)	3 (1 %)
I.3	Strukturierung der thematisierten Inhalte	7 (4 %)	22 (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 (3 %)	7 (3 %)
III.1	Unterstützung bei der Festigung von Lern- und Anwendungsprozessen	1 (1 %)	7 (3 %)
III.2	Konstruktiver Umgang mit Fehlern und Schwierigkeiten von Schülerinnen und Schülern beim Üben	5 (3 %)	0
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 lateinindaktischer 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)

Bem.: 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)	–

Bem.: 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), so dass 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(8). <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.
- 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. <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, Hildesheim University.
- 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, Hildesheim University.
- Kersten, K., Ponto, K., Bruhn A.-C. & Adler, M. (2019). Projekt SMILE – Abschlussbericht. *Studies on Multilingualism in Language Education*, 5, Hildesheim University.
- 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. Aufgabekultur 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-Bau-
lecke (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 FALKO-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.), *Konzepte und Bilanzen gelingenden Lehrens und Lernens – Grundlagen der Qualität von Schule* (Bd. 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. <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*, 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*. <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. & 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, E. (2021). Fachspezifische Aspekte von Unterrichtsqualität im Schulfach Englisch. *Unterrichtswissenschaft*, 49(2), 211–219. <https://doi.org/10.1007/s42010-021-00105-y>

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.

