

Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften – Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik

Alfred Lindl  · Laura Simböck  · Mara Rader  ·
Maria Hofrichter  · Maximilian Gutmiedl  · Stefan Böhringer  ·
Patrick Ehrich 

Eingegangen: 22. März 2024 / Überarbeitet: 21. März 2025 / Angenommen: 23. April 2025
© The Author(s) 2025

Zusammenfassung Der vorliegende Beitrag stellt eine Onlinebefragung zu Merkmalen guten Unterrichts bei Lehramtsstudierenden und Lehrkräften ($N=348$) der Fächer Deutsch, Englisch, Evangelische Religion, Latein und Musik vor. In dieser sollten die Teilnehmenden einerseits angeben, welche Merkmale sie selbst mit gutem Unterricht in ihrem Fach assoziieren, und andererseits in der Bildungswissenschaft und jeweiligen Fachdidaktik diskutierte generische und fachspezifische Qualitäts-

✉ Dr. Alfred Lindl · Laura Simböck · Mara Rader · Dr. Maria Hofrichter
Fakultät für Humanwissenschaften, Forschungsgruppe FALKO-PV, Universität Regensburg,
Sedanstraße 1, 93055 Regensburg, Deutschland
E-Mail: alfred.lindl@ur.de

Laura Simböck
E-Mail: laura.simboeck@ur.de

Mara Rader
E-Mail: mara.rader@ur.de

Dr. Maria Hofrichter
E-Mail: maria.hofrichter@ur.de

Maximilian Gutmiedl
Hochschule für katholische Kirchenmusik und Musikpädagogik Regensburg,
Andreasstraße 9, 93059 Regensburg, Deutschland
E-Mail: maximilian.gutmiedl@hfk-m-regensburg.de

Stefan Böhringer
Fakultät für Humanwissenschaften, Lehrstuhl Educational Data Science, Universität Regensburg,
Sedanstraße 1, 93055 Regensburg, Deutschland
E-Mail: stefan.boehringer@ur.de

Dr. Patrick Ehrich
Fakultät für Philosophie, Kunst-, Geschichts- und Gesellschaftswissenschaften, Institut für
Musikwissenschaft, Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Deutschland
E-Mail: patrick.ehrich@ur.de

aspekte hinsichtlich der Relevanz für ihre Disziplin einschätzen. Die Ergebnisse, deren Darstellungsfokus auf einer interdisziplinären Synopse liegt, bestätigen zwar grundsätzlich viele der in der Forschung erwogenen Kriterien aus Praxissicht. Sie machen aber neben fachübergreifenden Schwerpunktsetzungen auf einzelnen generischen Merkmalen auch fachbezogene und -kulturelle Unterschiede deutlich. Zudem geben die Freitextantworten der Teilnehmenden in allen betrachteten Domänen wertvolle Hinweise auf bislang wenig beachtete Aspekte guten Unterrichts. Damit leistet dieser Artikel einen Beitrag zum bildungswissenschaftlichen und fachdidaktischen Diskurs über die Generik bzw. Fachspezifik von Unterrichtsmerkmalen.

Schlüsselwörter Merkmale guten Unterrichts · Unterrichtsqualität · Fachspezifisch · Generisch · Onlinebefragung

Characteristics of good teaching from the perspective of (pre-service) teachers—results of an online survey in the five subjects German, English, Religious Education, Latin, and Music Education

Abstract This article presents an online survey on the characteristics of good teaching among pre-service and in-service teachers ($N=348$) in the subjects German as a First Language, English as a Foreign Language, Protestant Religious Education, Latin, and Music Education. The participants were asked to indicate which characteristics they associate with good teaching in their subject and to assess the relevance of generic and subject-specific aspects discussed in educational science and subject-matter didactics for their respective discipline. The results, which focus on an interdisciplinary synopsis, confirm many research-based criteria from a practical perspective. However, in addition to an interdisciplinary focus on individual generic aspects, they also highlight subject-related differences. Furthermore, the participants' free-text responses in all the domains analysed provide valuable information on characteristics of good teaching that have so far received scant scholarly attention. The article thus contributes to the educational science and subject-matter didactic discourse on the generic and subject-specific nature of teaching characteristics.

Keywords Characteristics of good teaching · Teaching quality · Subject-specific · Generic · Online survey

1 Forschungshintergrund und -fragestellungen

Zur analytischen Beschreibung von Unterricht sind in der bildungswissenschaftlichen Forschung Wirkungs- und Angebots-Nutzungs-Modelle (z.B. Helmke 2022; Krauss et al. 2020; Vieluf et al. 2020) wie auch eine grundlegende Unterscheidung von Oberflächen- und Tiefenstrukturen mittlerweile etabliert (z.B. Decristan et al. 2020; Drechsel und Schindler 2019; Kunter und Trautwein 2013). Unter Oberflächenstrukturen werden üblicherweise alle Merkmale subsummiert, die sich bei Betrachtung des unterrichtlichen Geschehens ad hoc und ohne aufwändiges Beob-

achtungsinstrumentarium erkennen lassen, wie Methoden (z. B. Projektarbeit, direkte Instruktion), Sozial- (z. B. Partner- oder Gruppenarbeit) oder Organisationsformen (z. B. jahrgangshomo- oder -heterogen). Tiefenstrukturen heben hingegen auf die inhaltliche Qualität von Lehr- und Lernprozessen ab und erfordern zu ihrer Beurteilung eine sorgfältige, systematische Analyse des Unterrichts (Drechsel und Schindler 2019; Kunter und Trautwein 2013). Zudem gelten eben diese als wesentliche Determinanten von Lern- und Leistungserfolgen (z. B. Blömeke et al. 2022; Hattie 2024; Kunter et al. 2013) und erfüllen damit eine zentrale Bedingung für Merkmale qualitativvollen Unterrichts im Sinne Berliners (2005). Denn dieser erachtet Unterricht als qualitativvoll, wenn er nicht nur *gut* ist, d. h. normativen Standards genügt, die moralisch, gesellschaftlich, kulturell wie auch fachlich geprägt sein können, sondern zugleich auch *effektiv*, d. h. nachweisbar lernwirksam.

1.1 Unterrichtsqualitätsforschung in der Bildungswissenschaft

Als Resultat explorativer Faktorenanalysen zur TIMS-Videostudie (Trends in International Mathematics and Science) schlugen Klieme et al. (2001) (effektive) Klassenführung, konstruktive Unterstützung und kognitive Aktivierung als grundlegende tiefenstrukturelle Merkmale von Unterricht vor, die seither zu den drei Basisdimensionen von Unterrichtsqualität avancierten (zsf. Lindl et al. 2024b; Praetorius et al. 2018). Ihre sparsame Modellstruktur und Gültigkeit wurden in zahlreichen Studien, die vorwiegend in mathematisch-naturwissenschaftlichen Fachkontexten entstanden, mit unterschiedlichsten methodischen Designs und heterogenen Operationalisierungen überprüft. Dabei zeigten sich hinsichtlich psychometrischer Gütekriterien, faktorieller Struktur, aber auch der Zusammenhänge mit schülerseitigen Lernerfolgen differente Ergebnisse (zsf. Praetorius et al. 2018). Angesichts dieser inkonsistenten, fachlich einseitigen Befundlage werden Exhaustivität, Validität und Transferierbarkeit der drei Basisdimensionen bezüglich anderer Schulfächer – nicht zuletzt auch aufgrund fehlender Replizierbarkeit beispielsweise in der DESI- (Deutsch-Englisch-Schülerleistungen-International: Helmke und Klieme 2008) oder PERLE-Studie (Persönlichkeits- und Lernentwicklung von Grundschulkindern: Gabriel 2014) – in der Forschung zunehmend kritisch hinterfragt (z. B. Begrich et al. 2023; Lipowsky und Bleck, 2019; Praetorius et al. 2020a, b; Praetorius und Gräsel 2021; Rothland 2024). Einerseits werden differenziertere Modelle vorgeschlagen, in denen die konstruktive Unterstützung in eine konstruktiv-kognitive und eine konstruktiv-emotionale Komponente aufgespalten (Kleickmann et al. 2020), ein zusätzlicher fachspezifischer Aspekt ergänzt (Lipowsky und Bleck 2019) oder andere Strukturierungen mit mehr Dimensionen entwickelt werden (z. B. Charalambous und Praetorius 2020; Wisniewski et al. 2020). So unterscheidet beispielsweise das auf das erweiterte Syntheseframework (Praetorius et al. 2018, 2020a) zurückgehende MAIN-TEACH-Modell (**M**ulti-layered **a**nd **i**ntegrated in conceptualizing the quality of **t**eaching, Charalambous und Praetorius 2020) die sieben Dimensionen Auswahl und Thematisierung von Inhalten und Fachmethoden, kognitive Aktivierung, Unterstützung des Übens, formatives Assessment, Unterstützung des Lernens aller Schüler:innen, sozio-emotionale Unterstützung und Klassenführung. Diese sieben Aspekte lassen sich in 21 Subdimensionen unterteilen, zu denen sich aufgrund

ihrer Relevanz für die vorliegende Studie in Appendix I (Lindl et al. 2025a) eine ausführlichere Darstellung findet.

Andererseits wird unter Rekurs auf die immense Bedeutung des Fachs für Lehr-Lernprozesse (Schilcher et al. 2021; Seidel und Shavelson 2007) dessen substanziellere Berücksichtigung – insbesondere von bislang wenig beachteten (fremd-)sprachlichen und gesellschaftswissenschaftlichen Disziplinen – als vielversprechende Zukunftsperspektive für die Modellierung und Untersuchung verschiedener Merkmale von Unterrichtsqualität diskutiert (Begrich et al. 2023; Reusser 2024). Dies geht mit der Forderung nach einem „intensive[n] interdisziplinäre[n] Austausch von generischen sowie fachspezifischen Unterrichtsforschenden“ (Praetorius und Gräsel 2021, S. 178) und einem stärkeren Einbezug fachbezogener, -kultureller und -didaktischer Perspektiven (z. B. Begrich et al. 2023; Dreher und Leuders 2021; Praetorius et al. 2020a, b; Praetorius und Gräsel 2021; Rothgangel 2021) einher, sodass Reusser und Pauli (2021, S. 192) sogar von einer „fachdidaktische[n] Wende in der aktuellen empirischen Lehr-Lernforschung“ sprechen.

1.2 Unterrichtsqualitätsforschung in der Fachdidaktik

Die aufgeworfene Frage, welche Merkmale qualitätsvollen Fachunterricht ausmachen, sollte zwar, wie mit Wilden (2021) anzunehmen ist, generell im Zentrum fachdidaktischer Forschung liegen. Dennoch wird sie in den verschiedenen Bezugsdisziplinen uneinheitlich betrachtet und der diesbezügliche Forschungsstand ist äußerst different. Im Vergleich zur umfangreichen Studienlage in Mathematik und den Naturwissenschaften (z. B. Dreher und Leuders 2021; Heinitz und Nehring 2020; Mu et al. 2022; Neuhaus 2021; Praetorius et al. 2018) erscheint diejenige in Fächern wie Deutsch (z. B. Rader et al. *in Vorb.*; Wiprächtiger-Geppert 2021), Englisch (z. B. Guttko 2023; Hofrichter et al. 2024; Lindl et al. 2024b; Wilden 2021) oder Musik (z. B. Björk et al. 2024; Gutmiedl et al. 2024; Kranefeld 2021; Nonte et al. 2024) noch eher überschaubar; in Evangelischer Religion (z. B. Rothgangel 2021; Schweizer 2024) und Latein (z. B. Lindl und Krämer 2022, Riecke-Baulecke et al. 2021) ist sie defizitär, da kaum Ansätze existieren.

In entsprechenden Vorarbeiten werden erste Vorschläge für fachliche Konkretisierungen generischer Merkmale bzw. Erweiterungen um zusätzliche Dimensionen unter Berücksichtigung fachlicher Spezifika entworfen (z. B. ästhetische Aktivierung: Björk et al. 2024; Kranefeld 2021; Lindl et al. 2024a; Puffer und Hofmann 2017; kognitiv-emotionale Aktivierung: Hesse und Winkler 2022; Fremdsprache als Kommunikationsmedium und Lerngegenstand: Hofrichter et al. 2024; Wilden 2021). Systematische Untersuchungen zur fachbezogenen Validierung bestehender Kriterien und Modelle qualitätsvollen Unterrichts wie auch zur Identifikation und Modellierung spezifischer Aspekte innerhalb eines Fachs sowie im Kontrast zu anderen Disziplinen liegen bislang erst vereinzelt vor bzw. fokussieren selten mehr als zwei (Kern-)Fächer (z. B. Deutsch, Englisch: Helmke und Klieme 2008; Praetorius et al. 2016; Deutsch, Mathematik: Eckerth et al. 2012; Gabriel 2014; Wemmer-Rogh et al. 2024; Deutsch, Mathematik, Englisch: Maag Merki 2012; Schreiner et al. 2019; alle Schulfächer: Wisniewski et al. 2020).

