Zeitschrift für Erziehungswissenschaft (2026) 29:813–841
https://doi.org/10.1007/s11618-026-01380-0

ALLGEMEINER TEIL

Unterrichtsqualität und Schülerleistungen in der Grundschule: Effekte des Einsatzes schulartfremder, fachfremder und nicht-grundständig ausgebildeter Lehrkräfte

Eric Richter · Sofie Henschel · Sebastian Weirich · Stefan Schipolowski · Lars Hoffmann · Dirk Richter

Eingegangen: 18. Oktober 2024 / Überarbeitet: 7. Oktober 2025 / Angenommen: 5. Dezember 2025 / Online publiziert: 18. Februar 2026
© The Author(s) 2026

Zusammenfassung Um die Unterrichtsversorgung zu sichern, setzen viele Grundschulen in Deutschland Lehrkräfte fach- und schulartfremd oder ohne grundständige Lehramtsausbildung im Unterricht ein. Der vorliegende Beitrag untersucht die Auswirkungen dieses Einsatzes auf die Unterrichtsqualität und die Leistungen der Schüler:innen in der Jahrgangsstufe 4. Auf Basis der Daten des IQB-Bildungstrends 2021 erfolgte basierend auf Daten von $N = 26.844$ Viertklässler:innen und $N = 1738$ Lehrkräften zunächst eine Aufbereitung mittels Multilevel Propensity Score Matching. Anschließend wurden Multilevel-Regressionsanalysen durchgeführt, um die Effekte unterschiedlicher Einsätze von Lehrkräften jeweils im Vergleich zu regulär eingesetzten Lehrkräften unter Kontrolle von Merkmalen der Schüler:innen, der Klassenzusammensetzung und der Lehrkraft zu analysieren. Die Ergebnisse zeigen, dass sich die wahrgenommene Unterrichtsqualität durch die Schüler:innen in den Fächern Deutsch und Mathematik in Klassen mit fachfremden, schulartfremden oder

✉ Prof. Dr. Eric Richter
Katholische Universität Eichstätt-Ingolstadt, Ostenstr. 26, 85072 Eichstätt, Deutschland
E-Mail: eric.richter@ku.de

Dr. Sofie Henschel · Dr. Sebastian Weirich · Dr. Stefan Schipolowski
Institut zur Qualitätsentwicklung im Bildungswesen, Luisenstraße 56, 10117 Berlin, Deutschland

Dr. Sofie Henschel
E-Mail: sofie.henschel@iqb.hu-berlin.de

Dr. Sebastian Weirich
E-Mail: sebastian.weirich@iqb.hu-berlin.de

Dr. Stefan Schipolowski
E-Mail: stefan.schipolowski@iqb.hu-berlin.de

Prof. Dr. Dirk Richter
Professur Erziehungswissenschaftliche Bildungsforschung, Universität Potsdam, Karl-Liebknecht-Straße 24–25, 14476 Potsdam, Deutschland
E-Mail: dirk.richter@uni-potsdam.de

nicht grundständig ausgebildeten Lehrkräften gegenüber regulär eingesetzten Lehrkräften nicht unterscheidet. Hinsichtlich der Leistungen der Schüler:innen zeigen sich nur im Fach Deutsch und nur für die von schulartfremden Lehrkräften unterrichteten Schüler:innen etwas geringere Kompetenzen. Der Beitrag schließt mit einer Diskussion der Implikationen für die Bildungsadministration und einem Ausblick auf weitere Forschungsbedarfe.

Schlüsselwörter Lehrkräftemangel · Seiten- und Quereinstieg · Unterrichtsqualität · Schüler:innenleistungen · Propensity Score Matching

Instructional quality and student achievement in primary school: Effects of assigning cross-school, out-of-field, and non-traditionally trained teachers

Abstract To ensure adequate staffing in schools, many primary schools in Germany employ teachers who are either teaching subjects outside their field of expertise, working outside their trained school track, or who lack traditional teacher training in the subjects of German and mathematics. This paper examines the impact of such staffing on teaching quality and student performance in the 4th grade. Based on data from the IQB Trends in Student Achievement 2021, a Multilevel Propensity Score Matching approach is first applied using data from $N = 26{,}844$ fourth-grade students and $N = 1738$ teachers. Subsequently, multilevel regression analyses are conducted to estimate the effects of these different teacher assignments compared to regularly assigned teachers, controlling for student characteristics, class composition, and teacher characteristics. The results indicate that, from the students' perspective, there are no significant differences in teaching quality for German and mathematics between subject-out-of-field, school-out-of-field, or non-traditionally trained teachers compared to regularly assigned teachers. Regarding student performance, small negative effects are observed only in German and only for students taught by school-out-of-field teachers. The implications for educational administration and further research are discussed.

Keywords Teacher shortage · Lateral entry · Instructional quality · Student performance · Propensity Score Matching

1 Einleitung

Der Lehrkräftemangel in Deutschland ist seit Jahren akut, insbesondere an Grundschulen und in Ostdeutschland (Autorengruppe Bildungsberichterstattung 2022; Klemm 2022a). Der gestiegenen Lehrkräftenachfrage steht dabei eine unzureichende Zahl an Lehramtsabsolvent:innen gegenüber (KMK 2025). Um diesem Mangel zu begegnen, beschloss die Ständige Konferenz der Kultusminister der Länder (KMK) im Jahr 2013 sogenannte Sondermaßnahmen, die einen flexibleren Einsatz von Lehrkräften ermöglichen. Diese Maßnahmen erlauben es, auch

nicht grundständig ausgebildete Personen für den Schuldienst zu gewinnen (KMK 2013). Darüber hinaus gestatten viele Bundesländer den fachfremden Einsatz von Lehrkräften, etwa auf Grundlage landesspezifischer Regelungen (z. B. Nordrhein-Westfalen: MSW 2012, § 12). In der Praxis wird zunehmend auch von schulart-übergreifenden Einsätzen Gebrauch gemacht, bei denen Lehrkräfte außerhalb ihrer originären Ausbildungsrichtung eingesetzt werden (Driesner und Arndt 2020). Solche Maßnahmen – etwa der Einsatz von Gymnasiallehrkräften an Grundschulen – werden unter anderem auch von der Ständigen Wissenschaftlichen Kommission der KMK empfohlen (SWK 2023).

Angesichts dieser Entwicklung stellt sich die Frage, wie sich der flexible Einsatz von Lehrkräften an Grundschulen auf die Unterrichtsqualität und die Kompetenzen von Schüler.innen auswirkt. Bisherige Studien in Deutschland untersuchten vor allem die Auswirkungen des fachfremden Unterrichts (z. B. Porsch und Wendt 2016; Rjosk et al. 2017) bzw. des Quer- und Seiteneinstiegs auf die Schüler:innenleistung (Hoffmann und Richter 2016; Richter et al. 2019). Diese Untersuchungen fokussieren jedoch überwiegend auf die Sekundarstufe, sodass kaum Befunde für die Grundschule vorliegen. Zudem basieren sie auf Querschnittsdaten, die mittels Regressionsmodellen analysiert wurden, jedoch ohne Methoden wie das Propensity Score Matching (PSM), das kausale Inferenzen ermöglichen könnte (Keele et al. 2021). Es fehlen zudem Studien, die die Effekte des schulartfremden Unterrichtseinsatzes oder verschiedene Maßnahmen des Lehrkräfteeinsatzes vergleichend untersuchen. Hier setzt die vorliegende Arbeit an, indem mithilfe von Multilevel Propensity Score Matching (MLPSM) die Auswirkungen verschiedener Formen des flexiblen Lehrkräfteeinsatzes auf die Unterrichtsqualität und die Kompetenzen von Schüler:innen untersucht werden. Der Schwerpunkt liegt dabei auf dem fach- und schulartfremden Einsatz sowie dem Einsatz von Quer- und Seiteneinsteigenden, wobei Effekte auf Basis vergleichbarer Lehrkräftepaare analysiert werden (Pimentel und Keele 2018).

Für die Analyse wird der Datensatz des IQB-Bildungstrends 2021 verwendet, der bundesweit repräsentative Daten zu Unterrichtsqualitätsmerkmalen und Kompetenzen von Schüler:innen der 4. Jahrgangsstufe enthält (Stanat et al. 2022). Die Ergebnisse dieser Untersuchung liefern erste Hinweise auf die Auswirkungen des flexiblen Einsatzes verschiedener Lehrkräftegruppen und den möglichen Steuerungsbedarf. Zu Beginn der Arbeit wird der aktuelle Stand des Lehrkräftemangels beschrieben, gefolgt von einem Überblick über die bisherigen Forschungsergebnisse zu den Effekten eines flexiblen Lehrkräfteeinsatzes auf Unterrichtsqualität und die Kompetenzen von Schüler:innen.

1.1 Lehrkräftemangel in Deutschland

Das deutsche Bildungssystem steht vor erheblichen Herausforderungen, von denen der Lehrkräftemangel eine der dringlichsten ist (Autorengruppe Bildungsberichterstattung 2022; Klemm 2022a). Dieser gefährdet nicht nur die Unterrichtsversorgung, sondern beeinträchtigt nach Einschätzung der SWK auch die Unterrichtsqualität (SWK 2023, S. 6).

Der Lehrkräftemangel war bereits seit mehreren Jahren absehbar und hat in den letzten Jahren alle Schularten betroffen, insbesondere die Grundschulen (Klemm

und Zorn 2019). Auch für die kommenden Jahre bis 2035 prognostizieren mehrere Studien eine anhaltende Mangellage, wobei das Ausmaß je nach Schulform unterschiedlich eingeschätzt wird (Geis-Thöne 2022; Klemm 2022b; KMK 2024). Die SWK (2023) stuft die Situation an Grundschulen weiterhin als kritisch ein, da nach aktueller Prognose unter anderem der KMK bis mindestens 2025 ein Mangel an grundständig ausgebildeten Lehrkräften besteht. Darunter werden Lehrkräfte verstanden, die sowohl ein Lehramtsstudium als auch den Vorbereitungsdienst erfolgreich abgeschlossen haben (Klemm und Zorn 2024; KMK 2024).

Für die Entwicklung an Grundschulen über das Jahr 2025 hinaus zeigen sich unterschiedliche Einschätzungen. Klemm und Zorn (2024) prognostizieren bis 2035 ein Überangebot an grundständig ausgebildeten Lehrkräften. Auch die KMK (2024) geht für den Zeitraum von 2027 bis 2030 aktuell von einer Entspannung der Situation aus. Im Gegensatz dazu rechnet Dohmen (2024) mit einem anhaltenden Mangel.

Unabhängig von der tatsächlichen Entwicklung in den nächsten Jahren sehen sich viele Grundschulen weiterhin mit erheblichen Herausforderungen im Zusammenhang mit dem Lehrkräftemangel konfrontiert. So wurden in den letzten Jahren vielerorts nicht-grundständig ausgebildete Lehrkräfte unbefristet in den Schuldienst eingestellt, die unabhängig von der zukünftigen Versorgung mit grundständig ausgebildeten Lehrkräften dauerhaft im System verbleiben und unterrichten werden.

1.2 Flexibler Einsatz von Lehrkräften

Angesichts des Mangels an grundständig ausgebildeten Lehrkräften hat die KMK im Jahr 2013 verschiedene Maßnahmen auf den Weg gebracht, um dem Lehrkräftemangel mit alternativen Ansätzen zu begegnen (KMK 2013). Eine zentrale Strategie ist die Gewinnung von Personen ohne Lehramtsstudium für den Schuldienst. In der „Vereinbarung über Sondermaßnahmen“ definierte die KMK Qualifikationsanforderungen für nicht grundständig ausgebildete Lehrkräfte, darunter ein Masterabschluss oder ein gleichwertiger Hochschulabschluss in zwei für das Lehramt relevanten Fächern (KMK 2013). Personen, die nur in einem Fach qualifiziert sind, können durch ein Ergänzungsstudium die Qualifikation für ein weiteres Fach erwerben und sich für den Vorbereitungsdienst bewerben. Diese Gruppe wird in der Bildungsforschung als Quereinsteiger:innen bezeichnet (Melzer et al. 2014). Die Vereinbarung erlaubt zudem landesspezifische Sondermaßnahmen, wie z. B. in Brandenburg, Mecklenburg-Vorpommern und Sachsen-Anhalt. Dort werden Personen ohne formale Qualifikation in lehramtsrelevanten Fächern eingestellt und durch eine teils berufsbegleitende pädagogische Zusatzqualifizierung statt eines regulären Vorbereitungsdienstes vorbereitet – dies wird als Seiteneinstieg bezeichnet. Zudem besteht in vielen Ländern die Möglichkeit, grundständig ausgebildete Lehrkräfte fachfremd oder schulartfremd einzusetzen, also in Fächern oder Schularten, für die sie nicht ausgebildet wurden.

Verschiedene Studien haben einige dieser Maßnahmen, insbesondere im Primarbereich, umfassend dokumentiert (vgl. Porsch 2016; Richter und Marx 2019; Richter et al. 2018 Ziegler et al. 2019). Untersuchungen des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) zeigen, dass bundesweit ein erheblicher Teil der Grundschullehrkräfte fachfremd unterrichtet, wobei deutliche Länderunterschiede bestehen. 2011 unterrichteten durchschnittlich 17 % der Deutsch- und 27 % der Ma-

thematiklehrkräfte in der vierten Klasse fachfremd; 2016 stiegen die Anteile auf 19 % bzw. 32 % (Richter et al. 2012; Rjosk et al. 2017). Besonders in Berlin war der fachfremde Unterricht verbreitet: Im Schuljahr 2016/17 lag der Anteil in Deutsch, Englisch und Mathematik bei 56, 32 bzw. 69 % (Ziegler et al. 2019). Auch wenn die Daten für einige Bundesländer eine Zunahme anzeigen, weisen sie gleichzeitig darauf hin, dass fachfremd erteilter Unterricht an Grundschulen nicht erst im Zusammenhang mit dem aktuellen Lehrkräftemangel entstanden ist, sondern bereits länger praktiziert wird. Während das Ausmaß fachfremden Unterrichts im Primarbereich somit gut dokumentiert ist, liegt für den schulartfremden Einsatz von Lehrkräften bisher weniger Forschung vor.

Diese Forschungslücke zeigt sich auch in der unzureichenden Datenlage zu Lehrkräften, die über alternative Wege in den Schuldienst gelangt sind. Die amtliche KMK-Statistik konzentriert sich auf Einstellungszahlen und liefert nur begrenzte Informationen, da weder der genaue Anteil von Quer- und Seiteneinsteiger:innen an der Lehrerschaft noch deren Verbleib im Schuldienst detailliert erfasst wird. Seit 2011 berichtet die KMK jährlich über den Anteil der neu eingestellten Seiteneinsteiger:innen, die seit 2021 als „sonstige (unbefristet beschäftigte) Lehrkräfte" ausgewiesen werden, darunter auch Personen ohne Hochschulabschluss auf Masterniveau (KMK 2023). Zaruba et al. (2023) stellen fest, dass diese Gruppe inzwischen auch Personen ohne formale Hochschulzugangsberechtigung umfasst. Im Jahr 2022 stellten die 3153 unbefristet eingestellten Seiteneinsteiger:innen etwa 9 % aller Neueinstellungen dar (KMK 2023). Die Datenlage zu Quereinsteiger:innen ist weniger klar, da diese zunächst den Vorbereitungsdienst absolvieren. Laut KMK-Statistik wurden 2022 insgesamt 1178 Personen mit nicht-lehramtsbezogenem Studienabschluss in den Vorbereitungsdienst aufgenommen, etwa 4 % aller Neueinstellungen (KMK 2023). Wie viele davon später in den regulären Schuldienst übergehen, wird nicht erfasst. Zudem sei ergänzend darauf hingewiesen, dass die KMK-Statistik (2023) ausweist, dass der Quereinstieg insbesondere an beruflichen Schulen schon seit längerer Zeit ein etabliertes Phänomen ist und nicht ausschließlich als Folge des aktuellen Lehrkräftemangels zu betrachten ist.

1.3 Effekte eines flexiblen Lehrkräfteeinsatzes auf die Unterrichtsqualität

Der flexible Einsatz von Lehrkräften und insbesondere die von der KMK initiierten Sondermaßnahmen zur Rekrutierung von Lehrkräften stehen im Fokus intensiver und kontroverser Debatten in der Bildungswissenschaft und der Öffentlichkeit. Dabei geht es vor allem um die Frage, ob Personen ohne grundständige oder fach- und schulartgerechte Lehramtsausbildung über die nötige professionelle Kompetenz verfügen, um eine hohe Unterrichtsqualität sicherzustellen (Kunter et al. 2013).

Aus professionstheoretischer Perspektive, unter Bezugnahme auf das Modell der Determinanten und Konsequenzen professioneller Kompetenz von Lehrkräften (Kunter et al. 2011), lässt sich argumentieren, dass ein fehlender oder eingeschränkter Zugang zu formellen Lerngelegenheiten in der Lehramtsausbildung den Erwerb professioneller Kompetenz negativ beeinträchtigen muss. Empirische Studien untermauern diesen Zusammenhang, indem sie zeigen, dass die Nutzung von Lerngelegenheiten positiv mit wesentlichen Facetten professioneller Kompetenz, wie päd-

agogischem Wissen (König et al. 2017) oder technologisch-pädagogischem Wissen (Gerhard et al. 2023), zusammenhängt.

Von hohem Interesse ist deshalb u. a. die Frage, inwieweit Lehrkräfte, die über alternative Qualifikationswege in den Schuldienst eintreten oder fach- bzw. schulartfremd eingesetzt werden, einen qualitativ vergleichbaren Unterricht erteilen wie grundständig ausgebildete Lehrkräfte, die fach- und schulartgerecht eingesetzt werden.

Um die Qualität des Unterrichts fachübergreifend zu beurteilen und zu vergleichen, haben Klieme et al. (2001) das Modell der drei Basisdimensionen entwickelt (für ähnliche Operationalisierungen vgl. z. B. Pianta et al. 2008). Das Modell basiert auf einem pragmatisch-integrativen Verständnis von Unterrichtsqualität und unterscheidet die drei Tiefenmerkmale Klassenführung, konstruktive Unterstützung und kognitive Aktivierung (Klieme 2019). Diese Merkmale beschreiben, wie sich Lernende mit dem Lerngegenstand auseinandersetzen und wie dies durch die Lehrkraft begleitet und unterstützt wird. Die Dimension der Klassenführung umfasst Vorgehensweisen der Lehrkraft, um Störungen zu vermeiden (z. B. durch klare Regeln und Routinen) und Unterrichtsinhalte gut zu strukturieren, um die Lernzeit möglichst gut auszuschöpfen. Eine effektive Klassenführung gilt als wesentliche Grundvoraussetzung für lernwirksamen Unterricht (Kunter und Trautwein 2013). Sie stellt die organisatorische und emotionale Basis dar, auf der anspruchsvolle kognitive und unterstützende Prozesse erst realisiert werden können. Kognitive Aktivierung betrifft die Anregung tiefgehender Denk- und Verstehensprozesse, z. B. durch herausfordernde Aufgabenstellungen, bei denen die Schüler:innen angeregt werden, Lernprozesse und Lösungsprozesse aktiv und selbstständig zu reflektieren und mit ihrem Vorwissen in Beziehung zu setzen. Konstruktive Unterstützung umfasst das emotionale, motivationale und fachliche Begleiten von Lernprozessen durch die Lehrkraft, z. B. durch ein wertschätzendes Lernklima, konstruktives Feedback und Motivierung.

Aus fachdidaktischer Perspektive wird die theoretische Fundierung der Basisdimensionen teils kritisch hinterfragt (z. B. Rothland 2024). Zudem wird gefordert, die generischen Qualitätsmerkmale um fachspezifische Aspekte zu ergänzen oder zu konkretisieren und stärker theoretisch zu fundieren, um die Qualität des jeweiligen Fachunterrichts adäquat erfassen zu können – etwa durch die Berücksichtigung korrekter fachlicher Repräsentationen oder typischer Fehlermuster im Mathematikunterricht (vgl. Praetorius und Gräsel 2021). Gleichzeitig weisen zahlreiche Studien auf die fachübergreifende Validität und den praktischen Nutzen des Modells hin, da die drei Basisdimensionen unabhängig vom Fach eng mit unterschiedlichen kognitiven, motivationalen und emotionalen Aspekten des Lernerfolgs der Schüler:innen zusammenhängen und daher eine verlässliche Grundlage für die Einschätzung von Unterrichtsqualität bieten (für einen Überblick siehe Praetorius und Charalambous 2018; Praetorius et al. 2020).