Neben dem bildungswissenschaftlichen und jeweiligen fachdidaktischen Diskurs sollten diese, um einer normativ-theoretischen Setzung und Auswahl der Forschenden vorzubeugen (Begrich et al. 2023), zudem die ökologisch valide Sicht praxisnaher und -erfahrener Protagonist:innen einbeziehen (Freitag und Theurer 2020; Prediger et al. 2022; Rothgangel 2021). Entsprechende Studien, die abseits von qualitativen Befragungsansätzen (z. B. Reinhardt et al. 2021; Rossa 2017) untersuchen, was (angehende) Lehrkräfte ihrer Erfahrung nach als Merkmale guten Unterrichts in ihrem Fach ansehen und nicht bestimmte Unterrichtsprozesse wie das Erklären (z. B. Findeisen 2017; Lindl et al. 2019; Schilcher et al. in Vorb.) anhand schon vorgegebener Kriterien bewerten, beschränken sich meist auf gewisse Lehr-Lerngegenstände (z. B. ästhetische Bildung: Freitag und Theurer 2020; Klassensingen: Ehrich 2025). In Bezug auf einzelne Fachbereiche existieren sie bislang nur für Sport (mit einem paneuropäischen Vergleich: Langer et al. in Vorb.) und auch ein (kontrastierender) Blick auf mehrere Fachgruppen steht bislang aus.

1.3 Forschungsfragen

An diesem Desiderat setzt die Forschungsgruppe FALKO-PV (**F**achspezifische **L**ehrkraft**k**ompetenzen – **P**rädiktive **V**alidierung) an, deren Fokus auf einer Untersuchung der Wirkungszusammenhänge zwischen domänenspezifischen Lehrkraftkompetenzen und Qualitätsmerkmalen wie auch Zielkriterien von Unterricht in den Fächern Deutsch, Englisch, Evangelische Religion, Latein, Mathematik und Musik liegt. Hierzu galt es in fünf der sechs beteiligten Disziplinen zunächst zu klären, was guten Unterricht jeweils ausmacht. In diese konzeptionellen Überlegungen gingen entsprechend der oben genannten Forderung nach einem Einbezug praxisnaher Protagonist:innen die Ergebnisse einer Onlinebefragung bei Lehramtsstudierenden und Lehrkräften – kurz: (angehende) Lehrkräfte – ein, deren Leitfragestellungen auch diejenigen dieses Beitrags sind:

1. Welche Merkmale assoziieren (angehende) Deutsch-, Englisch-, Ev. Religions-, Latein- und Musiklehrkräfte mit gutem Unterricht in ihrem Fach? Lassen sich diesbezüglich fachübergreifende oder -spezifische Aspekte identifizieren?
2. Wie relevant sind Merkmale guten Unterrichts, wie sie in Bildungswissenschaft und Fachdidaktik diskutiert werden, nach Einschätzung von (angehenden) Lehrkräften in Bezug auf ihr Fach? Welche Gemeinsamkeiten und Unterschiede werden hier erkennbar?

2 Studiendesign und methodisches Vorgehen

Um diese Forschungsfragen zu untersuchen, wurde im Frühsommer 2022 eine Onlinebefragung zu Merkmalen guten Unterrichts in den Fächern Deutsch, Englisch, Ev. Religion, Latein und Musik durchgeführt. Sie folgte einem ergebnisoffenen, explorativen Ansatz und richtete sich an Lehramtsstudierende und Lehrkräfte. Zur Illustration der nachstehenden Erläuterungen sind das eingesetzte Instrument und Auszüge aus dem Kodiermanual in den digitalen Appendices I und II (Lindl et al. 2025a, b) einsehbar.

2.1 Stichprobe und Durchführung

Die Einladung zur Onlinebefragung mittels LimeSurvey wurde über persönliche Kontakte, Mailinglisten (z. B. Schulnetzwerk der Universität Regensburg), professionsspezifische Gruppen und Kanäle in sozialen Medien (Facebook, Instagram), (über-)regionale Fortbildungsangebote, fachspezifische Netzwerke (z. B. Alumni der Fachrichtungen, Lehrbeauftragte und Betreuungslehrkräfte in erster oder zweiter Ausbildungsphase, religionspädagogisches Zentrum), Newsletter von Fachverbänden (z. B. Deutscher Altphilologenverband, Verband Bayerischer Schulmusiker) und schulbezogene Praxiszeitschriften (z. B. Lindl und Krämer 2022) in Deutschland und Österreich verbreitet. Die Mitwirkung erfolgte unter Einhaltung datenschutzrechtlicher Bestimmungen, freiwillig und unentgeltlich sowie ohne Bezug zu Studien- oder Qualifizierungsleistungen. Die Bearbeitung erforderte ungefähr zehn Minuten.

An der Onlinebefragung beteiligten sich insgesamt 348 Personen (299 aus Deutschland, davon 278 aus Bayern, und 49 aus Österreich), von denen 36 die Beantwortung bereits nach einigen Items zum bildungsbiographischen Hintergrund abbrachen. Unter den anderen 312 Personen absolvierten zehn keine lehramtsbezogene Ausbildung und verfügten über keine erkennbare pädagogische bzw. fachdidaktische Erfahrung (z. B. Studierende eines nichtlehramtsbezogenen Fachstudiengangs), weshalb sie aus den nachstehenden Analysen ausgeschlossen wurden. Um möglichst sicherzustellen, dass keine Person den Fragebogen wiederholt beantwortete und mehrfach im Datensatz enthalten war, wurde in LimeSurvey voreingestellt, dass der Fragebogen nur einmal pro IP-Adresse bearbeitet werden konnte. Zudem wurden die Daten – insbesondere die bildungsbiographischen Angaben – hinsichtlich wiederkehrender Muster gesichtet.

Die verbleibenden 302 Personen waren zum Zeitpunkt ihrer Teilnahme in einen universitären Lehramtsstudiengang für eines der genannten Fächer eingeschrieben ($N=106$) oder an der Schule ($N=165$), in der Lehrkräftebildung an einer Hochschule, z. B. im Bereich Fachdidaktik, bzw. mit Praxiserfahrung in der Bildungsadministration, z. B. als Fachreferent:in, tätig ($N=31$). 209 Teilnehmende fühlten sich dem weiblichen, 92 dem männlichen Geschlecht zugehörig; eine Person gab divers an. Das mittlere Alter der Stichprobe betrug 33,96 Jahre ($SD=12,13$) und ihre durchschnittliche Unterrichtserfahrung 10,28 Jahre ($SD=10,46$). Im Fach Deutsch beteiligten sich 70, in Englisch 55, in Religion 23, in Latein 61 und in Musik 93 Personen. Damit handelt es sich um eine nicht repräsentative Gelegenheitsstichprobe, mit der typischerweise gewisse Limitationen einhergehen (Abschn. 4.2).

2.2 Instrument: Onlinefragebogen

Zu Vergleichszwecken war der Fragebogen, der in Appendix II vollständig eingesehen werden kann, in allen fünf Fächern nahezu identisch und gliederte sich in zwei Abschnitte. Im ersten Teil wurde an die Teilnehmenden nach einigen Fragen zu ihrer Bildungsbiographie folgende Aufforderung gerichtet: ‚Beschreiben Sie in drei Schlagworten oder Stichpunkten, was Ihrer Meinung und Erfahrung nach guten <Deutsch>unterricht ausmacht.‘ Die Angabe in spitzen Klammern wurde an das entsprechende Fach angepasst und für die Antworten standen drei Freitextfelder

(mit einem jeweiligen Limit von 100 Zeichen) zur Verfügung. Die (angehenden) Lehrkräfte sollten also drei wichtige Aspekte benennen, die sie ad hoc und möglichst intuitiv mit gutem Unterricht in ihrem Fach assoziierten, weshalb im Vorfeld auf erläuternde Hinweise oder eine eventuell einflussnehmende Kontextualisierung verzichtet wurde. Insbesondere war mit Blick auf die Zielgruppe und bisherigen Erfahrungen mit dieser davon auszugehen, dass eine vorausgehende Nennung von Beispielen oder in der bildungswissenschaftlichen bzw. fachdidaktischen Forschung diskutierten Aspekten (wie im zweiten Teil des Fragebogens) zu Verzerrungen führen könnte. Ursachen hierfür könnten eine nicht intendierte inhaltliche Beeinflussung, soziale Erwünschtheit, (mangelnde) Erinnerungsleistungen und fehlende weiterführende Vorschläge oder Ideen sein (siehe jedoch nachstehend zum Freitextfeld am Ende des Fragebogens).

Im zweiten Fragebogenabschnitt wurden den Teilnehmenden je nach Fach 32 bis 40 Merkmale guten Unterrichts in randomisierter Listendarstellung angezeigt, um bei deren Bewertung wechselseitige Abgleiche zu ermöglichen. Diese Merkmale umfassten in allen Disziplinen die 21 generischen Subdimensionen des erweiterten Syntheseframeworks von Praetorius et al. (2020a, b) in deren Originalwortlaut (vgl. hierzu Tab. 4 und Appendices I und II). Hinzu traten elf (Deutsch) bis 19 (Religion) weitere Aspekte, die auf Basis von Literaturrecherchen im bildungswissenschaftlichen und fachdidaktischen Diskurs oder nach subjektiver Vorannahme der Mitglieder der Forschungsgruppe für ein oder mehrere Fächer als spezifisch gelten (z. B. aktiver Gebrauch der Fremdsprache als Kommunikationsmedium in Englisch bzw. als Reflexionsgegenstand in Latein, ästhetische Aktivierung in Deutsch und Musik: kurz dargestellt in Abschn. 1.1 und 1.2, ausführlicher z. B. bei Gutsmiedl et al. 2024; Hofrichter et al. 2024; Lindl 2024). Potenziell in mehreren Fächern relevante Kriterien wurden entsprechend in die jeweiligen Onlinefragebögen integriert, sodass manche Aspekte in allen oder mehreren, einige nur in einem Fach von (angehenden) Lehrkräften beurteilt wurden (Appendix II). Die konkrete Aufforderung an diese lautete: ‚Schätzen Sie bitte ein, wie wichtig nachstehende Aspekte Ihrer Erfahrung nach für guten Unterricht im Fach <Deutsch> sind.‘ Die Formulierung der Items war möglichst allgemein verständlich gehalten. Um Verzerrungseffekte durch Wissenschaftshörigkeit und Zustimmungstendenzen zu vermeiden, wurden die Teilnehmenden auch hier nicht über deren Hintergrund und Herkunft (z. B. Forschungsdiskurs, eigene Annahmen) informiert. Aus Gründen der Bearbeitungsdauer, Teilnahmebereitschaft und Akzeptanz wurden zudem keine weiterführenden terminologischen oder konzeptuellen Erklärungen implementiert. Damit wurde den Befragten die subjektive Interpretation der Items unter fachlicher Perspektive überantwortet. Auch wenn dies gewisse Freiheiten im interindividuellen Begriffsverständnis ermöglichte, wurde hierdurch Teilnehmenden aller Fächer eine gleichermaßen unbefangene Beurteilung auf Basis ihres domänenspezifischen Erfahrungshintergrunds und ihrer davon geprägten Auffassung zugestanden. Die Relevanzeinschätzungen erfolgten auf einer fünfstufigen Likert-Skala (0 unwichtig, ..., 4 wichtig).

Nach zwei weiteren Items zu Zielkriterien des jeweiligen Fachunterrichts (hier nicht fokussiert, Appendix II) erhielten die Teilnehmenden am Ende des Fragebogens ferner die Option, in einem Freitextfeld weitere, bislang unerwähnte und eventuell erst durch Bearbeitung dieser Umfrage ins Bewusstsein gerückte Merk-

Tab. 1 Auswertungsobjektivität bei Freitextantworten

	Deutsch	Englisch	Ev. Religion	Latein	Musik
Anzahl vergebener Antwortcodes	219	177	75	204	392
Prozentuale Übereinstimmung der Kodierer:innen	81 %	84 %	81 %	78 %	92 %

male guten Unterrichts zu ergänzen. In diesem Instrument wird also die Erfassung intuitiver Spontanantworten im ersten mit einer retrospektiv bilanzierenden Reflexionsmöglichkeit nach dem zweiten Fragebogenabschnitt kombiniert.

2.3 Kodierung der Freitextantworten

Die Angaben in den Freitextfeldern im ersten Abschnitt sowie gegen Ende des Fragebogens wurden zunächst deduktiv anhand der generischen (Sub-)Dimensionen des Syntheseframeworks von Praetorius et al. (2020a, b) und der darin enthaltenen Konkretisierungen bzw. Beispielfälle geordnet und kategorisiert (Appendix I). In diesem pragmatisch als Kodiermanual genutzten Schema nicht zu verortende Antworten ließen sich grob in zwei Gruppen unterteilen: Die einen bezogen sich nicht auf unterrichtliche Tiefenstrukturen, sondern auf Organisations- und Sozialformen, Persönlichkeitsmerkmale oder die Kompetenz von Lehrkräften. Sie wurden folglich nicht als Aspekte guten Unterrichts gewertet, aber aufgrund ihrer Häufigkeit zu einer eigenen Kategorie zusammengefasst (Tab. 2). Aus den anderen Antworten konnten interessante, neuartige und unter fachdidaktischen sowie -kulturellen Gesichtspunkten plausible Merkmale guten Unterrichts abgeleitet werden. Diese wurden zunächst einem iterativen Aushandlungsprozess mit allen Mitgliedern der Forschungsgruppe unterzogen, wobei interdisziplinäre Gemeinsamkeiten und Unterschiede herausgearbeitet wurden. Erst nach dieser diskursiven Validierung vor dem jeweiligen Fachhintergrund wurden ggf. neue Kategorien gebildet. Dies führte dazu, dass im Folgenden zwischen auf ein oder auch auf mehrere Disziplinen bezogenen Merkmalen (kurz: de, en, re, la, mu vs. fg „fächergemeinsam“ in Tab. 2) unterschieden werden kann.

Mithilfe des so induktiv ergänzten Kategorisierungssystems, das zu Vergleichszwecken in allen Fächern einheitlich angelegt wurde, werteten je Disziplin mindestens zwei Kodierer:innen (wiss. Mitarbeitende bzw. Lehramtsstudierende) alle offenen Antworten unabhängig voneinander aus. Jedem Freitextfeld konnten bei Bedarf auch mehrere Antwortcodes zugeordnet werden. Die erzielte Auswertungsobjektivität (proz. Übereinstimmung) war angemessen (Tab. 1); etwaige Abweichungen und Zweifelsfälle wurden abschließend diskutiert und konsensuell kodiert.

3 Ergebnisse

Dem explorativen, komparativen Zugang dieses Beitrags gemäß liegt der Schwerpunkt des nachstehenden Ergebnisteils auf einer deskriptiven Dokumentation domänenspezifischer Besonderheiten sowie auffälliger interdisziplinärer Gemeinsamkeiten und Unterschiede. Hierzu werden zunächst die Freitextantworten, daraufhin die

Relevanzeinschätzungen der (angehenden) Lehrkräfte betrachtet. Alle Auswertungen und Abbildungen wurden mit der Statistiksoftware *R* (R Core Team 2023) und Zusatzpaketen wie *ggplot2* (Wickham 2016) erstellt. Die Daten und Analysecodes sind in Appendix III (Lindl et al. 2025c) online verfügbar.

3.1 Freitextantworten

Aus Tab. 2 ist die absolute und relative Verteilung der 1067 Antwortcodes zu den Freitextantworten von (angehenden) Lehrkräften auf die in Abschn. 2.3 geschilderten Analysekatoren insgesamt sowie pro Fach zu entnehmen. Ein großer Anteil entfällt auf Antworten, die sich nicht auf Prozessmerkmale des unterrichtlichen Angebots beziehen (zwischen 21 % in Religion und 56 % in Deutsch bzw. Englisch), sondern Merkmale der Lehrkräftepersönlichkeit (z. B. Enthusiasmus, Fachwissen, Motivation), Sichtstrukturen (z. B. Methodenvielfalt, Sozialformen) oder Zielkriterien (z. B. Übersetzungskompetenz, Wortschatzerwerb) beschreiben. Ohne darauf angesichts der eng umrissenen Forschungsfragen dieses Beitrags im Detail eingehen zu können, stellt dies an sich schon einen diskussionswürdigen Befund dar (Abschn. 4.1).

3.1.1 Generische Dimensionen

Unter den Freitextantworten, die sich auf unterrichtliche Tiefenstrukturen beziehen, macht die Dimension *Auswahl und Thematisierung von Inhalten und Fachmethoden* mit 145 Nennungen (N.) den größten Anteil (16 %) aus, wobei deren Subdimensionen motivierende Einbettung (6 %) und Strukturierung (5 %) in allen Fächern mit Ausnahme von Latein ähnlich häufig Erwähnung finden. In Latein liegt der Fokus in höherem Maße auf Aspekten der Strukturierung (11 %). Akkuratheit und Korrektheit der thematisierten Inhalte spielen eine eher untergeordnete Rolle und werden in Englisch gar nicht genannt.

Der Dimension *kognitive Aktivierung* wird mit 20 Nennungen (2 %) vergleichsweise wenig Aufmerksamkeit zuteil; in den Antworten der (angehenden) Englischlehrkräfte ist sie sogar gar nicht zu finden. In den übrigen Fächern verteilen sich die wenigen Angaben vorwiegend auf die Subdimensionen der Aufgabenauswahl (1 %) und Unterstützung der kognitiven Aktivität der Schüler:innen (1 %).

Dem Aspekt *Unterstützung des Übens* werden über alle Fächer hinweg 31 Freitextantworten zugewiesen (3 %), die sich vor allem auf die Teilkategorie Festigung von Lern- und Anwendungsprozessen beziehen. Im Religionsunterricht ist Üben für (angehende) Lehrkräfte keine nennenswerte Kategorie guten Unterrichts, während sonst ein ähnlich großer Anteil an Antworten darauf entfällt.

Noch weniger als kognitive Aktivierung ist in den Freitextantworten der (angehenden) Lehrkräfte *formatives Assessment* mit gutem Unterricht assoziiert (7 N., 1 %). Wie schon das Üben erwähnen (angehende) Religionslehrkräfte diese Dimension nicht; lediglich die Subdimension eines qualitativ hochwertigen Feedbacks an die Schüler:innen findet in allen übrigen Disziplinen Erwähnung. Eine klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen und eine Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts sind kei-

Tab. 2 Absolute und relative Verteilung der Freitextantwortcodes je Analysekatgorie pro Fach

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
Anzahl vergebener Codes	1067	219	177	75	204	392
Kein Prozessmerkmal des Angebots	523 (49%)	122 (56%)	99 (56%)	16 (21%)	93 (46%)	193 (49%)
<i>I Auswahl und Thematisierung von Inhalten und Fachmethoden</i>	145 (16%)	35 (16%)	28 (16%)	14 (19%)	32 (16%)	36 (9%)
I.1 Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten, Realien sowie Fachmethoden	23 (2%)	4 (2%)	3 (2%)	1 (1%)	5 (2%)	10 (3%)
I.2 Lebensweltbezogene, motivierende Einbettung von Inhalten sowie Fachmethoden	60 (6%)	20 (9%)	18 (10%)	9 (12%)	3 (1%)	10 (3%)
I.3 Transparente und kohärente Strukturierung der im Unterricht thematisierten Inhalte sowie Fachmethoden	52 (5%)	8 (4%)	7 (4%)	2 (3%)	22 (11%)	13 (3%)
I.4 Akkuratheit und Korrektheit der vermittelten Inhalte	10 (1%)	3 (1%)	–	2 (3%)	2 (1%)	3 (1%)
<i>II Kognitive Aktivierung</i>	20 (2%)	5 (3%)	–	5 (7%)	4 (2%)	6 (2%)
II.1 Auswahl fachlich gehaltvoller und auf das kognitive Niveau der Schüler:innen abgestimmter Aufgaben	8 (1%)	4 (2%)	–	–	1 (0%)	3 (1%)
II.2 Einsatz fachlich gehaltvoller Aufgaben, die multiple Denkweisen, Bearbeitungsansätze und Lösungswege anregen	3 (0%)	–	–	–	–	3 (1%)
II.3 Unterstützung der kognitiven Aktivität der Schüler:innen (durch konstruktive Fehlerkultur, Explizierung von Denkprozessen etc.)	8 (1%)	1 (0%)	–	5 (7%)	2 (1%)	–
II.4 Unterstützung des metakognitiven Lernens der Schüler:innen (z. B. durch Möglichkeiten zur Reflexion des eigenen Handelns und verwendeter Strategien)	1 (0%)	–	–	–	1 (0%)	–
<i>III Unterstützung des Übens</i>	31 (3%)	8 (4%)	6 (3%)	–	7 (3%)	10 (3%)
III.1 Unterstützung bei der Festigung von Lern- und Anwendungsprozessen	23 (2%)	8 (4%)	1 (1%)	–	7 (3%)	7 (2%)
III.2 Konstruktiver Umgang mit Fehlern und Schwierigkeiten der Schüler:innen beim Üben	8 (1%)	–	5 (3%)	–	–	3 (1%)

Tab. 2 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latin	Musik
IV	7 (1%)	4 (2%)	2 (1%)	–	1 (0%)	1 (0%)
IV.1	–	–	–	–	–	–
IV.2	1 (0%)	1 (0%)	–	–	–	–
IV.3	6 (1%)	3 (1%)	1 (1%)	–	1 (0%)	1 (0%)
IV.4	–	–	–	–	–	–
V	39 (4%)	12 (6%)	3 (2%)	1 (1%)	6 (3%)	17 (4%)
V.1	4 (0%)	1 (0%)	–	–	3 (1%)	–
V.2	18 (2%)	6 (3%)	–	1 (1%)	3 (1%)	8 (2%)
V.3	17 (2%)	5 (2%)	3 (2%)	–	–	9 (2%)
VI	46 (4%)	4 (2%)	6 (3%)	15 (20%)	5 (2%)	16 (4%)
VI.1	39 (4%)	4 (2%)	6 (3%)	11 (15%)	5 (2%)	13 (3%)
VI.2	7 (1%)	–	–	4 (5%)	–	3 (1%)
VII	6 (1%)	–	2 (1%)	–	1 (0%)	3 (1%)
VII.1	4 (0%)	–	1 (1%)	–	1 (0%)	2 (1%)
VII.2	2 (0%)	–	1 (1%)	–	–	1 (0%)

Tab. 2 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
<i>Fachgemeinschafts/-spezifisches</i>	250 (23%)	29 (13%)	32 (18%)	24 (32%)	55 (27%)	110 (28%)
fg.1 Didaktische Reduktion unter Berücksichtigung zukünftiger fachlicher Lernschritte	7 (1%)	4 (2%)	–	2 (3%)	1 (0%)	–
fg.2 Kognitiv-ästhetische Aktivierung	44 (4%)	9 (4%)	–	2 (3%)	–	33 (8%)
fg.3 Fremdheitsverfahrung und reflektierender Umgang mit (fremd-)kulturellen Schemata	14 (1%)	4 (2%)	–	3 (4%)	–	7 (2%)
fg.4 Aktiver Gebrauch der Fremdsprache als Medium der Unterrichtskommunikation	8 (1%)	–	7 (4%)	–	1 (0%)	–
fg.5 Verwendung der Fremdsprache als Reflexionsgegenstand	11 (1%)	–	1 (1%)	–	10 (5%)	–
fg.6 Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses	10 (1%)	9 (4%)	–	–	1 (0%)	–
de.1 Intermedialität	1 (0%)	1 (0%)	–	–	–	–
de.2 Modelllernen	2 (0%)	2 (1%)	–	–	–	–
en.1 Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch	3 (0%)	–	3 (2%)	–	–	–
en.2 Schaffung hoher Sprechanteile der Schüler:innen durch echte kommunikative Anlässe	21 (2%)	–	21 (12%)	–	–	–
re.1 Dialogorientierung	4 (0%)	–	–	4 (5%)	–	–
re.2 Interreligiöses Lernen	1 (0%)	–	–	1 (1%)	–	–
re.3 Subjektorientierung	2 (0%)	–	–	2 (3%)	–	–
re.4 Christlicher Gegenstandsbezug	10 (1%)	–	–	10 (13%)	–	–

Tab. 2 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
la.1	1 (0%)	-	-	-	1 (0%)	-
la.2	24 (2%)	-	-	-	24 (12%)	-
la.3	1 (0%)	-	-	-	1 (0%)	-
la.4	1 (0%)	-	-	-	1 (0%)	-
la.5	6 (1%)	-	-	-	6 (3%)	-
la.6	3 (0%)	-	-	-	3 (1%)	-
la.7	5 (0%)	-	-	-	5 (2%)	-
la.8	1 (0%)	-	-	-	1 (0%)	-
mu.1	25 (2%)	-	-	-	-	25 (6%)
mu.2	5 (1%)	-	-	-	-	5 (1%)
mu.3	9 (1%)	-	-	-	-	9 (2%)
mu.4	1 (0%)	-	-	-	-	1 (0%)
mu.5	16 (1%)	-	-	-	-	16 (4%)
mu.6	14 (1%)	-	-	-	-	14 (4%)

ne Aspekte, die Lehrkräfte der fünf Fächer spontan mit der Frage nach den drei wichtigsten Kriterien guten Unterrichts verknüpfen.