Nur wenige Studien haben die Auswirkungen einzelner Maßnahmen auf die Unterrichtsqualität untersucht. Für den Quer- und Seiteneinstieg ist anzunehmen, dass die fehlende Lehramtsausbildung und das damit verbundene geringere pädagogische Wissen (Kleickmann und Anders 2011; Lucksnat et al. 2020; Oettinghaus et al. 2014) mit einer niedrigeren Unterrichtsqualität einhergehen. Eine empirische Studie von Lucksnat et al. (2024) vergleicht auf Basis von IQB-Daten die Unterrichtsqualität

grundständig und nicht-grundständig ausgebildeter Lehrkräfte im Fach Mathematik in der Sekundarstufe I und findet dabei jedoch keine signifikanten Unterschiede in der Bewertung des Unterrichts aus Sicht der Schüler:innen in den Dimensionen kognitive Aktivierung, Klassenmanagement und konstruktive Unterstützung. Vergleichbare Ergebnisse zeigten sich in Studien mit amerikanischen Lehrkräften (z. B. Hill et al. 2015; McCarty und Dietz 2011).

Fachfremd unterrichtende Lehrkräfte verfügen aufgrund der fehlenden Fakultas im unterrichtenden Fach über geringere fachliche und fachdidaktische Kenntnisse (Krauss et al. 2008). Dementsprechend wäre zu erwarten, dass die kognitive Aktivierung im Unterricht niedriger ausfällt als bei regulär qualifizierten Lehrkräften. Für die konstruktive Unterstützung und das Klassenmanagement sind diese Unterschiede nicht zu erwarten, da diese Qualitätsmerkmale weniger stark fachlich gebunden sind. Demgegenüber konnten Porsch und Wilden (2022) hinsichtlich fachfremd erteilten Unterrichts zeigen, dass sich fachfremd unterrichtende Grundschullehrkräfte in ihrer Selbsteinschätzung der Unterrichtsqualität in den Dimensionen kognitive Aktivierung und konstruktive Unterstützung nicht von fachgerecht unterrichtenden Lehrkräften unterscheiden. Im Bereich der Klassenführung zeigten sich sogar Vorteile zugunsten fachfremd unterrichtender Lehrkräfte. Die Autor:innen interpretieren dieses Ergebnis als Hinweis darauf, dass fachfremd unterrichtende Lehrkräfte ihren Unterricht im Hinblick auf effektives Klassenmanagement weniger kritisch beurteilen würden.

Für den schulartfremden Einsatz wäre zu erwarten, dass Personen ohne eine schulartspezifische Ausbildung über anderes fachliches bzw. fachdidaktisches Wissen verfügen und daher der Unterricht von schulartfremden Lehrkräften eine geringere Qualität aufweist (Blömeke et al. 2010). Untersuchungen zu dieser Frage fehlen jedoch bislang.

1.4 Effekte eines flexiblen Lehrkräfteeinsatzes auf die Schüler:innenleistung

Im Mittelpunkt dieses Kapitels steht die Frage, wie sich der flexible Einsatz von Lehrkräften, insbesondere die von der KMK initiierten Sondermaßnahmen, auf die Leistungen der Schüler:innen auswirkt. Aus theoretischer Sicht ist anzunehmen, dass Defizite in der Unterrichtsqualität – wie sie etwa bei schulartfremd eingesetzten, fachfremd unterrichtenden oder im Quer- und Seiteneinstieg tätigen Lehrkräften zu erwarten wären – mit geringeren Schüler:innenleistungen einhergehen (Kunter et al. 2011). Vor diesem Hintergrund widmet sich eine Reihe von Studien der Frage, inwieweit einzelne Maßnahmen im Zusammenhang mit einem flexiblen Lehrkräfteeinsatz mit dem Lernerfolg der Schüler:innen zusammenhängen. Die empirische Befundlage hierzu ist jedoch bislang uneinheitlich und weist teils erhebliche Forschungslücken auf. Hinsichtlich des Quer- und Seiteneinstiegs liefern Untersuchungen, die auf Daten des IQB-Bildungstrends basieren, unterschiedliche Ergebnisse. Während einige Studien keine signifikanten Leistungsunterschiede zwischen Schüler:innen feststellen, deren Lehrkräfte entweder grundständig ausgebildet wurden oder nicht, deuten andere auf spezifische Unterschiede hin. Insbesondere in weiterführenden Schulen zeigen die Ergebnisse für die Fächer Deutsch (Hoffmann und Richter 2016), Mathematik und Physik (Richter et al. 2019; Ziegler et al. 2022) sowie

Biologie und Chemie (Richter et al. 2019) keine signifikanten Unterschiede in den Schüler:innenleistungen, wenn die Klassenzusammensetzung berücksichtigt wird. Im Fach Englisch werden hingegen geringe Nachteile im Lese- und Hörverstehen bei Schüler:innen der Sekundarstufe I erfasst, wenn diese von nicht-grundständig ausgebildeten Lehrkräften unterrichtet werden (Hoffmann und Richter 2016). Für den Primarbereich liegen für Deutschland bislang keine Erkenntnisse zum Effekt von Lehrpersonen im Quer- und Seiteneinstieg auf die Schüler:innenleistungen vor. Diese nationalen Befunde korrespondieren mit internationalen Studien, die tendenziell ebenfalls nur minimale oder keine Unterschiede in den Schüler:innenleistungen zwischen grundständig und nicht grundständig ausgebildeten Lehrkräften feststellen (z. B. Constantine et al. 2009; Kane et al. 2008).

Die Effekte von fachfremd erteiltem Unterricht auf Schüler:innenleistungen auf Basis der IQB-Bildungstrends erweisen sich als inkonsistent und variieren je nach Schulform. Für die Sekundarstufe I konnte gezeigt werden, dass Schüler:innen mit fachfremd unterrichtenden Lehrkräften an nichtgymnasialen Schularten geringere Kompetenzen in den Fächern Mathematik, Biologie und Physik im Jahr 2012 (Richter et al. 2013) sowie in den Fächern Deutsch und Englisch im Jahr 2015 (Hoffmann und Richter 2016) erreichen als Schüler:innen mit Fachlehrkräften. Im Jahr 2018 zeigte sich dies allerdings nur noch im Fach Mathematik (Richter et al. 2019) und im Jahr 2022 im Fach Englisch (Hoffmann et al. 2023), während in den anderen Fächern keine Unterschiede mehr zu verzeichnen waren.

Ähnlich uneinheitlich präsentieren sich die Ergebnisse für den Primarbereich. Ergebnisse des IQB-Ländervergleichs aus dem Jahr 2011 und der TIMS-Studie deuten darauf hin, dass Schüler:innen in fachfremd unterrichteten Klassen durchschnittlich geringere Kompetenzen im Fach Deutsch (Lesen und Zuhören) und im Fach Mathematik sowie den Naturwissenschaften erzielen (Porsch und Wendt 2015; Richter et al. 2012). Diese Ergebnisse zeigten sich jedoch nicht mehr unter Kontrolle der Klassenzusammensetzung und konnten auch in Folgeuntersuchungen nicht repliziert werden (Porsch und Wendt 2016; Rjosk et al. 2017, Ziegler und Richter 2017). Auch die internationale Forschung zum fachfremden Unterricht zeigt heterogene Ergebnisse (Hanfstingl et al. 2023; Porsch und Whannell 2019; Van Overschelde 2022).

Zusammenfassend zeigt sich, dass die Forschung zu den Effekten der von der KMK initiierten Sondermaßnahmen sowie weiterer Maßnahmen wie fachfremdem oder schulartfremdem Unterricht differenzierte Analysen erfordert und bislang ein uneinheitliches Bild ergibt. Bei den Analysen handelt es sich überwiegend um Regressionsanalysen, die auf Querschnittsdaten beruhen und unterschiedliche Kovariaten mit einbeziehen und somit nur eingeschränkt miteinander vergleichbar sind. Auffällig ist wiederum, dass bislang keine Studien zu Effekten des schulartfremden Einsatzes von Lehrkräften vorliegen, was eine bedeutende Forschungslücke darstellt. In der vorliegenden Arbeit greifen wir ebenfalls auf Querschnittdaten zurück, jedoch nutzen wir für deren Analyse ein spezifisches Verfahren, welches die Unterschiede zwischen den betrachteten Gruppen durch die Bildung von statistischen Zwillingen minimiert.

1.5 Debatte um Ursache und Wirkung

In der öffentlichen Diskussion weisen Gewerkschaften und Verbände von Lehrkräften und Wissenschafter:innen darauf hin, dass nicht traditionell ausgebildete Lehrkräfte besondere Unterstützung benötigen, um eine Deprofessionalisierung und negative Auswirkungen auf die Unterrichtsqualität und den Lernerfolg der Schüler:innen zu vermeiden (z. B. Deutscher Philologenverband 2023; GEBF 2023; GEW Landesverband Sachsen-Anhalt 2024). Gängige theoretische Modelle, wie das Angebots-Nutzungs-Modell (Helmke 2015), unterstützen diese Annahme, da sie einen Zusammenhang zwischen den in der Ausbildung erworbenen Kompetenzen – für die die Qualifikation als Indikator dient – und der Unterrichtsqualität nahelegen. Allerdings kann diese Wirkrichtung auf Basis der vorhandenen empirischen Befunde nicht eindeutig bestätigt werden. Zunächst ist festzustellen, dass es sich bei den meisten Arbeiten um Querschnittsstudien handelt, in denen die Qualifikation von Lehrkräften als unabhängige Variable und die Unterrichtsqualität bzw. die Leistung der Schüler:innen als abhängige Variable betrachtet werden, um so den Effekt des entsprechendes Lehrkräftemerkmals zu bestimmen. Dabei bleibt in den bislang vorliegenden querschnittlichen Studien oft unberücksichtigt, dass nicht-grundständig ausgebildete Lehrkräfte nicht gleichermaßen auf Schulen eines Bundeslandes verteilt sind. Studien mit Daten des Landes Berlin konnten beispielweise zeigen, dass nicht-grundständig ausgebildete Lehrkräfte in der Sekundarstufe häufiger an nichtgymnasialen Schulen eingesetzt werden und dass diese häufiger in Klassen mit einem durchschnittlich geringeren sozio-ökonomischen Status zum Einsatz kommen (Helbig und Nikolai 2019; Ziegler et al. 2022). Auch für die Berliner Grundschulen zeigte sich, dass Quereinsteiger:innen häufiger an den Schulen zu finden sind, an denen häufiger Schüler:innen mit Lernmittelbefreiung und nicht-deutscher Herkunftssprache unterrichtet werden (Richter et al. 2018). Vor dem Hintergrund dieser Befunde wird deutlich, dass der Einsatz nicht-grundständig ausgebildeter Lehrkräfte nicht zufällig erfolgt, sondern häufiger an Schulen, die einen erhöhten Bedarf an Lehrkräften haben, der nicht vollständig durch grundständig ausgebildete Lehrkräfte gedeckt werden kann (Richter et al. 2024).

Um Unterschiede in der Unterrichtsqualität beziehungsweise im Lernerfolg kausal auf unterschiedliche Qualifikationen der Lehrkräfte zurückführen zu können, müsste letztere als (im statistischen Sinne) unabhängige Variable gelten können, was aufgrund der oben skizzierten Abhängigkeit des Einsatzes nicht-grundständig ausgebildeter Lehrkräfte nicht möglich ist. Eine solche Unabhängigkeit wäre etwa durch eine experimentell zufällige Zuweisung der Lehrkräfte auf die Klassen zu erreichen, die eine echte statistische Unabhängigkeit der Variable „Qualifikation" von allen anderen Kovariaten bedeuten würde (Campbell und Stanley 1963), jedoch praktisch nicht möglich ist, weil die Zuweisung natürlich bedarfsgemäß stattfindet. Alternativ könnten in Längsschnittstudien mittels wiederholter Befragungen die Veränderungen in den abhängigen Variablen „Unterrichtsqualität" und „Lernerfolg" zwischen Klassen verglichen werden, die von Lehrkräften mit unterschiedlicher Qualifikation unterrichtet werden. Für eine solche, sehr aufwändige Erhebung bestünde dennoch die Einschränkung, dass bei nicht zufälliger Zuweisung der Lehrkräfte zu den Klassen Unterschiede in den abhängigen Variablen bereits *vor* der Zuwei-

sung bestanden haben könnten, und dass Unterschiede in der Lernentwicklung auch darauf zurückzuführen sein könnten, dass Klassen mit unterschiedlichem Ausgangsniveau unterschiedliche Lernfortschritte erzielen, unabhängig von der Qualifikation der Lehrkräfte, die diese Klassen unterrichten (Holland 1986; Rubin 1974).

1.6 Forschungsfragen

1. Unterscheiden sich die Unterrichtsqualität und die Kompetenzen der Schüler:innen an Grundschulen zwischen Lehrkräften mit grundständiger Lehramtsausbildung, die fachfremd oder fachgerecht eingesetzt werden?
2. Unterscheiden sich die Unterrichtsqualität und die Kompetenzen der Schüler:innen an Grundschulen zwischen Lehrkräften mit grundständiger Lehramtsausbildung, die schulartfremd oder schulartgerecht eingesetzt werden?
3. Unterscheiden sich die Unterrichtsqualität und die Kompetenzen der Schüler:innen an Grundschulen zwischen Lehrkräften mit und ohne grundständiger Lehramtsausbildung?

2 Methode

2.1 Studiendesign und Stichproben

Die Datengrundlage für die Analysen bildete der IQB-Bildungstrend 2021, der zwischen April und Juni 2021 zur Überprüfung der KMK-Bildungsstandards im Primarbereich durchgeführt wurde (Stanat et al. 2022). Eine repräsentative Stichprobe von 26.844 Viertklässler:innen aus 1464 Schulen in allen 16 Bundesländern bearbeitete an zwei Erhebungstagen Kompetenztests in Deutsch und Mathematik. Zusätzlich füllten die Schüler:innen sowie ihre Eltern Begleitfragebögen aus. Die Teilnahme an den Tests war verpflichtend, während das Ausfüllen der Fragebögen je nach Bundesland freiwillig war (Teilnahmerate Test: 93 %, Schüler:innenfragebogen: 83 %, Elternfragebogen: 65 %). Zudem wurden 1738 Lehrkräfte befragt, die zum Zeitpunkt der Erhebung in den gezogenen Klassen Deutsch oder Mathematik unterrichteten. Die Teilnahme war in den meisten Ländern verpflichtend, in einigen jedoch freiwillig (Teilnahmerate: 77 %). Eine detaillierte Beschreibung des Erhebungsdesigns liefern Sachse et al. (2022).

Die Analysen wurden getrennt für Deutsch und Mathematik durchgeführt, wobei nur Daten von Schüler:innen einbezogen wurden, die klar einer Lehrkraft zugeordnet werden konnten, die nach den relevanten Kategorien (grundständige Ausbildung, fach- oder schulartfremder Einsatz bzw. nicht-grundständige Ausbildung) klassifiziert war. Da der Anteil der Lehrkräfte mit Quer- oder Seiteneinstieg gering war, wurde diese Gruppe zusammengefasst. Die entsprechenden Fallzahlen sind in Tab. 1 zu finden. Zusätzliche Informationen zur Zusammensetzung der Lehrkräftegruppen sind in Appendix A dargestellt.

Tab. 1 Stichprobegröße für Lehrkräfte und Schüler:innen

	Deutsch	*Mathematik*
	N	*N*
Lehrkräfte		
Grundständig ausgebildet und fach-/schulartgerecht eingesetzt	630	563
Fachfremd eingesetzt, aber grundständig ausgebildet	186	257
Schulartfremd eingesetzt, aber grundständig ausgebildet	79	70
Nicht-grundständig ausgebildet (Quer-/Seiteneinstieg)	48[1]	55[3]
Schüler:innen in Klassen von Lehrkräften, die ...		
Grundständig ausgebildet und fach-/schulartgerecht eingesetzt werden	11.778	10.610
Fachfremd eingesetzt, aber grundständig ausgebildet sind	3463	4663
Schulartfremd eingesetzt, aber grundständig ausgebildet sind	1432	1331
Nicht-grundständig ausgebildet sind (Quer-/Seiteneinstieg)	894[2]	1027[4]

[1]Quereinstieg = 21; Seiteneinstieg = 27
[2]Quereinstieg = 378; Seiteneinstieg = 516
[3]Quereinstieg = 24; Seiteneinstieg = 31
[4]Quereinstieg = 442; Seiteneinstieg = 585

2.2 Instrumente

Soziodemografischer und beruflicher Hintergrund der Lehrkräfte Die Lehrkräftebefragung umfasste demographische Angaben wie Alter und Geschlecht sowie Fragen zur Ausbildung und Berufstätigkeit. Erfasst wurden die gesamte Berufserfahrung der Lehrkraft und die Anzahl der Schulhalbjahre, in denen sie die getestete Klasse in Deutsch bzw. Mathematik unterrichtet hatte (Unterrichtsdauer in der Klasse). Die Qualifikation wurde durch Fragen zu lehramtsbezogenen Studienabschlüssen (Art und Fächer des Erst- bzw. Zweitstudiums) und zum Vorbereitungsdienst ermittelt. Zudem wurde abgefragt, welchem Lehramtstyp (z. B. Lehramt Grundschule/Primarbereich, Lehramt Sekundarstufe I) die Person aufgrund ihres Studiums oder Vorbereitungsdienstes zugeordnet werden kann.

Lehrkräfte, die ein grundschulspezifisches Lehramtsstudium und den Vorbereitungsdienst erfolgreich abgeschlossen hatten, wurden der Gruppe „grundständig ausgebildet und fach- beziehungsweise schulartgerecht eingesetzt“ zugeordnet. Als fachfremd wurden Lehrkräfte eingestuft, die ein Lehramtsstudium und den Vorbereitungsdienst abgeschlossen hatten, jedoch das von ihnen unterrichtete Fach (Deutsch oder Mathematik) nicht im Studium belegt hatten.[1] Lehrkräfte mit einem lehramtsbezogenen Studium und Vorbereitungsdienst, die jedoch nicht dem Lehramtstyp Lehramt Grundschule/Primarbereich zugeordnet werden konnten, wurden als schulartfremd erfasst.

[1] In Deutschland werden Grundschullehrkräfte grundsätzlich als Generalist:innen ausgebildet, das heißt, sie erwerben die Befähigung zum Unterrichten aller Fächer. Im Rahmen des Studiums erfolgt jedoch in der Regel eine vertiefte fachwissenschaftliche und fachdidaktische Ausbildung in zwei bis vier Fächern bzw. Lernbereichen (Liegmann et al. 2022). Die Bezeichnung „fachfremd“ bezieht sich in diesem Zusammenhang auf Lehrkräfte, die ein bestimmtes Fach – hier Deutsch oder Mathematik – nicht vertiefend studiert haben, dieses jedoch unterrichten.

Lehrkräfte ohne lehramtsbezogenes Studium, aber mit abgeschlossenem Vorbereitungsdienst, wurden der Gruppe der Quereinsteiger:innen zugeordnet. Lehrkräfte, die weder ein Lehramtsstudium noch den Vorbereitungsdienst absolviert hatten, wurden als Seiteneinsteiger:innen erfasst. Aufgrund der geringen Fallzahlen in beiden Gruppen (Deutsch: Quereinstieg $n = 21$, Seiteneinstieg $n = 27$; Mathematik: Quereinstieg $n = 24$, Seiteneinstieg $n = 31$) wurden Quer- und Seiteneinsteiger:innen in der Analyse zu einer gemeinsamen Kategorie „nicht grundständig ausgebildet" zusammengefasst. Diese Zusammenfassung ermöglicht eine analytisch tragfähige Unterscheidung zwischen grundständig und nicht grundständig qualifizierten Lehrkräften, ohne die Aussagekraft durch zu geringe Subgruppenstärken zu beeinträchtigen.

Soziodemografischer Hintergrund der Schüler:innen Die Schüler:innen- und Elternfragebögen erfassten den sozioökonomischen und kulturellen Hintergrund der Familie, u. a. durch die Anzahl der Bücher im Haushalt als etablierten Indikator für kulturelles Kapital (Heppt et al. 2022). Zudem wurden Angaben zu den Berufen der Eltern erhoben, um den International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom 2010) zu berechnen, der Bildung und Einkommen typischerweise mit einer Berufstätigkeit verknüpft. In den Analysen wurde der höchste ISEI-Wert der Eltern (HISEI) als Indikator des sozioökonomischen Status verwendet. Weiterhin wurden Fragen zum Geschlecht der Schüler:innen, ihrem Geburtsland, den Geburtsländern der Eltern und zu den Familiensprachen gestellt. Wo möglich, wurden die Angaben aus Schüler:innen- und Elternfragebögen zusammengeführt, um fehlende Werte zu minimieren. Informationen zum sonderpädagogischen Förderbedarf wurden direkt bei den Schulen erfragt.