In allen Fächern wurden Antworten gegeben, die sich der Dimension *Unterstützung des Lernens aller Schüler:innen* (39 N., 4%) zuordnen lassen. Darunter entfielen die meisten Nennungen auf Individualisierung und Anpassung (in Englisch unerwähnt) sowie auf die aktive Mitwirkung aller Schüler:innen (in Religion und Latein unerwähnt), während die Bereitstellung eines geeigneten Lernumfelds nur in Deutsch und Latein genannt wurde.

Merkmale der *sozio-emotionalen Unterstützung* (46 N., 4%) ziehen vor allem (angehende) Religionslehrkräfte in Betracht (15 N., 20%), wobei diesen wie auch den Teilnehmenden der anderen Fächer die Beziehung zwischen Lehrperson und Lernenden wichtig ist. Die Beziehungen der Schüler:innen untereinander werden in den Sprachfächern Deutsch, Englisch und Latein nicht thematisiert.

Schließlich finden sich unter den drei von (angehenden) Lehrkräften genannten Kriterien nur sehr wenige Bezüge zur *Klassenführung* (6 N., 1%). Auf Deutsch und Religion entfällt diesbezüglich sogar keine Angabe.

3.1.2 Fachgemeinsame und -spezifische Merkmale

Unter den in mehreren Fächern genannten Merkmalen, die sich nicht in das Syntheseframework einordnen ließen (Tab. 2), kommt am häufigsten der Aspekt kognitiv-ästhetische Aktivierung vor. Die meisten Nennungen stammen aus dem Fach Musik (33 N., 8%), während sich in den Sprachfächern Englisch und Latein überhaupt keine Antworten dieser Kategorie zuordnen lassen (Lindl et al. 2024a). Insgesamt zehn oder mehr Nennungen entfallen ferner auf die Aspekte Fremdheitserfahrung und reflektierender Umgang mit (fremd-)kulturellen Schemata (14, aber nicht in Englisch und Latein), Verwendung der Fremdsprache als Reflexionsgegenstand (11, davon 10 in Latein) und Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses, der mit neun von zehn Nennungen vor allem in der Wahrnehmung der (angehenden) Deutschlehrkräfte ein weiteres Merkmal guten Unterrichts darstellt. Wie die genannten, so weisen auch die übrigen fachgemeinsamen Kategorien, die Tab. 2 listet, klare Kristallisationspunkte in bestimmten Disziplinen auf (z. B. aktiver Gebrauch der Fremdsprache als Medium der Unterrichtskommunikation in Englisch). Sie tragen also trotz Berücksichtigung und Auftreten in mehreren Disziplinen eine gewisse fachspezifische Nuance.

Darüber hinaus enthält Tab. 2 auch Merkmale, welche von Lehrkräften in nur einem Fach (aber zum Teil mehrfach) genannt wurden und in diesem unter fachdidaktischer Perspektive als Gütekriterium gelten können. Hierzu zählen beispielsweise die Schaffung hoher Sprechanteile der Schüler:innen durch echte kommunikative Anlässe in Englisch (21 N., 12%), ein christlicher Gegenstandsbezug in Religion (10 N., 13%), die Förderung einer Faszination für das Fach in Latein (6 N., 3%) oder ein sinnvoller Einsatz von Klassenmusizieren in Musik (25 N., 6%). Häufiger tritt auch die Kategorie regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen (24 N., 3%) auf, die einem Schwerpunkt des Fachdiskurses entsprechend überwiegend in Latein als wichtig empfunden wird.

3.2 Relevanzeinschätzungen

3.2.1 Generische Dimensionen

In Abb. 1 sind die Relevanzeinschätzungen der (angehenden) Lehrkräfte aus dem zweiten Fragebogenabschnitt nach den sieben generischen Dimensionen aggregiert dargestellt. Diese liegen alle über dem theoretischen Skalenmittelwert, erfahren also erwartungsgemäß eher Zustimmung ($M \geq 2,89$). Zugleich sind jedoch prononcierte Unterschiede zwischen den Fächern erkennbar, die in einfaktorischen Varianzanalysen (auch bei einer Bonferroni-Holm-Adjustierung des Signifikanzniveaus) für alle Dimensionen außer der *Auswahl und Thematisierung von Inhalten und Fachmethoden* signifikant sind (Tab. 3). Die (angehenden) Lehrkräfte scheinen die Kriterien demnach mit einer gewissen fachbezogenen Differenziertheit zu bewerten, was auch auf Ebene der Subdimensionen in Tab. 4 sichtbar wird.

Eine *Auswahl und Thematisierung von Inhalten und Fachmethoden* spielt für (angehende) Lehrkräfte in Latein im Mittel die größte, in Religion die kleinste Rolle (Abb. 1), wobei hier die Einschätzungen bezüglich Strukturierung sowie fachlicher Akkuratheit und Korrektheit vergleichsweise hoch, dort etwas niedriger ausfallen (Tab. 4). In Religion ist zudem ein besonderer Fokus auf der motivierenden Einbettung der thematisierten Inhalte zu erkennen. Auch die *kognitive Aktivierung* wird durchschnittlich als relevant eingeschätzt, wobei alle vier zugehörigen Subdimensionen in Deutsch die höchsten, in Musik die niedrigsten Bewertungen erhalten (Abb. 1). Diese sind somit pro Fach konsistent, auch wenn die Beurteilung der Unterstützung des metakognitiven Lernens in Musik im Mittel relativ niedrig ausfällt ($M = 2,60$, $SD = 0,94$; Tab. 4). Die *Unterstützung des Übens* ist eine Unterrichtsdi-

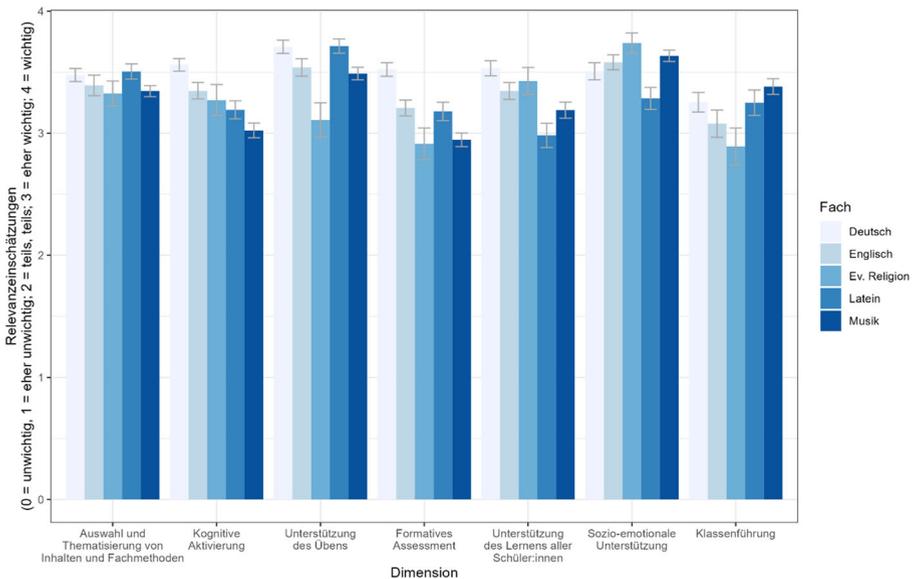


Abb. 1 Arithmetische Mittelwerte (inkl. Standardfehler) pro Dimension und Fach

Tab. 3 Einfaktorielle Welch-Varianzanalysen mit der abhängigen Variable Dimension und der unabhängigen Variable Fach

		$df_{\text{betw.}}$	df_{res}	F	p
I	Auswahl und Thematisierung von Inhalten und Fachmethoden	4	100,39	1,63	.173
II	Kognitive Aktivierung	4	101,12	12,08	< .001
III	Unterstützung des Übens	4	99,61	6,22	< .001
IV	Formatives Assessment	4	100,89	14,59	< .001
V	Unterstützung des Lernens aller Schüler:innen	4	103,93	6,97	< .001
VI	Sozio-emotionale Unterstützung	4	104,07	4,13	.011
VII	Klassenführung	4	99,96	2,99	.044

df Freiheitsgrade, F F -Statistik, p (nach Bonferroni-Holm adjustierter) p -Wert

Aufgrund der Verletzung der Varianzhomogenität bei einzelnen Dimensionen wurden einheitlich robuste Welch-ANOVAS genutzt.

mension, die in den Sprachfächern Deutsch und Latein besonders gewürdigt wird (Abb. 1). Dies gilt für einen konstruktiven Umgang mit Fehlern und Schwierigkeiten der Schüler:innen, aber auch für die Unterstützung bei der Festigung von Lern- und Anwendungsprozessen (Tab. 4). Im Vergleich der fünf Fächer beurteilen (angehende) Religionslehrkräfte die Bedeutung des Übens für ihren Unterricht geringer. Auf Dimensionsebene schneidet das *formative Assessment* bei den Relevanzeinschätzungen im Mittel am niedrigsten ab (Abb. 1), was mitunter durch geringere Zustimmungswerte zu einzelnen Facetten in Religion und Musik bedingt ist (Tab. 4). Die höchsten Werte erreicht es in Deutsch und zumindest bezüglich einer regelmäßigen Überprüfung des Verständnisses der Schüler:innen auch in Latein.

Der *Unterstützung des Lernens aller Schüler:innen* kommt nach Ansicht der Teilnehmenden im Fach Latein insgesamt (Abb. 1) wie auch konsequent in all seinen Facetten die geringste Bedeutung zu (Tab. 4). Umgekehrt vergeben hier (angehende) Deutsch- und in Bezug auf die Förderung der aktiven Mitwirkung aller Schüler:innen auch Religionslehrkräfte die höchsten Relevanzeinschätzungen. Von allen Dimensionen wird die sozio-emotionale Unterstützung am relevantesten bewertet ($M=3,54$, $SD=0,54$; Abb. 1), wobei insbesondere eine gute Beziehung zwischen Lehrkräften und Lernenden im Vordergrund steht ($M=3,81$, $SD=0,48$; Tab. 4). Die höchsten Bewertungen für beide Subdimensionen vergeben (angehende) Religionslehrkräfte. Hingegen scheint die Beziehung der Schüler:innen untereinander nach Einschätzung der Teilnehmenden in Latein weniger relevant zu sein ($M=2,84$, $SD=0,99$). Merkmale der Klassenführung werden schließlich in Musik und Latein als besonders wichtig erachtet (Abb. 1), und zwar in Bezug auf Verhaltens- wie auch Zeitmanagement. (Angehende) Religionslehrkräfte taxieren die Bedeutung beider Kriterien niedriger.