Unterrichtsqualität Die Merkmale der Unterrichtsqualität wurden im Rahmen der Schüler:innenbefragung auf Basis des im deutschsprachigen Raum etablierten Modells der drei Basisdimensionen erfasst (Klieme et al. 2009). Als Indikator für die *Klassenführung* machten die Schüler:innen für die Fächer Deutsch und Mathematik auf einer vierstufigen Likert-Skala (1 = *stimmt gar nicht* bis 4 = *stimmt ganz genau*) Angaben dazu, inwieweit ihr Unterricht durch Störungen beeinträchtigt wird (6 Items, $\alpha_{\text{Deutsch/Mathematik}} = 0{,}88/0{,}90$, $\text{ICC1}_{\text{Deutsch/Mathematik}} = 0{,}28/0{,}28$, $\text{ICC2}_{\text{Deutsch/Mathematik}} = 0{,}85/0{,}85$; z. B. „In unserem Deutschunterricht/Mathematikunterricht stört keiner den Unterricht.").

Als Indikator für die *konstruktive Unterstützung* gaben die Schüler:innen auf der gleichen Antwortskala für die Fächer Deutsch und Mathematik an, inwieweit ihre Lehrkraft sie beim Lernen unterstützt, motiviert und ihre Bedürfnisse berücksichtigt (7 Items, $\alpha_{\text{Deutsch/Mathematik}} = 0{,}87/0{,}88$, $\text{ICC1}_{\text{Deutsch/Mathematik}} = 0{,}16/0{,}09$, $\text{ICC2}_{\text{Deutsch/Mathematik}} = 0{,}65/0{,}60$; z. B. „Meine Lehrerein/mein Lehrer macht mir Mut, wenn ich eine Aufgabe schwierig finde.").

Die *kognitive Aktivierung* wurde ebenfalls anhand der zuvor genannten vierstufigen Antwortskala erfasst. Im Fach Deutsch gaben die Schüler:innen für die Kompetenzbereiche Lesen und Orthografie sowie für das Fach Mathematik jeweils an, inwieweit im Unterricht Aktivitäten und Aufgaben eingesetzt werden, die vertiefte Verstehensprozesse anregen (Lesen/Orthografie/Mathematik: 6/6/9 Items, $\alpha_{\text{Lesen/Orthografie/Mathematik}} = 0{,}67/0{,}68/0{,}76$, $\text{ICC1}_{\text{Lesen/Orthografie/Mathematik}} = 0{,}07/0{,}07/0{,}08$,

$ICC2_{Lesen/Orthografie/Mathematik} = 0{,}55/0{,}52/0{,}56$; z. B. Lesen „In unserem Deutschunterricht sollen wir oft erklären, was ein Text bedeutet."; Orthografie: „In unserem Deutschunterricht sollen wir oft erklären, nach welcher Regel ein Wort geschrieben wird."; Mathematik: „In unserem Mathematikunterricht machen wir Aufgaben, über die ich ganz genau nachdenken muss.").

Kompetenztests Die Testaufgaben zu den Bildungsstandards wurden von interdisziplinären Teams entwickelt sowie vor dem Einsatz im Bildungstrend erprobt und validiert. Im Fach Deutsch umfassten die Tests 155 Items zum Lesen, 115 zum Zuhören und 204 zur Orthografie. Die mathematischen Kompetenzen wurden anhand von 406 Items zu den fünf Leitideen der Bildungsstandards erfasst und zu einem Globalwert zusammengefasst. Wie in großen Bildungsstudien üblich, wurde die Aufgabenmenge auf verschiedene Hefte verteilt (Multiple-Matrix-Design), sodass jede:r Schüler:in nur eine Teilmenge bearbeitete. Details zu den Konstrukten, ihrer Operationalisierung und dem Testdesign bietet Stanat et al. (2022).

2.3 Statistische Analysen

Die folgenden Analysen wurden getrennt für Deutsch und Mathematik durchgeführt. Zunächst wurden fehlende Werte auf Ebene der Schüler:innen (Level 1) und Klassenebene (Level 2) mittels Multiple Imputation by Chained Equations (MICE; van Buuren und Groothuis-Oudshoorn 2011; van Buuren 2018) in R (R Core Team 2020) ersetzt. Es wurden insgesamt 15 Imputationen erstellt. Anschließend wurden die Gesamtdatensätze für Deutsch und Mathematik jeweils in drei Teildatensätze unterteilt, die separat analysiert wurden. Jeder Teildatensatz umfasste alle grundständig ausgebildeten Lehrkräfte sowie eine der folgenden Gruppen: nicht grundständig ausgebildete (Quer- und Seiteneinsteigende), fachfremde oder schulartfremde Lehrkräfte. Diese Unterteilung ermöglichte die gezielte Untersuchung der Effekte unterschiedlicher Lehrkraftmerkmale auf die Unterrichtsqualität und die Schüler:innenleistung.

2.3.1 Multilevel Propensity Score Matching

Im zweiten Schritt wurde ein MLPSM mit dem Paket matchMulti in R durchgeführt (Pimentel 2023). Der Einsatz eines Propensity Score Matching (PSM) Ansatzes wurde gewählt, da es sich um ein nicht-randomisiertes Beobachtungsdesign handelt und daher nicht angenommen werden kann, dass sich die Vergleichsgruppen (z. B. grundständig vs. nicht-grundständig ausgebildete Lehrkräfte) ausschließlich in der interessierenden Variable unterscheiden. PSM ermöglicht es, beobachtete Kovariaten zwischen den Gruppen auszugleichen und so Verzerrungen durch selektive Zuweisung zu reduzieren (Rosenbaum und Rubin 1983).

Um der hierarchischen Struktur der Daten gerecht zu werden, kam ein multilevel-fähiges Verfahren (MLPSM) zum Einsatz, das eine Balance sowohl auf Schüler:innen- als auch auf Klassenebene anstrebt. MLPSM verringert systematische Unterschiede zwischen Treatment- und Kontrollgruppen und minimiert Verzerrungen in der Schätzung von Treatmenteffekten. Es berücksichtigt die verschachtelte

Datenstruktur und reduziert potenzielle Verzerrungen durch unbeobachtete heterogene Effekte (Keele et al. 2020, 2023).

Um vergleichbare Gruppen zu schaffen, wurden Anpassungen auf Schüler:innen- (L1) und Klassenebene (L2) vorgenommen. Auf L1 umfassten die analysierten Kovariaten Geschlecht, sonderpädagogischen Förderbedarf, Zuwanderungshintergrund, Familiensprache, sozioökonomischen Status (HISEI) und Anzahl der Bücher im Haushalt. Auf L2 wurden diese Variablen aggregiert und zusätzlich Lehrkraftmerkmale wie Geschlecht, Berufserfahrung und Unterrichtsdauer in der Klasse berücksichtigt (Keele et al. 2021). Das MLPSM erfolgte in zwei Phasen: Zuerst wurden die Lehrkräfte auf Basis der L2-Variablen abgeglichen, indem für jede Lehrkraft ein Propensity Score berechnet wurde, der ihre Zugehörigkeit zur Treatment- oder Kontrollgruppe angibt. Diese Scores dienten dazu, eine vergleichbare Verteilung der Lehrkräfte sicherzustellen (Keele et al. 2023; Pimentel 2023). Anschließend wurden die Unterschiede der Propensity Scores überprüft, um sicherzustellen, dass die standardisierten Mittelwertsdifferenzen unter dem Schwellenwert von 0,10 lagen (Cochran und Rubin 1973; Rosenbaum und Rubin 1985). In der zweiten Phase wurden die Schüler:innen innerhalb der abgeglichenen Lehrkraftpaare auf Basis individueller Merkmale abgeglichen, um Unterschiede innerhalb der Lehrkraftpaare zu minimieren (Pimentel und Keele 2018).

Nach dem MLPSM wurde die Balance der Kovariaten überprüft, um sicherzustellen, dass die Gruppen für die weiteren Analysen vergleichbar waren. Die Balance wurde auf beiden Ebenen bewertet, indem Mittelwerte und standardisierte Differenzen vor und nach dem MLPSM verglichen wurden (Keele et al. 2023, Appendix B und C). Die Ergebnisse zeigten, dass das MLPSM die Unterschiede auf L1 fast vollständig beseitigte, während auf L2 größere Unterschiede, insbesondere bei Lehrkraftmerkmalen wie Ausbildungstyp und Einsatzart, bestehen blieben. Um verbleibende Verzerrungen zu minimieren, wurden in den nachfolgenden Analysen alle Kovariaten kontrolliert, was als „doppeltes Adjustieren“ (Double Adjustment) bezeichnet wird (Austin 2011).

2.3.2 Multilevel-Regressionsanalysen

Im dritten Schritt wurden Mehrebenenanalysen auf den gematchten Datensätzen durchgeführt, um die Effekte des Lehrkräfteeinsatzes zu untersuchen. Doppelt latente Modelle der Unterrichtsqualität (Marsh et al. 2012) wurden geschätzt, wobei die Variablen auf L1 und L2 als latente, messfehlerbereinigte Konstrukte modelliert wurden, was präzisere Schätzungen ermöglichte. Alle Modelle wurden in M*plus* 8.8 (Muthén und Muthén 1998–2024) mit robuster Maximum-Likelihood-Schätzung (MLR) zur Korrektur möglicher Nicht-Normalität berechnet.