3.2.2 Fachgemeinsame und -spezifische Merkmale

Eine Übersicht zu den Relevanzeinschätzungen der Kriterien, die auf Basis des bildungswissenschaftlichen oder fachdidaktischen Forschungsstands formuliert und als

Tab. 4 Mittelwerte (*M*) und Standardabweichungen (*SD*) der Relevanzschatzungen pro Fach und (Sub-)Dimension

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
Gesamtanzahl der Antworten	288	67	51	23	56	91
<i>I</i> Auswahl und Thematisierung von Inhalten und Fachmethoden	3,41 (0,48)	3,48 (0,44)	3,39 (0,60)	3,33 (0,49)	3,51 (0,47)	3,35 (0,43)
<i>I.1</i> Auswahl von bedeutungsvollen, dem Lernstand angemessenen Inhalten, Realien sowie Fachmethoden	3,43 (0,75)	3,53 (0,68)	3,43 (0,81)	3,30 (0,70)	3,39 (0,85)	3,40 (0,72)
<i>I.2</i> Lebensweltbezogene, motivierende Einbettung von Inhalten sowie Fachmethoden	3,57 (0,68)	3,66 (0,59)	3,64 (0,60)	3,70 (0,56)	3,51 (0,88)	3,47 (0,66)
<i>I.3</i> Transparente und kohärente Strukturierung der im Unterricht thematisierten Inhalte sowie Fachmethoden	3,25 (0,83)	3,36 (0,85)	3,12 (0,91)	3,09 (0,90)	3,54 (0,66)	3,09 (0,78)
<i>I.4</i> Akkuratheit und Korrektheit der vermittelten Inhalte	3,40 (0,76)	3,34 (0,83)	3,40 (0,81)	3,22 (0,90)	3,59 (0,60)	3,38 (0,74)
<i>II</i> Kognitive Aktivierung	3,26 (0,56)	3,56 (0,42)	3,35 (0,48)	3,27 (0,61)	3,19 (0,56)	3,02 (0,58)
<i>II.1</i> Auswahl fachlich gehaltvoller und auf das kognitive Niveau der Schüler:innen abgestimmter Aufgaben	3,60 (0,61)	3,77 (0,52)	3,60 (0,53)	3,57 (0,73)	3,58 (0,63)	3,50 (0,66)
<i>II.2</i> Einsatz fachlich gehaltvoller Aufgaben, die multiple Denkweisen, Bearbeitungsansätze und Lösungswege anregen	3,04 (0,91)	3,45 (0,78)	3,02 (0,85)	2,87 (0,87)	2,91 (1,03)	2,86 (0,88)
<i>II.3</i> Unterstützung der kognitiven Aktivität der Schüler:innen (durch konstruktive Fehlerkultur, Explizierung von Denkprozessen etc.)	3,43 (0,75)	3,63 (0,63)	3,60 (0,70)	3,52 (0,59)	3,39 (0,76)	3,18 (0,84)
<i>II.4</i> Unterstützung des metakognitiven Lernens der Schüler:innen (z. B. durch Möglichkeiten zur Reflexion des eigenen Handelns und verwendeter Strategien)	2,97 (0,94)	3,42 (0,70)	3,08 (0,68)	3,13 (0,97)	2,88 (1,11)	2,60 (0,94)
<i>III</i> Unterstützung des Übens	3,56 (0,52)	3,71 (0,45)	3,54 (0,51)	3,11 (0,67)	3,71 (0,44)	3,49 (0,49)
<i>III.1</i> Unterstützung bei der Festigung von Lern- und Anwendungsprozessen	3,39 (0,72)	3,58 (0,65)	3,30 (0,74)	2,87 (0,81)	3,64 (0,62)	3,28 (0,71)
<i>III.2</i> Konstruktiver Umgang mit Fehlern und Schwierigkeiten der Schüler:innen beim Üben	3,73 (0,56)	3,84 (0,45)	3,80 (0,49)	3,35 (0,78)	3,79 (0,62)	3,69 (0,53)

Tab. 4 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
IV <i>Formatives Assessment</i>	3,17 (0,56)	3,52 (0,45)	3,21 (0,46)	2,91 (0,62)	3,18 (0,56)	2,95 (0,54)
IV.1 Klare Ausrichtung der Beurteilung auf die zu erlernenden Kompetenzen	2,79 (1,02)	3,17 (0,99)	2,96 (0,88)	2,22 (1,04)	2,79 (1,19)	2,58 (0,91)
IV.2 Regelmäßige Überprüfung des Verständnisses der Schüler:innen	3,41 (0,76)	3,58 (0,68)	3,44 (0,67)	3,17 (0,94)	3,59 (0,65)	3,21 (0,82)
IV.3 Qualitativ hochwertiges Feedback an die Schüler:innen	3,50 (0,69)	3,78 (0,45)	3,52 (0,61)	3,35 (0,71)	3,52 (0,66)	3,30 (0,82)
IV.4 Nutzung des Feedbacks als Grundlage für die Ausrichtung des weiteren Unterrichts	2,98 (0,96)	3,55 (0,72)	2,90 (0,87)	2,91 (1,08)	2,82 (1,06)	2,72 (0,90)
V <i>Unterstützung des Lernens aller Schüler:innen</i>	3,28 (0,62)	3,53 (0,51)	3,35 (0,49)	3,43 (0,53)	2,98 (0,74)	3,19 (0,63)
V.1 Bereitstellung eines Lernumfelds, welches verantwortungsvolle Autonomie und produktives Kompetenzerleben fördert	3,27 (0,86)	3,58 (0,68)	3,46 (0,68)	3,43 (0,73)	2,88 (1,08)	3,15 (0,82)
V.2 Individualisierung und Anpassung von unterrichtlichen Inhalten, Prozessen, Zielen und Erwartungshaltungen	3,10 (0,84)	3,46 (0,66)	3,08 (0,80)	3,09 (0,95)	2,82 (0,92)	3,01 (0,84)
V.3 Gestaltung und Förderung von Partizipationsmöglichkeiten und aktiver Mitwirkung aller Schüler:innen	3,47 (0,77)	3,55 (0,72)	3,49 (0,74)	3,77 (0,43)	3,27 (0,99)	3,44 (0,72)
VI <i>Sozio-emotionale Unterstützung</i>	3,54 (0,54)	3,51 (0,57)	3,58 (0,43)	3,74 (0,40)	3,29 (0,67)	3,63 (0,44)
VI.1 Beziehung zwischen Lehrperson und Schüler:innen (z. B. Wertschätzung, Gleichbehandlung)	3,81 (0,48)	3,73 (0,48)	3,88 (0,33)	3,87 (0,46)	3,73 (0,67)	3,87 (0,38)
VI.2 Beziehung der Schüler:innen untereinander (z. B. Hilfsbereitschaft, soziales Kohärenzempfinden)	3,26 (0,83)	3,28 (0,83)	3,27 (0,71)	3,61 (0,50)	2,84 (0,99)	3,40 (0,77)
VII <i>Klassenführung</i>	3,23 (0,71)	3,25 (0,65)	3,08 (0,80)	2,89 (0,72)	3,25 (0,78)	3,38 (0,61)
VII.1 Verhaltensmanagement (z. B. durch Prävention und Intervention von Störungen)	3,25 (0,80)	3,25 (0,75)	3,06 (0,91)	2,96 (0,64)	3,12 (0,94)	3,52 (0,66)
VII.2 Zeitmanagement (z. B. durch Maximierung der Lernzeit, klare Regeln und Routinen)	3,21 (0,91)	3,25 (0,84)	3,08 (0,91)	2,83 (1,03)	3,38 (1,00)	3,24 (0,84)

Tab. 4 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
	3,10 (0,43)	3,11 (0,46)	3,17 (0,34)	3,43 (0,28)	2,82 (0,47)	3,14 (0,36)
fig.1 Fachgemeinsames/-spezifisches	3,23 (0,88)	3,09 (0,87)	3,18 (0,75)	3,30 (0,82)	3,23 (1,08)	3,34 (0,82)
Didaktische Reduktion unter Berücksichtigung zukünftiger fachlicher Lernschritte						
fig.2 Kognitiv-ästhetische Aktivierung	2,95 (0,88)	2,88 (0,86)	2,78 (0,78)	2,82 (0,85)	2,73 (0,96)	3,28 (0,82)
fig.3 Fremdheitserfahrung und reflektierender Umgang mit (fremd-)kulturellen Schemata	3,07 (0,94)	3,09 (0,83)	3,20 (0,86)	3,48 (0,79)	3,00 (1,07)	2,93 (0,98)
fig.4 Aktiver Gebrauch der Fremdsprache als Medium der Unterrichtskommunikation	2,48 (1,55)	–	3,84 (0,47)	–	1,29 (1,12)	–
fig.5 Verwendung der Fremdsprache als Reflexionsgegenstand	2,97 (1,11)	–	3,18 (0,99)	–	2,78 (1,18)	–
fig.6 Förderung eines gegenstandsbezogenen dialogischen Unterrichtsdiskurses	3,10 (0,86)	3,22 (0,78)	3,35 (0,72)	3,57 (0,79)	2,83 (0,95)	2,90 (0,85)
fig.7 Adäquate Auswahl und Berücksichtigung bedeutsamer Rezeptionsdokumente	3,17 (0,83)	3,01 (0,84)	–	3,70 (0,56)	2,91 (0,90)	3,31 (0,76)
fig.8 Regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen	3,39 (0,84)	–	–	3,78 (0,52)	3,62 (0,78)	3,14 (0,87)
fig.9 Kognitiv-kommunikative Auseinandersetzung mit Fachinhalten	3,34 (0,75)	3,49 (0,69)	3,41 (0,73)	3,48 (0,73)	3,24 (0,88)	3,22 (0,70)
fig.10 Kognitiv-emotionale Betrachtung von Fachinhalten	2,96 (0,89)	2,94 (0,87)	2,78 (0,81)	–	2,93 (0,97)	3,09 (0,90)
fig.11 Förderung einer Faszination für das Fach	3,59 (0,72)	–	–	3,70 (0,63)	3,64 (0,67)	3,52 (0,77)
fig.12 Kognitiv-motorische Auseinandersetzung mit Fachinhalten	2,71 (1,03)	2,73 (0,88)	2,45 (1,02)	3,22 (0,74)	1,98 (1,18)	3,17 (0,79)
fig.13 Kritische Reflexion über die gegenwärtige (und zukünftige) kulturelle wie auch gesellschaftliche Relevanz des Fachs von Fachinhalten	2,67 (1,14)	–	–	–	2,84 (1,09)	2,57 (1,16)
fig.14 Zeitintensive, stark fokussierte und analytisch genaue Betrachtung von Fachinhalten	2,29 (0,95)	–	–	2,70 (1,02)	2,48 (1,04)	2,07 (0,82)
fig.15 Evozierung kognitiver Konflikte und Förderung entdeckenden Lernens	2,82 (0,89)	–	–	–	2,80 (1,03)	2,82 (0,79)
fig.16 Achten der Würde des Einzelnen	3,89 (0,43)	–	–	4,00 (0,00)	–	3,87 (0,48)

Tab. 4 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
de.1 Kognitiv-kreative Aktivierung	-	3,33 (0,77)	-	-	-	-
de.2 Berücksichtigung von (negativen) (Lebens-)Erfahrungen der Schüler:innen und entwicklungspsychologischer Aufgaben beim Einsatz und bei der Auswahl von Texten	-	3,18 (0,89)	-	-	-	-
de.3 Unterstützung der Schüler:innen bei emotionalen Erfahrungen mit Literatur	-	3,30 (0,70)	-	-	-	-
en.1 Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch	-	-	3,70 (0,51)	-	-	-
en.2 Schaffung hoher Sprechanteile der Schüler:innen durch echte kommunikative Anlässe	-	-	3,69 (0,58)	-	-	-
en.3 Anpassung der Verwendung der Fremdsprache seitens der Lehrkraft auf das Kompetenz- und Sprachniveau der Schüler:innen	-	-	3,18 (1,00)	-	-	-
en.4 Gelegenheiten zur Selbstkorrektur, um die eigene Sprachverwendung reflektieren zu können	-	-	3,08 (0,80)	-	-	-
en.5 Schaffung und Nutzung interdisziplinärer Bezüge zwischen (schulischen) Fremdsprachen als Lerngelegenheiten	-	-	2,66 (0,92)	-	-	-
en.6 Repetitives Üben zur Automatisierung von Redemitteln und intelligentes Üben unter Variation der Anwendungskontexte	-	-	3,04 (0,86)	-	-	-
en.7 Unterstützung bei emotionalen, (inter-)kulturellen (Fremdheits-)Erfahrungen	-	-	3,22 (0,77)	-	-	-

Tab. 4 (Fortsetzung)

	Insgesamt	Deutsch	Englisch	Ev. Religion	Latein	Musik
re.1	Offene Lernprozesse mit Texten der christlichen Tradition inszenieren	-	-	2,87 (0,97)	-	-
re.2	Interreligiöses Lernen	-	-	3,45 (0,60)	-	-
re.3	Förderung der Reflexionsfähigkeit der Schüler:innen auf das eigene religiöse (Er-)Leben und Handeln	-	-	3,70 (0,56)	-	-
re.4	Förderung der Reflexionsfähigkeit der Schüler:innen auf ethisches Handeln	-	-	3,70 (0,47)	-	-
re.5	Sensibilisierung der Schüler:innen im Hinblick auf Fehlformen/Schattenseiten von Religion	-	-	2,87 (0,81)	-	-
re.6	Kommunikation des Evangeliums	-	-	3,04 (0,88)	-	-
re.7	Raum für Seelsorge	-	-	3,00 (0,85)	-	-
re.8	Raum für Spiritualität	-	-	2,91 (0,90)	-	-
mu.1	Sinnvoller Einsatz von musikalischem Instrumentarium	-	-	-	-	3,59 (0,69)
mu.2	Raum- und Materialmanagement im Kontext von Klassenmusizieren	-	-	-	-	3,36 (0,80)
mu.3	Anbahnung von Tiefenverständnis mittels Musikpraxis	-	-	-	-	3,26 (0,86)

Skalierung: 0 unwichtig, ..., 4 wichtig; freie Tabellenzelle mit „-“: Aspekt in diesem Fach aufgrund inhaltlicher oder forschungspragmatischer Erwägungen nicht administriert. Trotz unabhängiger Erhebung wurden die Variablenabkürzungen der quantitativen Items nachträglich an die qualitativen Kategorienbezeichnungen (Tab. 2) angepasst, um – sofern möglich – Vergleiche zu erleichtern.