Zur Beantwortung der ersten Forschungsfrage führten wir Multilevel-Regressionsanalysen durch, wobei die Unterrichtsqualitätsdimensionen (kognitive Aktivierung, konstruktive Unterstützung, Klassenführung) und die Schüler:innenleistung in Deutsch (Lesen, Orthografie, Zuhören) und Mathematik als abhängige Variablen betrachtet wurden. Die Unterrichtsitems wurden standardisiert ($M = 0$, $SD = 1$). Auf L1 wurden Geschlecht der Schüler:innen, sonderpädagogischer Förderbedarf, Zuwanderungshintergrund, Erstsprache Deutsch, HISEI und die Anzahl der Bücher zu Hause

als Kovariaten einbezogen. Auf L2 wurden die Klassenmittelwerte dieser Variablen und zusätzlich das Geschlecht der Lehrkräfte, z-standardisierte Unterrichtserfahrung, Unterrichtsdauer in der Klasse (in Monaten) sowie die Art des Lehrkräfteeinsatzes berücksichtigt. Alle Variablen auf L2 wurden grand-mean-zentriert, um L1-Abweichungen vom Klassendurchschnitt und L2-Unterschiede zwischen Klassen im Vergleich zum Gesamtdurchschnitt zu erfassen.

Das Vorgehen zur Beantwortung der zweiten und dritten Forschungsfrage war analog, wobei ebenfalls Multilevel-Regressionsanalysen zur Untersuchung der Unterrichtsqualität und Schüler:innenleistung in Deutsch und Mathematik durchgeführt wurden, unter Berücksichtigung der gleichen Kovariaten zur Vorhersage des Lehrkräfteeffekts.

3 Ergebnisse

Forschungsfrage 1 zielt auf Unterschiede in der Unterrichtsqualität und den erreichten Kompetenzen zwischen fachfremd und fachgerecht eingesetzten Lehrkräften, wobei alle betrachteten Lehrkräfte eine grundständige Lehramtsausbildung durchlaufen haben. Die Ergebnisse in Tab. 2 weisen darauf hin, dass sich fachfremd unterrichtende Lehrkräfte in Deutsch und Mathematik weder in der durch die Schüler:innen wahrgenommenen Unterrichtsqualität noch in den erreichten Kompetenzen von ihren fachgerecht eingesetzten Lehrkräften mit grundständiger Lehramtsausbildung unterscheiden.

In Forschungsfrage 2 wird untersucht, ob es Unterschiede in der Unterrichtsqualität und den erreichten Kompetenzen von Schüler:innen zwischen Lehrkräften gibt, die entweder für die Grundschule oder eine andere Schulart qualifiziert wurden. Ähnlich wie bei den fachfremd unterrichtenden Lehrkräften zeigt sich für die Unterrichtsqualität kein Unterschied zwischen den schulartfremden und den schulartgerecht eingesetzten Lehrkräften in beiden Fächern (Tab. 2). Bei den erreichten Kompetenzen schneiden Klassen von schulartfremden Lehrkräften im Bereich Orthografie und im Zuhören signifikant schlechter ab als Klassen von ausgebildeten Grundschullehrkräften (Orthografie: $\beta = -0{,}17$, Zuhören: $\beta = -0{,}14$). Für das Fach Mathematik zeigten sich hingegen keine bedeutsamen Leistungsunterschiede zwischen Schüler:innen in Klassen, die von schulartfremd oder schulartgerecht eingesetzten Lehrkräften unterrichtet wurden.

Zur Beantwortung von Forschungsfrage 3 werden Lehrkräfte ohne grundständige Lehramtsausbildung mit grundständig ausgebildeten Lehrkräften bezogen auf die Unterrichtsqualität und die erreichten Kompetenzen ihrer Schüler:innen verglichen. Die Ergebnisse in Tab. 2 machen deutlich, dass sich keine statistisch bedeutsamen Unterschiede zwischen den beiden Gruppen in den verschiedenen Merkmalen der Unterrichtsqualität und den erreichten Kompetenzen ihrer Schüler:innen in den Fächern Deutsch und Mathematik feststellen lassen.

Tab. 2 Ergebnisse der Multilevel-Regressionsanalysen zur Vorhersage der Unterrichtsqualität und Kompetenzen der Schüler:innen

	Fachfremd vs. fachgerecht β (*SE*)	Schulartfremd vs. schulartgerecht β (*SE*)	Nicht-grundständig ausgebildet vs. grundständig ausgebildet β (*SE*)
Unterrichtsqualität			
Deutsch			
Kognitive Aktivierung (Lesen)	–0,03 (0,11)	0,04 (0,16)	–0,14 (0,24)
Kognitive Aktivierung (Orthografie)	0,20 (0,18)	0,38 (0,26)	0,14 (0,27)
Konstruktive Unterstützung	0,11 (0,22)	0,11 (0,25)	0,00 (0,38)
Klassenführung	0,15 (0,26)	0,23 (0,31)	–0,02 (0,37)
Mathematik			
Kognitive Aktivierung	–0,01 (0,02)	0,04 (0,05)	0,08 (0,05)
Konstruktive Unterstützung	–0,01 (0,03)	0,01 (0,07)	0,02 (0,07)
Klassenführung	0,06 (0,05)	0,01 (0,07)	0,05 (0,09)
Kompetenzen			
Deutsch			
Lesen	0,03 (0,04)	–0,11 (0,06)	–0,11 (0,06)
Orthografie	0,03 (0,05)	*–0,17* (0,07)*	–0,14 (0,09)
Zuhören	–0,05 (0,04)	*–0,14* (0,06)*	–0,08 (0,08)
Mathematik			
Globalwert	–0,01 (0,03)	–0,08 (0,06)	–0,07 (0,08)

Anm.: Dargestellt sind die Effekte des Lehrkräfteeinsatzes bzw. der Art der Ausbildung auf die Unterrichtsqualität bzw. auf die Kompetenz der Schüler:innen unter Kontrolle der Merkmale der Schüler:innen (L1), der Klassenkomposition (L2) und der Lehrkraft (L2). Die vollständigen Ergebnistabellen befinden sich im Appendix D, E und F
* $p < 0{,}05$, ** $p < 0{,}01$

4 Diskussion

Diese Arbeit untersucht, ob die Flexibilisierung der Lehrkräftegewinnung durch die im Jahr 2013 von der KMK beschlossenen Sondermaßnahmen sowie durch weitere Maßnahmen wie fachfremden oder schulartfremden Unterricht systematische Unterschiede in der Unterrichtsqualität und den Leistungen der Schüler:innen in der vierten Klasse zur Folge hat. Dazu wurde der repräsentative Datensatz des IQB-Bildungstrends 2021 genutzt, der sowohl Daten der Schüler:innen als auch Informationen zur Lehrkräftequalifikation enthält. Der Datensatz eignete sich besonders, da standardisierte Instrumente zur Erfassung von Unterrichtsqualität und Kompetenzen verwendet und Klassen aus allen Bundesländern zufällig ausgewählt wurden. Zudem wurde MLPSM angewendet, um Unterschiede nach Ausbildung (grundständig oder nicht-grundständig) und Einsatz (fachfremd/fachgerecht, schulartfremd/schulartgerecht) zu minimieren und kausale Effekte abzuschätzen. Diese Analyse

geht über bisherige Arbeiten hinaus, die vor allem für ausgewählte Hintergrundmerkmale der Schüler:innen kontrollierten.

Die Ergebnisse der Arbeit machen deutlich, dass sich die von den Schüler:innen wahrgenommenen Merkmale der Unterrichtsqualität weder für die fach- und schulartfremden Lehrkräfte noch für die nicht-grundständig ausgebildeten Lehrkräfte im Quer- oder Seiteneinstieg von grundständig ausgebildeten und sowohl fach- als auch schulartgerecht eingesetzten Lehrkräften signifikant unterscheiden. Lediglich im Fach Deutsch, in den Bereichen Orthografie und Zuhören, zeigten sich signifikante Leistungsnachteile bei Klassen, die von schulartfremden Lehrkräften unterrichtet wurden. In allen anderen Kompetenzbereichen und für andere Lehrkräftetypen wurden keine signifikanten Unterschiede festgestellt. Im Einklang mit früheren Studien (Porsch und Wilden 2022) deuten die Ergebnisse darauf hin, dass Schüler:innen den Unterricht bei fachfremden Lehrkräften als ebenso qualitätsvoll wahrnehmen wie bei fachgerecht unterrichtenden Lehrkräften. Systematische Kompetenzunterschiede in fachfremd unterrichteten Klassen wurden auch in früheren Studien nicht festgestellt (Porsch und Wendt 2016; Rjosk et al. 2017). Diese Studie bestätigt daher frühere Ergebnisse, die ohne PSM und teils basierend auf Lehrkräfteselbstberichten durchgeführt wurden.

Eine Erklärung für die fehlenden Unterschiede zwischen fachfremd und fachgerecht unterrichtenden Lehrkräften könnte sein, dass Grundschullehrkräfte zwar vertieft in zwei bis drei Fächern ausgebildet werden (Liegmann et al. 2022), in der Praxis jedoch häufig als Generalisten auch andere Fächer unterrichten (Künzli David et al. 2020). Dies wird häufig mit dem Klassenlehrprinzip begründet, das durch zahlreiche gemeinsame Unterrichtsstunden eine enge Bindung zu den Schüler:innen fördern soll (Hammel 2011). Dadurch sammeln sie Erfahrungen im fachfremden Unterricht, was Unterschiede zwischen den Gruppen verringert. Zudem könnten auch Fortbildungen Defizite ausgleichen (Porsch 2016).

Schulartfremde Lehrkräfte stehen dagegen vor der Herausforderung, Schüler:innen zu unterrichten, für die sie keine spezifische Qualifizierung haben. In dieser Studie zeigen sich jedoch nur geringe Kompetenznachteile in Deutsch (Orthografie und Zuhören), nicht in anderen Bereichen. Diese Unterschiede sind schwer erklärbar, aber die Literatur weist auf bestehende Unterschiede zwischen Grundschul- und Sekundarschullehrkräften hin. Anderson und Taner (2023) zeigten, dass Grundschullehrkräfte flexibler bei der Anpassung von Unterrichtsmaterialien sind und ein tieferes Verständnis der Entwicklungsstadien und Lernbedürfnisse ihrer Schüler:innen haben. Zudem setzen sie häufiger konstruktivistische Ansätze ein. Tatto und Senk (2011) betonen zudem, dass sich die Kompetenzen aufgrund unterschiedlicher Ausbildungsstrukturen unterscheiden, wobei Sekundarschullehrkräfte stärker auf spezialisierte Inhalte fokussiert sind. Die Unterschiede lassen sich möglicherweise auch darauf zurückführen, dass Lehrkräfte mit einer Lehrbefähigung für unterschiedliche Schularten auch über unterschiedliche Niveaus fachlichen und fachdidaktischen Wissens verfügen (Blömeke 2010). Somit könnte das spezifische fachdidaktische Wissen von Lehrkräften der Sekundarstufe I nur eingeschränkt für die Vermittlung von Unterrichtsinhalten in der Grundschule geeignet sein. Hinzu kommt, dass Lehrkräfte der Grundschule die spezifischen Herausforderungen im Schriftspracherwerb kennen sollten, um bei entsprechenden

Lerndefiziten angemessene Fördermaßnahmen ableiten zu können (Valtin 2020). Darüber hinaus benötigen Lehrkräfte der Grundschule Kenntnisse über die Wirksamkeit unterschiedlicher fachdidaktischer Konzepte (Weinhold 2009) sowie über die Bedeutsamkeit der fachspezifischen Ausdifferenzierung der Unterrichtsqualität und des fachspezifischen professionellen Lehrer:innenwissens (Hanke und König 2024). Es kann daher vermutet werden, dass Lehrkräfte, die nicht spezifisch für die Primarstufe qualifiziert wurden, nur begrenzte Kenntnisse in diesem Bereichen besitzen, und somit die Ergebnisse in den Kompetenztests im Fach Deutsch signifikant schlechter ausfallen. Genauere Erklärungen erfordern jedoch zusätzliche Daten der unterrichtenden Lehrkräfte (z. B. fachdidaktisches Wissen), die in der aktuellen Untersuchung nicht vorlagen.

Auch bei nicht-grundständig ausgebildeten Lehrkräften (Quer- und Seiteneinsteigende) wurden keine Nachteile in der Unterrichtsqualität oder den erworbenen Kompetenzen festgestellt. Diese Ergebnisse stimmen mit anderen Studien überein, die ebenfalls keine systematischen Unterschiede in der Unterrichtsqualität in der Sekundarstufe in Deutschland (Lucksnat et al. 2024) oder den USA fanden (Hill et al. 2015; McCarty und Dietz 2011). Erstmals zeigt diese Studie für Grundschulen in Deutschland, dass in Deutsch und Mathematik keine systematischen Leistungsnachteile feststellbar sind, was auch den Befunden für die Sekundarstufe I entspricht (Deutsch: Hoffmann und Richter 2016; Mathematik und Physik: Richter et al. 2019; Biologie, Chemie: Richter et al. 2019; Ziegler et al. 2022). Lediglich im Fach Englisch wurden in früheren Studien Nachteile für Schüler:innen in Klassen von Quer- und Seiteneinsteigenden festgestellt (Hoffmann und Richter 2016). Diese Studie liefert keine Hinweise darauf, dass Lehrkräfte ohne grundständige Ausbildung systematisch mit Nachteilen im Kompetenzaufbau der Schüler:innen einhergehen.

Die fehlenden Leistungsunterschiede zwischen Klassen von nicht-grundständig und grundständig ausgebildeten Lehrkräften könnten darauf zurückzuführen sein, dass Quer- und Seiteneinsteigende eine sehr motivierte Gruppe sind (Lucksnat et al. 2022), die bewusst den Lehrberuf gewählt haben. Diese hohe intrinsische Motivation könnte helfen, fachliche oder fachdidaktische Defizite durch Fortbildungen zu kompensieren. Daten aus dem IQB-Bildungstrend 2018 unterstützen dies, da Quereinsteigende häufiger Fortbildungen zu fachdidaktischen Themen besuchten als grundständig ausgebildete Lehrkräfte (Richter et al. 2019).

Die hier untersuchte Gruppe von Quer- und Seiteneinsteigenden ist zudem vermutlich positiv selektiert, da sie trotz bekannter Herausforderungen im Lehrberuf verblieben ist. Studien aus den USA (Redding und Smith 2016) zeigen, dass Lehrkräfte ohne grundständige Ausbildung häufiger frühzeitig aus dem Beruf ausscheiden. Ähnliche Tendenzen könnten auch in Deutschland bestehen. Richter et al. (2022) weisen darauf hin, dass Seiteneinsteigende mit höherer Selbstwirksamkeit eher im Beruf verbleiben. Westphal et al. (2024) zeigen zudem, dass Lehramtsstudierende mit höherer Selbstwirksamkeit und emotionaler Stabilität geringere Abbruchsabsichten haben. Da die betrachteten Lehrkräfte im Beruf geblieben sind, handelt es sich vermutlich um eine resilientere Gruppe mit höherer Kompetenz. Systematische Daten zum Dropout von Quer- und Seiteneinsteigenden gibt es in Deutschland jedoch nicht, es existieren nur vereinzelte Berichte, etwa aus der Pres-

se. So endeten im Schuljahr 2022/23 in Sachsen 121 von 330 neu eingestellten Seiteneinsteigenden ihren Dienst vorzeitig (Olbrisch 2024).

Ein weiterer Erklärungsansatz für die fehlenden Unterschiede in der von Schüler:innen wahrgenommenen Unterrichtsqualität könnte in der eingeschränkten Urteilsfähigkeit von Grundschulkindern hinsichtlich der im Fragebogen abgefragten Konzepte liegen. Die valide Einschätzung von Merkmalen zur Unterrichtsqualität setzt ein gewisses Maß sprachlicher Differenzierungsfähigkeit voraus, welche bei jüngeren Schüler:innen nicht immer vollständig ausgeprägt ist. Wie Lenske (2016) zeigt, kann es insbesondere in der Grundschule vorkommen, dass Schüler:innen die intendierten Inhalte der Fragebogenitems nicht in der gewünschten Weise erfassen. Gleichwohl belegen andere Studien, dass Grundschulkinder durchaus in der Lage sind, die Qualität ihres Unterrichts differenziert und valide einzuschätzen (Fauth 2021), und dass diese Einschätzungen substanzielle Zusammenhänge mit ihren Lernleistungen aufweisen (Fauth et al. 2019).

Eine weitere mögliche Erklärung für die weitgehend ausbleibenden Unterschiede in der Unterrichtsqualität und den Schüler:innenleistungen könnte darin liegen, dass die traditionelle, grundständige Lehrkräftebildung in ihrer derzeitigen Ausgestaltung möglicherweise nicht in allen Fällen zu einem systematischen Kompetenzvorsprung führt. Diese These stützt eine Vergleichsstudie, die bereits im Vorbereitungsdienst zeigt, dass sich Quereinsteigende und regulär ausgebildete Lehrkräfte nur in wenigen Merkmalen (z. B. im pädagogisch-psychologischen Wissen) unterscheiden (Lucksnat et al. 2020). Die vorliegenden Nullbefunde könnten demnach nicht nur auf fehlende Nachteile der untersuchten Maßnahmen hindeuten, sondern auch darauf, dass die universitäre Ausbildung bislang nicht in dem Maße wirksam ist, wie es angesichts ihres hohen Aufwands zu erwarten wäre.

In diesem Zusammenhang wäre ebenfalls zu vermuten, dass Lehrkräfte sich im Zuge der Ausübung ihrer Tätigkeit professionalisieren. Unterschiede in der Unterrichtsqualität zwischen grundständig ausgebildeten und nicht-grundständig ausgebildeten Lehrkräften würden sich demnach am ehesten zu Beginn des Berufseinstiegs zeigen, könnten jedoch durch praxisnahe Lerngelegenheiten wie kollegiale Unterrichtshospitationen oder gezielte Fortbildungen zunehmend nivelliert werden. Die Befunde zur Wirkung von Berufserfahrung auf Unterrichtsqualität und Schülerleistungen sind insgesamt jedoch heterogen. Einzelne Studien belegen positive Effekte wachsender Berufserfahrung auf die Lernleistungen der Schüler:innen (z. B. Ladd und Sorensen 2017). Andere Untersuchungen relativieren diese Ergebnisse und kommen zu dem Schluss, dass nach Kontrolle von Kontextfaktoren – etwa der sozialen Zusammensetzung der Klasse oder der Schulumgebung – keine signifikanten Effekte zusätzlicher Berufsjahre mehr nachweisbar sind (z. B. Klusmann et al. 2016). Somit lässt sich nicht abschließend beurteilen, wie sich die Berufserfahrung auf die Unterrichtsqualität und die Schüler:innenleistung auswirkt.

4.1 Limitationen

Die vorliegende Untersuchung weist trotz der insgesamt hohen Datenqualität und dem methodisch anspruchsvollen Vorgehen verschiedene Einschränkungen auf, die nachfolgend diskutiert werden sollen. Bei der durchgeführten Studie handelt es sich

um eine Querschnittsuntersuchung, bei der die Merkmale der Lehrkräfte und die Qualität des Unterrichts sowie die erreichten Leistungen zeitgleich erhoben wurden. Um die Befunde weiter zu untermauern, bedürfte es anderer Studiendesigns (z. B. experimentelle Längsschnittstudien), die jedoch für den Bereich des Lehrkräfteeinsatzes nur schwer realisierbar und forschungsethisch nicht vertretbar sind. Trotz dieser Limitation im vorliegenden Studiendesign haben wir es durch das gewählte Auswertungsverfahren erreicht, eine möglichst große Zahl an Unterschieden zwischen den betrachteten Gruppen zu minimieren.

Eine weitere Einschränkung der Arbeit besteht in der Fokussierung auf nur zwei ausgewählte Fächer und eine Jahrgangsstufe. Um die Effekte der untersuchten Maßnahmen ganzheitlicher untersuchen zu können, wäre es notwendig gewesen, auch weitere Fächer und Jahrgangsstufen, insbesondere den Anfangsunterricht, der aus grundschulpädagogischer Sicht eine zentrale Bedeutung hat (Hess und Elting 2024), in den Fokus zu rücken. Für solche speziellen Untersuchungen bedarf es eigener Datenerhebungen, da die vorliegenden Datensätze des IQB nur punktuelle Einblicke erlauben.

Anzumerken ist weiterhin, dass die Fallzahlen in den Gruppen der Quer- und Seiteneinsteiger:innen nur sehr gering ausfielen. Aus diesem Grund wurden beide Gruppen in den Analysen zusammengefasst, wodurch keine differenzierten Aussagen zu möglichen Unterschieden zwischen diesen beiden Qualifikationswegen getroffen werden können. Künftige Studien sollten hier gezielt differenzieren, um spezifischere Erkenntnisse zu ermöglichen.