nicht generisch angenommen wurden (Abschn. 1.1, 1.2 und 2.2), bietet Tab. 4; im Folgenden wird nur auf einzelne Aspekte eingegangen, die in bestimmten Fächern auffallend viel oder wenig Zuspruch erfuhren. Von (angehenden) Deutschlehrkräften wurde die kognitiv-kommunikative Auseinandersetzung mit Fachinhalten als besonders relevant eingeschätzt ($M=3,49$, $SD=0,69$). Auch eine kognitiv-kreative Aktivierung ($M=3,33$, $SD=0,77$) bzw. eine Unterstützung der Schüler:innen bei emotionalen Erfahrungen mit Literatur ($M=3,30$, $SD=0,70$) erhielt größere Zustimmung, nicht aber eine kognitiv-motorische Auseinandersetzung mit Fachinhalten ($M=2,73$, $SD=0,88$). Diese ist auch aus Sicht von (angehenden) Englischlehrkräften von geringer Relevanz ($M=2,45$, $SD=1,02$; Tab. 4). Aber diese halten einen aktiven Gebrauch der Fremdsprache als Kommunikationsmedium ($M=3,84$, $SD=0,99$), eine Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch ($M=3,70$, $SD=0,51$) und eine Schaffung hoher Sprechanteile für sehr wichtig ($M=3,69$, $SD=0,58$). Gerade diese sprachproduzierenden Merkmale sind allerdings in Latein nach Einschätzung der Teilnehmenden irrelevant ($M=1,29$, $SD=1,12$); im Vordergrund stehen hier (Tab. 4) eine Förderung der Faszination für das Fach ($M=3,64$, $SD=0,67$), regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen ($M=3,62$, $SD=0,78$) und eine kognitiv-kommunikative Auseinandersetzung mit Fachinhalten ($M=3,24$, $SD=0,88$).

Für (angehende) Religionslehrkräfte besitzt das Achten der Würde des Einzelnen, d. h. die Vermeidung von Bloßstellung aufgrund religiöser Zugehörigkeit oder Ansichten im Unterricht univok oberste Priorität ($M=4,00$, $SD=0$), gefolgt von regelmäßigen Aktualitätsbezügen zur Lebenswelt der Schüler:innen ($M=3,78$, $SD=0,52$) und der Förderung einer Faszination für das Fach ($M=3,70$, $SD=0,56$), der Reflexionsfähigkeit der Schüler:innen mit Bezug auf das eigene religiöse (Er-)Leben und Handeln ($M=3,70$, $SD=0,56$) bzw. auf ethisches Handeln ($M=3,70$, $SD=0,47$; Tab. 4). Eine zeitintensive, stark fokussierte und analytisch genaue Betrachtung von Fachinhalten ist hier ($M=2,70$, $SD=1,02$) wie auch in Musik ($M=2,07$, $SD=0,82$) hingegen ein weniger relevanter Aspekt. Ihre Einschätzungen teilen (angehende) Religions- und Musiklehrkräfte auch darin (Tab. 4), dass letztgenannte ebenfalls die Achtung der Würde des Einzelnen, d. h. die Vermeidung von Bloßstellung beim (Vor-)Musizieren ($M=3,87$, $SD=0,48$) sowie die Förderung einer Faszination für das Fach ($M=3,52$, $SD=0,77$) hoch einschätzen. Zudem sprechen sie dem Musizieren und sinnvollen Einsatz von Musikinstrumenten einen hohen Stellenwert zu ($M=3,59$, $SD=0,69$).

4 Diskussion und Limitationen

4.1 Diskussion

Die vorliegende Studie knüpft an die in Bildungswissenschaft und Fachdidaktik aktuell geführte Diskussion zu Verallgemeinerbarkeit und Fachspezifik von Merkmalen und Modellen von Unterrichtsqualität an (z. B. Begrich et al. 2023; Charalambous und Praetorius 2018; Praetorius und Gräsel 2021) und bringt in diese Perspektiven aus den fachdidaktischen Diskursen sowie (angehender) Lehrkräfte

der Unterrichtsfächer Deutsch, Englisch, Ev. Religion, Latein und Musik ein. Damit folgt sie einem wiederholt formulierten Postulat, die theoretischen Annahmen und empirischen Befunde dieses von mathematisch-naturwissenschaftlichen Fächern dominierten Forschungsbereichs unter dem Blickwinkel anderer Fächer zu reflektieren und um (bislang unberücksichtigte) Konzepte und Spezifika der jeweiligen Fächer zu bereichern (z.B. Begrich et al. 2023; Reusser 2024; Rothland 2024). Hierzu bezieht die Studie neben den einschlägigen fachdidaktischen Forschungsdiskursen, die teils distal zur Unterrichtsqualitätsforschung stehen (Abschn. 1.2), die Praxissicht und -erfahrung von (angehenden) Lehrkräften ein und kommt damit der unter anderem von Freytag und Theurer (2020), Prediger et al. (2022) oder Rothgangel (2021) geäußerten Forderung nach, gerade bei einem derart praxisrelevanten Forschungsfeld den Unterrichtsalltag und dessen Protagonist:innen nicht unbeachtet zu lassen. Vielmehr können durch Einbezug von Praxiserfahrungen, Augenscheinvalidierungen und dialogische Aushandlung von Konzepten (gerade bei ohnehin defizitärem Forschungsstand) wesentliche Impulse und Innovationen ausgehen und den Forschungsdiskurs (z.B. wissenschaftliche Strukturierung und Abstrahierung subjektiver Annahmen und Theorien, Transfer von Modellen und Qualitätskriterien zwischen Fächern bzw. von der Wissenschaft in die Praxis) voranbringen.

Diesbezüglich stellt diese Onlinebefragung einen ersten explorativen Schritt dar. Ihr einheitliches Design mit offenen Antwortmöglichkeiten und Relevanzeinschätzungen zu generischen wie auch fachspezifischen Unterrichtsmerkmalen und die aufeinander abgestimmte methodische Vorgehensweise bei der Auswertung ermöglichen direkte Vergleiche und erleichtern die Identifikation von Gemeinsamkeiten und Unterschieden oder zumindest fachspezifischen Kristallisationspunkten. Während für detaillierte Erläuterungen zu Einzelbefunden oder -phänomenen (Abschn. 3) unter einer stärker fachbezogenen Perspektive, die im Rahmen dieses Beitrags nicht zu leisten sind, beispielsweise auf Hofrichter et al. (2024), Lindl et al. (2024a) oder Gutsmiedl et al. (2024) verwiesen sei, liegt der Fokus hier auf einer komparativen fachübergreifenden Synthese der Ergebnisse zu generischen und fachspezifischen Merkmalen guten Unterrichts.

Der generischen Dimension Auswahl und Thematisierung von Inhalten und Fachmethoden wird von (angehenden) Lehrkräften in allen Fächern eine gewisse Bedeutung zugesprochen. Gemäß der inhaltlichen Interpretation dieses Qualitätsaspekts von Ufer und Praetorius (2022) impliziert dies, dass ein sorgfältiges Unterrichtsarrangement als Indikator für dessen potenzielle Güte angesehen werden kann. Der kognitiven Aktivierung wird gerade in Deutsch erhöhte Aufmerksamkeit zuteil, was vielleicht weniger auf deren fachspezifische Relevanz als eher darauf hindeuten mag, dass Deutschlehrkräfte infolge wiederholt unerfreulicher PISA-Ergebnisse bereits häufiger mit derartigen Konzepten der Unterrichtsqualitätsentwicklung konfrontiert wurden (Schilcher et al. 2021) als Lehrkräfte der Fächer Latein, Religion oder Musik, die keinem vergleichbaren Evaluations- und Innovationsdruck unterliegen. Während Praetorius und Gräsel (2021) die Berücksichtigung einer Dimension Unterstützung des Übens noch begründungspflichtig erscheint, erachten (angehende) Lehrkräfte aller Fächer diesen Aspekt als relevant, wenn auch in den Sprachen Deutsch und Latein mehr als in Religion. Von allen zur Auswahl gestellten generischen Merkmalen erfährt das formative Assessment in den Antworten der Lehrkräfte die geringste

Beachtung und Wertschätzung. Dies ist gerade mit Blick auf die Tragweite dieser Dimension für eine lernwirksame Unterrichtsgestaltung (z. B. Hattie 2024; Hattie und Timperley 2007) kritisch zu sehen, deckt sich aber gewissermaßen mit den Ergebnissen des Deutschen Schulbarometers 2024 (Robert Bosch Stiftung 2024). Die ebenfalls darin als stärkste Faktoren für schulisches Wohlbefinden (Robert Bosch Stiftung 2024) reklamierten Aspekte Unterstützung des Lernens aller Schüler:innen und sozio-emotionale Unterstützung weisen in der vorliegenden Studie fachbezogene Ausprägungen und Unterschiede (v. a. Religion vs. Latein) auf, welche zum Teil die jeweiligen Fachkulturen und -traditionen widerspiegeln. Die Bedeutsamkeit der Klassenführung wird insbesondere von Musik – im Gegensatz zu Religionslehrkräften betont, was ähnlich wie in Sport (Herrmann und Gerlach, 2020) unter anderem auf ein anforderungsreiches Raum- und Materialmanagement zurückzuführen ist (Gutsmiedl et al. 2024; Puffer 2021).

Daneben wurden Freitextantworten und Relevanzeinschätzungen dazu genutzt, in der bildungswissenschaftlichen und fachdidaktischen Forschung bereits diskutierte fachspezifische Aspekte guten Unterrichts (z. B. ästhetische Aktivierung; Abschn. 1.1 und 1.2) aus Sicht von (angehenden) Lehrkräften zu beurteilen oder aus deren Antworten bislang wenig beachtete Kriterien abzuleiten. Hierbei erwies sich die Kombination aus beiden Zugängen als effektiv, da sie sich teils gegenseitig bestätigten, teils komplementär ergänzten und in jedem Fach wertvolle Anregungen zu berücksichtigungsfähigen fachspezifischen Merkmalen boten. So treten bei einer Zusammenschau der höchsten Nennungshäufigkeiten bzw. Beurteilungen in Englisch dieselben drei Merkmale hervor: aktiver Gebrauch der Fremdsprache als Medium der Unterrichtskommunikation, Kommunikation der Lehrkraft in sprachlich und fachlich korrektem Englisch und Schaffung hoher Sprechanteile der Schüler:innen durch echte kommunikative Anlässe. In Latein stimmen zwei der drei Aspekte überein (Förderung einer Faszination für das Fach, regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen), in Religion und Musik je einer (regelmäßige Aktualitätsbezüge zur Lebenswelt der Schüler:innen bzw. Musizieren und sinnvoller Einsatz von Musikinstrumenten). Deutsch weist keine augenfälligen Überschneidungen auf, jedoch könnte die absolute Nennungshäufigkeit eines gegenstandsbezogenen Unterrichtsdiskurses wie auch die tendenziell hohe Zustimmung zu einer kognitiv-kommunikativen Auseinandersetzung mit Fachinhalten als Hinweis auf eine wechselseitige Bestätigung angesehen werden. Außerdem wird hier wie auch in Musik beispielsweise auf das Potenzial ästhetisch-emotionalen Lernens verwiesen, auf das an anderer Stelle bereits ausführlicher eingegangen wurde (Lindl et al. 2024a). In Religion und Musik ist zudem der Aspekt der Achtung der Würde von Lernenden besonders relevant (Gutsmiedl et al. 2024), wenn auch jeweils in etwas anderer Ausprägung. In Latein liegt ein didaktischer Schwerpunkt eher auf der Sprachreflexion als auf der -produktion, welche in Englisch im Vordergrund steht (Hofrichter et al. 2024). Bemerkenswert sind schließlich auch die teils recht hohen Anteile an Antworten von (angehenden) Lehrkräften, die nicht auf unterrichtliche Prozessmerkmale bezogen waren (Tab. 2).

4.2 Limitationen und Ausblick

Die noch unter pandemischen Bedingungen rekrutierte Stichprobe weist hinsichtlich ihres Umfangs und ihrer Zusammensetzung nach bestimmten Kontextfaktoren (z. B. Fachzugehörigkeit, berufliche Stellung bzw. Erfahrung, Land der Teilnehmenden) Einschränkungen auf, aus denen in weiterführenden Analysen jedoch keine systematischen und substanziellen Unterschiede im Antwortverhalten ersichtlich werden (z. B. dt. vs. nicht-dt. Lehrkräfte, Lehramtsstudierende vs. Lehrkräfte). Obwohl die hier fokussierten fünf Schulfächer im bildungswissenschaftlichen Diskurs zu Unterrichtsqualität bislang eher wenig beachtet wurden (Abschn. 1.1 und 1.2), ist deren Auswahl nur durch pragmatische Gründe bedingt und wäre eine konzertierte Untersuchung des ganzen schulischen Fächerspektrums unter einem einheitlichen Rahmendesign zu begrüßen, um generalisierbare Aussagen zu erhalten. Dies ist zumindest in Ansätzen durch die theoretische Anbindung an das erweiterte Syntheseframework (Praetorius et al. 2020a, b) möglich, an dessen Konzeption bereits die Domänen Mathematik, Naturwissenschaften, Geschichte und Sport beteiligt waren.

Eine weitere Herausforderung stellten die eindeutige Kodierung und objektive Interpretation der oftmals schlagwortartig, kurz und prägnant formulierten Angaben in den Freitextfeldern dar. Hierzu verständigten sich alle Kodierenden mehrfach über die Kategorien des Auswertungsrasters, wurde jede offene Antwort mindestens zweifach kodiert (Tab. 1) und wurden uneindeutige Fälle bis zur Konsensfindung diskutiert. Zudem ist bei der Interpretation der Ergebnisse zu den Freitextantworten zu bedenken, dass nur drei Merkmale guten Fachunterrichts benannt werden sollten; die Angaben repräsentieren also immer eine individuell priorisierte Selektion aus einer Vielzahl weiterer möglicher Aspekte. Die verhältnismäßig geringe Nennungshäufigkeit generischer Aspekte ist außerdem auch damit erklärbar, dass in der Instruktion ein klarer Fachbezug hergestellt wurde. Die doppelte Abfrage von Merkmalen guten Unterrichts im Freitextformat vor und nach Präsentation der Relevanzeinschätzungen erlaubte es aber, einerseits spontan, intuitiv und unbeeinflusst auf eigener Erfahrung fußende Ideen einzubringen sowie andererseits vor dem gedanklichen Hintergrund der soeben bewerteten Kriterien noch ungenannte, zusätzliche oder konkretisierende Merkmale anzuführen.

Wie bei jedem Fragebogeninstrument, das testökonomischen Restriktionen unterliegt, ist auch bei diesem nicht mit letzter Gewissheit sicherzustellen, dass alle Teilnehmenden sämtliche Instruktionen und Items identisch und im Sinne der Urheber:innen verstehen. Hierfür erforderliche ausführliche Erläuterungen würden einerseits den Umfang sowie zeitlichen Bearbeitungsaufwand des Fragebogens erhöhen sowie dessen Akzeptanz und die Teilnahmebereitschaft daran schmälern. Andererseits bestünde die Gefahr, dass Hinweise und notwendigerweise fachlich orientierte Konkretisierungen bereits in bestimmte Denkrichtungen lenken und so je nach disziplinärem Standpunkt der teilnehmenden Person eher zur Annahme bzw. Ablehnung des Kriteriums führen könnten. Dies würde die interfachliche Vergleichbarkeit der Items reduzieren. Zudem wäre selbst durch umfangreiche Kommentierung bei (angehenden) Lehrkräften, welche mit aktuellen Entwicklungen der Bildungsforschung eher wenig vertraut sind, ein nicht nur fachbezogen, sondern sogar interindividuell einheitliches Verständnis von Konzepten wie beispielsweise kognitiver, kognitiv-

emotionaler (z. B. Hesse und Winkler 2022), kognitiv-ästhetischer (z. B. Björk et al. 2024; Kranefeld 2021; Lindl et al. 2024a), kognitiv-kreativer (z. B. Haas 2021; Praetorius und Gräsel 2021; Rakoczy et al. 2021) oder kognitiv-kommunikativer Aktivierung (z. B. Kranefeld 2021; Praetorius und Gräsel 2021; Wilden et al. 2024) nicht zu erwarten, die selbst auf wissenschaftlicher Ebene noch unterbestimmt sind und teils sophistisch diskutiert werden (z. B. Praetorius et al. 2024; Praetorius und Gräsel 2021; Schreyer 2024).

Eine Einführung in diesen speziellen Diskurs war jedoch weder Ziel dieser Onlinebefragung noch zwingende Voraussetzung hierfür. Vielmehr war von Interesse, ob die Teilnehmenden den präsentierten Begriffen und Konzepten in intuitiver Interpretationsfreiheit und aus ihrer subjektiven, eher praxisorientierten Vorstellungs- und Erfahrungswelt fachbezogene Bedeutung beimessen können. Dies scheint ihnen erstens mit Blick auf die interdisziplinäre Varianz in den mittleren Relevanzeinschätzungen pro Item (wie auch zwischen den Items) und mindestens für sechs der sieben Dimensionen zu gelingen, da sich hier trotz aller potenziellen interindividuellen Verständnisunterschiede, die sich bezogen auf die Gesamtstichprobe gegebenenfalls ausmitteln, gewisse Facheigenheiten und -strukturen abzeichnen (Abb. 1 und Tab. 2). Zweitens lässt sich anhand von Items, die eher dem didaktischen Diskurs einer bestimmten Disziplin entstammen (z. B. Englisch: aktiver Gebrauch der Fremdsprache als Medium der Unterrichtskommunikation; z. B. Helmke und Klieme 2008; Wilden 2021), aber prinzipiell auch für andere Fächer bedeutsam sein könnten (z. B. Latein: *Latine loqui*; Böttcher 2018) und daher in mehreren Disziplinen abgefragt wurden, auf eine fachangemessene Auslegung und differenzielle Bewertung schließen. Drittens geht aus der schriftlichen Rückmeldung einer Lateinlehrkraft zum Onlinefragebogen, „die Umfrage zeigt(e) leider auch für den Laien sehr deutlich die sog. Kontrollfragen“ (L_077), indirekt hervor, dass gerade bei dem zuletzt beschriebenen Itemtyp Sinnrekonstruktion und Einschätzung nicht willkürlich erfolgten, sondern entsprechende Unterrichtsmerkmale dezidiert als nicht facheigen wahrgenommen wurden. Vor diesem Hintergrund sind die Befunde auf konzeptueller Ebene durchaus anschlussfähig und liefern aussagekräftige Indizien aus Lehrkraftsicht für die Relevanz bestimmter Aspekte von Lehr-Lernprozessen in den jeweiligen Fächern.

Somit leistet diese Studie einen kleinen Beitrag zum Diskurs über die Generik und Fachspezifik von Merkmalen guten Unterrichts, und zwar insbesondere aus dem Blickwinkel von Fächern, die diesbezüglich in der Forschung eher wenig Beachtung erfuhren. Dementsprechend sind die in eben diesen Disziplinen aus Freitextantworten extrahierten und in Relevanzeinschätzungen als besonders wichtig oder unwichtig eingeschätzten Kriterien von gesteigertem Interesse und sollten von der weiteren bildungswissenschaftlichen wie auch fachdidaktischen Forschung aufgegriffen, im Einzelnen untersucht, in bestehenden Modellen (additiv oder integrativ) verortet und diskutiert werden. Des Weiteren ist fraglich und für die fünf hier betrachteten Schulfächer erstmalig zu überprüfen, ob und welche der identifizierten Kriterien guten Unterrichts sich empirisch als lernförderlich erweisen und damit nach Berliners (2005) Definition Merkmale qualitätvollen Unterrichts sind. Während dies einen wichtigen Validierungsschritt darstellt, um nicht vorschnell konkrete Handlungsempfehlungen abzuleiten, deuten einige Ergebnisse der Onlinebefragung jedoch darauf hin, dass der seit über 20 Jahren evidenzbasiert geführte, äußerst praxisrelevante wis-

senschaftliche Diskurs über die Wirksamkeit von Oberflächen- und Tiefenstrukturen von Unterricht (angehende) Lehrkräfte zum Teil noch zu wenig erreicht hat. Daraus ergibt sich die Frage, wie es fachdidaktischer und bildungswissenschaftlicher Unterrichtsforschung künftig gelingen kann, ihre Erkenntnisse in die Aus-, Weiter- und Fortbildung zu implementieren und in einem kokonstruktiven Dialog mit der Praxis weiterzuentwickeln, um die Qualität von Unterricht zu sichern und zu optimieren (GFD und ÖGFD 2024). Ein erster Ansatzpunkt hierfür könnte beispielsweise sein, dass (angehende) Lehrkräfte vonseiten der Schulinspektionen und Landesinstitute bereits frühzeitig in ihrer beruflichen Laufbahn und in regelmäßigen Zeitabständen zu einer kritischen (Eigen-, Fremd- oder kollegialen) Evaluation und Reflexion ihrer pädagogischen Praxis angehalten und ihnen entsprechende Instrumente an die Hand gegeben werden (z. B. amadeus.falko-pv.de; argev.ch/insula; feedbackschule.com; qualitepe.phzh.ch; unterrichtsdiagnostik.de). Denn aus der Expertiseforschung zu Lehrkräften sind positive Zusammenhänge zwischen einer reflektierten Praxis und der Leistungsentwicklung von Schüler:innen bekannt (Anderson und Taner 2023; Schön 1983).

Förderung 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 Autoren verantwortlich.

Funding Open Access funding enabled and organized by Projekt DEAL.

Interessenkonflikt A. Lindl, L. Simböck, M. Rader, M. Hofrichter, M. Gutsmiel, S. Böhringer und P. Ehrlich bestätigen, dass sie keine Verbindungen zu oder Beteiligung an Organisationen oder Einrichtungen haben, die ein finanzielles oder nicht finanzielles Interesse an den in dieser Veröffentlichung behandelten Themen oder Gegenständen haben.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Anderson, J., & Taner, G. (2023). Building the expert teacher prototype: A metasummary of teacher expertise studies in primary and secondary education. *Educational Research Review*, 38, 100485. <https://doi.org/10.1016/j.edurev.2022.100485>.
- Begrich, L., Praetorius, A.-K., Decristan, J., Fauth, B., Göllner, R., Herrmann, C., Kleinknecht, M., Taut, S., & Kunter, M. (2023). Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft. *Unterrichtswissenschaft*, 51(1), 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>.

- Björk, C., Hofmann, B., Knigge, J., & Puffer, G. (2024). Towards modelling instructional quality for music classrooms: Exploring subject-specific application of the TALIS/GTI generic framework. *Beiträge empirischer Musikpädagogik*, 16(2), 1–17. <https://doi.org/10.62563/bem.v16i2.246>.
- 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>.
- Böttcher, E. (2018). *Lateinisch sprechen im Unterricht. Praktische Ansätze des „Latine loqui“*. Vandenhoeck & Ruprecht.
- 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>.
- 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 1), 102–116. <https://doi.org/10.25656/01:25867>.
- Drechsel, B., & Schindler, A. K. (2019). Unterrichtsqualität. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 353–372). 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>.
- Eckerth, M., Hanke, P., & Hein, A. K. (2012). Schulische Bedingungen des Lehrens und Lernens im Anfangsunterricht der Grundschule – Ausgewählte Ergebnisse aus dem FiS-Projekt. In F. Hellmich, S. Förster & F. Hoya (Hrsg.), *Bedingungen des Lehrens und Lernens in der Grundschule. Bilanz und Perspektiven* (S. 65–68). Springer VS.
- Ehrich, P. (2025). *Zielvorstellungen in Bezug auf Klassensingen: Eine Mixed-Methods-Studie*. Universität Augsburg. <https://nbn-resolving.org/urn:nbn:de:bvb:384-opus4-1214298>.
- Findeisen, S. (2017). *Fachdidaktische Kompetenzen angehender Lehrpersonen. Eine Untersuchung zum Erklären im Rechnungswesen*. Springer. <https://doi.org/10.1007/978-3-658-18390-5>.
- 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). Springer. https://doi.org/10.1007/978-3-658-30602-1_3.
- Gabriel, K. (2014). *Videobasierte Erfassung von Unterrichtsqualität im Anfangsunterricht der Grundschule. Klassenführung und Unterrichtsklima in Deutsch und Mathematik*. Kassel Univ. Press.
- GFD/ÖGFD (Gesellschaft für Fachdidaktik/Österreichische Gesellschaft für Fachdidaktik) (2024). Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung. Gemeinsames Positionspapier der Gesellschaft für Fachdidaktik (GFD) und Österreichischen Gesellschaft für Fachdidaktik (ÖGFD). In M. Hemmer, C. Angele, C. Bertsch, S. Kapelari, G. Leitner & M. Rothgangel (Hrsg.), *Fachdidaktik im Zentrum von Forschungstransfer und Transferforschung* (S. 11–21). Waxmann.
- Gutsmiedl, M., Ehrich, P., & Lindl, A. (2024). Merkmale guten Musikunterrichts – ein Validierungsansatz aus Lehrkraftsicht. In B. Clausen, J. Ehninger & M. Sachsse (Hrsg.), *45. Jahresband des Arbeitskreises Musikpädagogische Forschung* (S. 29–43). Waxmann. <https://doi.org/10.31244/9783830999188.03>.
- Guttko, J. (2023). Kognitive Aktivierung im Fremdsprachenunterricht. Ein systematisches Review von Forschungsarbeiten aus dem deutschsprachigen Raum. *Zeitschrift für Fremdsprachenforschung*, 32(2), 145–176.
- Haas, G. (2021). *Handlungs- und produktionsorientierter Literaturunterricht: Theorie und Praxis eines „anderen“ Literaturunterrichts für die Primar und Sekundarstufe*. Klett.
- Hattie, J. (2024). *Visible learning 2.0*. Schneider Hohengehren.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Heinitz, B., & Nehring, A. (2020). Kriterien naturwissenschaftsdidaktischer Unterrichtsqualität – ein systematisches Review videobasierter Unterrichtsforschung. *Unterrichtswissenschaft*, 48(3), 319–360. <https://doi.org/10.1007/s42010-020-00074-8>.
- Helmke, A. (2022). *Unterrichtsqualität und Professionalisierung. Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Klett.

- 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). Beltz. <https://doi.org/10.25656/01:3149>.
- Herrmann, C., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport – Ein Überblicksbeitrag zum Forschungsstand in Theorie und Empirie. *Unterrichtswissenschaft*, 48(3), 361–384. <https://doi.org/10.1007/s42010-020-00080-w>.
- Hesse, F., & Winkler, I. (2022). Fachliche Qualität im Literaturunterricht. *Zeitschrift für Sprachlich-Literarisches Lernen und Deutschdidaktik*, 2, 1–29. <https://doi.org/10.46586/SLLD.Z.2022.9605>.
- Hofrichter, M., Lindl, A., Rader, M., & Krämer, A. (2024). Was macht guten Fremdsprachenunterricht aus? Theoretische Überlegungen und erste Ergebnisse einer interdisziplinären Onlinebefragung in den Fächern Englisch und Latein. In V. Lohe, A. Lindl & P. Kirchhoff (Hrsg.), *Unterrichtsqualität in schulischen Fremdsprachen. Theoretische Ansätze und empirische Ergebnisse aus den Fachdidaktiken* (S. 109–137). Waxmann. <https://doi.org/10.31244/9783830999201.05>.
- Kleickmann, T., Steffensky, M., & Praetorius, A.-K. (2020). Quality of teaching in science education. More than three basic dimensions? *Zeitschrift für Pädagogik*, 66(Beiheft 1), 37–55. <https://doi.org/10.25656/01:25862>.
- 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). BMBF.
- 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., 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(2), 311–327. <https://doi.org/10.1007/s11858-020-01151-z>.
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Schöningh. <https://doi.org/10.36198/9783838538952>.
- 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>.
- Langer, W., Scheuer, C., Schnitzler, C., Adamakis, M., Borghouts, L., Costa, J., Crapa, A., Herrmann, C., Janíková, M., Jurak, G., Kohake, K., Lefèvre, L., Lemling, A., Lundvall, S., Masarykova, D., O'Brien, W., Ries, F., Slingerland, M., & Gerlach, E. (In Vorbereitung). *Quality of Teaching in Physical Education: A European Delphi Study*.
- Lindl, A. (2024). Lateinunterricht (*ars docendi*) im Spiegel der Wissenschaft – aktuelle Forschungsperspektiven aus dem Vorhaben FALCO-PV. *Ars docendi*, 18, 1–17.
- Lindl, A., & Krämer, A. (2022). Was macht guten Lateinunterricht aus? Eine Delphi-Studie zu Unterrichtsqualität im Fach Latein. *Die Alien Sprachen im Unterricht*, 69(1), 9–15.
- Lindl, A., Gaier, L., Weich, M., Frei, M., Ehra, C., Gastl-Pischetsrieder, M., Elmer, M., Asen-Molz, K., Ruck, A.-M., Heinze, J., Murmann, R., Gunga, E., & Röhl, S. (2019). Eine ‚gute‘ Erklärung für alle?! Gruppenspezifische Unterschiede in der Beurteilung von Erklärqualität – erste Ergebnisse aus dem interdisziplinären Forschungsprojekt FALKE. In T. Ehmke, P. Kuhl & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten. Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 128–141). Beltz Juventa.
- Lindl, A., Ehrich, P., Gutmiedl, M., Rader, M., Simböck, L., Gürtner, M., Böhringer, S., Krämer, A., Kirchhoff, J., & Frei, M. (2024a). 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). Waxmann.
- Lindl, A., Lohe, V., & Kirchhoff, P. (2024b). Fachspezifische Unterrichtsqualität in schulischen Fremdsprachen – Grundlagen und Forschungsansätze. In V. Lohe, A. Lindl & P. Kirchhoff (Hrsg.), *Unterrichtsqualität in schulischen Fremdsprachen* (S. 9–37). Waxmann. <https://doi.org/10.31244/9783830999201.01>.
- Lindl, A., Simböck, L., Rader, M., Hofrichter, M., Gutmiedl, M., Böhringer, S., & Ehrich, P. (2025a). *Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften – Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik. Appendix I: Syntheseframework als Kategoriensystem*. OSF. <https://doi.org/10.17605/OSF.IO/S6BQ5>.

- Lindl, A., Simböck, L., Rader, M., Hofrichter, M., Gutmiedl, M., Böhringer, S., & Ehrich, P. (2025b). *Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften – Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik. Appendix II: LimeSurvey-Fragebogen „Unterrichtsqualität im Fach“*. OSF. <https://doi.org/10.17605/OSF.IO/S6BQ5>.
- Lindl, A., Simböck, L., Rader, M., Hofrichter, M., Gutmiedl, M., Böhringer, S., & Ehrich, P. (2025c). *Merkmale guten Unterrichts aus Sicht von (angehenden) Lehrkräften – Ergebnisse einer Onlinebefragung in den fünf Unterrichtsfächern Deutsch, Englisch, Evangelische Religion, Latein und Musik. Appendix III: Analysecode und -daten*. OSF. <https://doi.org/10.17605/OSF.IO/S6BQ5>.
- 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). Waxmann.
- 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). Springer VS. <https://doi.org/10.1007/978-3-531-94023-6>.
- Mu, J., Bayrak, A., & Ufer, S. (2022). Conceptualizing and measuring instructional quality in mathematics education: A systematic literature review. *Frontiers in Education*, 7, 994739. <https://doi.org/10.3389/educ.2022.994739>.
- Neuhaus, B. J. (2021). Unterrichtsqualität aus der Perspektive der Biologiedidaktik. *Unterrichtswissenschaft*, 49(2), 273–283. <https://doi.org/10.1007/s42010-021-00114-x>.
- Nonte, S., Willems, A. S., & Stubbe, T. C. (2024). Analysen zu Oberflächen- und Tiefenmerkmalen des Musikunterrichts an Gymnasien und Gesamtschulen in musikalischen Profilklassen sowie in Klassen ohne Profil. *Beiträge empirischer Musikpädagogik*, 15, 1–17. <https://doi.org/10.62563/bem.v2024242>.
- 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., 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., Wemmer-Rogh, W., Schreyer, P., & Brinkmann, M. (Hrsg.). (2024). *Kognitive Aktivierung unter der Lupe: Bestandsaufnahme und Möglichkeiten der Weiterentwicklung eines prominenten Konstrukts*. Waxmann.
- Praetorius, A.-K., Vieluf, S., Saß, S., Bernholt, A., & Klieme, E. (2016). The same in German as in English? Investigating the subject-specificity of teaching quality. *Zeitschrift für Erziehungswissenschaft*, 19, 191–209. <https://doi.org/10.1007/s11618-015-0660-4>.
- 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., 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>.
- Prediger, S., Götze, D., Holzäpfel, L., Rösken-Winter, B., & Selter, C. (2022). Five principles for high-quality mathematics teaching: Combining normative, epistemological, empirical, and pragmatic perspectives for specifying the content of professional development. *Frontiers in Education*, 7, 969212. <https://doi.org/10.3389/educ.2022.969212>.
- Puffer, G. (2021). Professionelle Kompetenzen von Musiklehrkräften: Grundzüge einer Modellierung. *Beiträge empirischer Musikpädagogik*, 12, 1–71. <https://doi.org/10.62563/bem.v2021207>.
- Puffer, G., & Hofmann, B. (2017). FALKO-M. Entwicklung und Validierung eines Testinstruments zum domänenspezifischen Professionswissen von Musiklehrkräften. 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. 245–289). Waxmann.
- R Core Team (2023). *R. A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rader, M., Lindl, A., Ehrich, P., & Schilcher, A. (In Vorbereitung). *Unterrichtsqualität im Fach Deutsch – ein systematisches Review*.

- Rakoczy, K., Wagner, E., & Frick, U. (2021). Wie in Mathe so auch in Kunst? Zur Konzeption von Unterrichtsqualität im Kunstunterricht. *Unterrichtswissenschaft*, 49(2), 235–241. <https://doi.org/10.1007/s42010-021-00104-z>.
- Reinhardt, V., Rehm, M., & Wilhelm, M. (Hrsg.). (2021). *Wirksamer Fachunterricht. Eine metaanalytische Betrachtung von Expertisen aus 17 Schulfächern*. Schneider Hohengehren.
- Reusser, K. (2024). Empirische Unterrichtsforschung – quo vadis? *Unterrichtswissenschaft*, 52(2), 167–182. <https://doi.org/10.1007/s42010-024-00209-1>.
- 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). Klett.
- Robert Bosch Stiftung (2024). *Deutsches Schulbarometer: Befragung Schüler:innen. Ergebnisse von 8- bis 17-Jährigen und ihren Erziehungsberechtigten zu Wohlbefinden, Unterrichtsqualität und Hilfsverhalten*. Robert Bosch Stiftung.
- 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 in-service and pre-service primary EFL teachers. National and international research* (S. 197–218). Waxmann.
- 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*, 14(1), 41–53. <https://doi.org/10.1007/s35834-024-00411-2>.
- Schilcher, A., Lindl, A., Krauss, S., & Hilbert, S. (Hrsg.). (In Vorbereitung). *FALKE – Fachspezifische Lehrkraftkompetenzen im Erklären: Untersuchungen zur Beurteilung und zu Kriterien unterrichtlicher Erklärqualität aus der Perspektive von 13 Fachbereichen*. Beltz Juventa.
- 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., Memminger, J., Rank, A., Tepner, O., & Thim-Mabrey, C. (2021). FALKE. Experiences from transdisciplinary educational research by fourteen disciplines. *Frontiers in Education*, 5, 579982. <https://doi.org/10.3389/educ.2020.579982>.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. Temple Smith.
- Schreiner, C., Wiesner, C., Kiefer, T., Helm, C., Ivanova, M., Kemethofer, D., Illetschko, M., Freller-Töglhofer, M., & Paasch, D. (2019). Merkmale des fachlichen Unterrichts und Schülerkompetenzen. In A. C. George, C. Schreiner, C. Wiesner, M. Pointinger & K. Pacher (Hrsg.), *Fünf Jahre flächendeckende Bildungsstandardüberprüfungen in Österreich. Vertiefende Analysen zum Zyklus 2012 bis 2016* (S. 115–136). Waxmann.
- Schreyer, P. (2024). *Kognitive Aktivierung in der Unterrichtsinteraktion. Eine qualitativ-rekonstruktive Analyse zu Passungsverhältnissen im Mathematikunterricht*. Waxmann. <https://doi.org/10.25656/01:29151>.
- Schweitzer, F. (2024). *Lernen im Religionsunterricht. Was der RU leisten kann und wie er seine Ziele erreicht*. Vandenhoeck & Ruprecht. <https://doi.org/10.36198/9783838560885>.
- 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>.
- 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). 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 1), 63–80. <https://doi.org/10.3262/ZPB2001063>.
- Wemmer-Rogh, W., Grob, U., Charalambous, C. Y., & Praetorius, A.-K. (2024). Measurement invariance between subjects: what can we learn about subject-related differences in teaching quality? *ZDM – Mathematics Education*, 56(5), 831–844. <https://doi.org/10.1007/s11858-024-01622-7>.
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.
- 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., Guttke, J., & Wellmanns, L. (2024). Das Potential zur kommunikativ-kognitiven Aktivierung im Englischunterricht – Befunde einer Analyse von Lernaufgaben aus Lehrwerken der

- Grundschule. *Fremdsprachen Lehren und Lernen*, 53(2), 120–139. <https://doi.org/10.24053/FLuL-2024-0026>.
- 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>.

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.