Darüber hinaus wurde Unterrichtsqualität in der vorliegenden Studie anhand generischer Merkmale (Klassenführung, konstruktive Unterstützung, kognitive Aktivierung) und damit fachübergreifend erfasst. Dieser Zugang ermöglicht vergleichende Einschätzungen über die hier betrachteten Fächer (Deutsch und Mathematik) und Kompetenzbereiche hinweg. Aus fachdidaktischer Perspektive wird dies mitunter kritisch diskutiert, weil fachspezifische Aspekte dadurch nicht hinreichend berücksichtigt werden (vgl. Praetorius und Gräsel 2021; Rothland 2024). Die Operationalisierung der kognitiven Aktivierung erfolgte zumindest für das Fach Deutsch differenziert für die Kompetenzbereiche Lesen und Orthografie, weitere fachbezogene Konkretisierungen und Ergänzungen waren aber auch aufgrund der Datenerhebung im Rahmen des nationalen Bildungsmonitorings nicht möglich. Inwieweit sich die hier untersuchten Gruppen von Lehrkräften in weiteren, hier nicht betrachteten Aspekten stärker fachspezifischer Unterrichtsqualitätsmerkmale unterscheiden, bleibt somit unklar und muss in Folgestudien untersucht werden.

Zu berücksichtigen ist weiterhin, dass die Unterrichtsqualität ausschließlich aus Sicht der Schüler:innen erfasst wurde, da sie den Unterricht direkt und kontinuierlich erleben und die primären Adressat:innen der Lehr-Lernprozesse sind. Studien haben wiederholt gezeigt, dass Schüler:innen die realisierte Umsetzung zentraler Qualitätsdimensionen wie kognitive Aktivierung, Klassenführung und konstruktive Unterstützung und deren Wirkung auf ihr Lernen valide und zuverlässig beurteilen können und diese Einschätzungen mit Indikatoren des Lernerfolgs im Zusammenhang stehen (Fauth et al. 2014; Praetorius et al. 2020). Gleichzeitig ist nicht auszuschließen, dass Lehrkräfte oder externe Beobachter:innen zu anderen Bewertungen kommen und die Einschätzungen auch innerhalb der Schüler:innenschaft variieren

können, etwa in Abhängigkeit vom individuellen Leistungsniveau. Ergebnisse der TIMS-Studie 2023 zeigen hierzu ein differenziertes Bild: Während leistungsstärkere Schüler:innen ihren Mathematikunterricht als störungsärmer wahrnahmen und angaben, sich stärker durch ihre Lehrkraft unterstützt zu fühlen als leistungsschwächere Schüler:innen, fanden sich keine signifikanten Unterschiede zwischen unterschiedlich leistungsstarken Schüler:innengruppen bezogen auf die wahrgenommene kognitive Aktivierung (Scholz et al. 2024). Urteile zu sozialen Organisations- und Unterstützungsprozessen im Unterricht könnten möglicherweise leistungssensitiver sein als solche zur kognitiven Aktiviertheit. Um insbesondere leistungsbezogenen Einflüssen Rechnung zu tragen, wurden die Kompetenzen in den Analysen kontrolliert.

Weiterhin ist zu berücksichtigen, dass sich die Einschätzung der Unterrichtsqualität auf das laufende Schuljahr bezieht, in dem der Unterricht pandemiebedingt teils als Präsenz-, teils als Distanz- bzw. Wechselunterricht stattfand. Analysen des IQB-Bildungstrends 2022 zeigen, dass sich die Schüler:innenurteile zur Unterrichtsqualität in der Sekundarstufe I (Fach Deutsch) im Vergleich zu 2012 bzw. 2015 kaum verändert haben (Hoffmann et al. 2023). Auch in TIMSS 2023 ergaben sich für den Mathematikunterricht in der Primarstufe keine nennenswerten Unterschiede in der wahrgenommenen Unterrichtsqualität im Vergleich zu 2019 (Scholz et al. 2024). Diese Befunde deuten darauf hin, dass der Unterricht aus Sicht der Schüler:innen trotz pandemiebedingter Veränderungen nicht grundsätzlich als weniger qualitätsvoll wahrgenommen wurde. Gleichzeitig zeigen die Ergebnisse des IQB-Bildungstrends 2021, dass sich das durchschnittliche Kompetenzniveau zwischen 2016 und 2021 in allen untersuchten Bereichen deutlich verringert hat (Stanat et al. 2022). Es ist nicht auszuschließen, dass diese Entwicklung zumindest teilweise auf ungünstige Effekte veränderter Lernsettings im Distanz- und Wechselunterricht zurückzuführen ist. Aus Sicht von Lehrkräften und Eltern war der Unterricht in dieser Zeit häufig durch ein reduziertes Anforderungsniveau, einen verstärkten Fokus auf Übung und Wiederholung sowie eingeschränkte Möglichkeiten zur kognitiven Aktivierung und Motivierung gekennzeichnet (vgl. zusammenfassend Voss und Wittwer 2020). Vor diesem Hintergrund sollten die Befunde der vorliegenden Studie mit einer gewissen Zurückhaltung interpretiert und durch zukünftige Untersuchungen weiter abgesichert werden.

Bei der Interpretation der Befunde ist weiterhin nicht auszuschließen, dass andere schulische Einflussfaktoren (zum Beispiel Lehrmaterialien) ebenfalls bedeutsam sind. So konnten beispielsweise van den Ham und Heinze (2018) zeigen, dass in der Grundschule, insbesondere im Fach Mathematik, nicht die Qualifikation der Lehrkräfte, sondern die Wahl des eingesetzten Schulbuchs signifikante Unterschiede im Lernzuwachs der Schüler:innen erklärte. Dies verweist auf die Bedeutung curricularer und didaktischer Rahmenbedingungen und legt nahe, dass die Wirkung professioneller Merkmale von Lehrkräften stets im Zusammenspiel mit weiteren Kontextfaktoren zu betrachten ist.

4.2 Implikationen für Bildungspraxis und Forschung

Aus den Ergebnissen dieser Studie ergeben sich einige Implikationen, insbesondere für die Bildungsadministration und die Schulleitungen, die für den Lehrkräfteeinsatz verantwortlich sind. Zunächst liefert diese Arbeit empirische Evidenz zu einer Diskussion, die häufig nur mit wenig oder ohne wissenschaftliche Evidenz geführt wird. Auf Basis dieser Ergebnisse lässt sich festhalten, dass sich keine systematischen Hinweise auf eine geringere Unterrichtsqualität oder ein niedrigeres Kompetenzniveau der Schüler:innen im Zusammenhang mit dem Einsatz fachfremder oder nicht-grundständig ausgebildeter Lehrkräfte finden. Die Befunde sind jedoch vor dem Hintergrund der dargelegten Limitationen mit Zurückhaltung zu interpretieren. Aus diesem Grund lassen sich auf Basis dieser Studie keine Argumente dafür ableiten, den fachfremden Unterrichtseinsatz und den Einsatz von Quer- und Seiteneinsteigenden grundsätzlich zu unterbinden. Da die Effekte des Quer- und Seiteneinstiegs in dieser Arbeit jedoch für alle im System befindlichen Personen betrachtet wurden, kann kein Beleg dafür geliefert werden, dass diese Ergebnisse auch für Noviz:innen im Quer- und Seiteneinstieg gelten. Vor dem Hintergrund anderer Untersuchungen zu Kompetenzen und Herausforderungen von neu eingestellten Quer- und Seiteneinsteigenden lässt sich ableiten, dass gerade zu Beginn der beruflichen Karriere besondere Unterstützung, beispielsweise im Bereich der Differenzierung sowie dem Umgang mit Störungen (Richter et al. 2023), notwendig ist. Aus diesem Grund sollte insbesondere die große Zahl von Quer- und Seiteneinsteigenden durch umfassende Qualifizierungsangebote unterstützt werden. Aufgrund der negativen Effekte, die bei schulartfremden Lehrkräften in den Bereichen Orthografie und Zuhören festgestellt wurden, könnten sie von zusätzlicher Qualifizierung für den Grundschulunterricht profitieren. Solche Qualifizierungsangebote könnten z. B. Inhalte zu fachdidaktischen Besonderheiten bei der Vermittlung von fachlichen Inhalten an Grundschüler:innen umfassen.

Neben den Schlussfolgerungen für die Bildungspraxis lassen sich auch Implikationen für die Forschung ableiten. Zunächst wäre es wünschenswert, genauer zu verstehen, warum es zwischen den verschiedenen Lehrkräftegruppen kaum systematische Unterschiede in der Unterrichtsqualität und in den erworbenen Kompetenzen gibt. Ein besonderer Schwerpunkt sollte hierbei auf dem Vergleich zwischen grundständig ausgebildeten Lehrkräften und Quer- und Seiteneinsteigenden liegen. Vor dem Hintergrund der starken Unterschiede in der von beiden Gruppen durchlaufenen Qualifizierung stellt sich die Frage, warum dennoch keine systematischen Unterschiede festgestellt werden können. Zur Untersuchung dieser Frage sollten Vergleichsstudien zu Beginn der beruflichen Tätigkeit ansetzen und sowohl professionelle Kompetenzen als auch das berufliche Handeln im Unterricht in den Blick nehmen. Bisher liegen keine Befunde darüber vor, in welchen Bereichen regulär qualifizierte Lehramtsanwärter:innen nach dem Hochschulstudium über umfassendere Fähigkeiten verfügen als nicht-grundständig ausgebildete Lehrkräfte zu Beginn ihrer Tätigkeit. Solche Untersuchungen könnten jedoch wichtigen Aufschluss darüber geben, ob und inwiefern die Lehrkräftebildung in Deutschland so aufgestellt ist, dass diese zu einem systematischen Aufbau von Professionswissen beiträgt.

4.3 Fazit

Die Studie deutet darauf hin, dass weder die Umsetzung der von der KMK beschlossenen Sondermaßnahmen noch weitere Maßnahmen zur Bewältigung des Lehrkräftemangels in der Grundschule, etwa fachfremder oder schulartfremder Unterricht, systematisch mit negativen Auswirkungen auf die Unterrichtsqualität verbunden sind. Nur in Einzelfällen wurde ein geringeres Kompetenzniveau bei den Schüler:innen festgestellt. Ähnlich wie frühere Untersuchungen in der Sekundarstufe I werfen die Ergebnisse Fragen zur Qualität des regulären Lehramtsstudiums auf. Dies könnte Anlass geben, kritisch zu hinterfragen, welche Kompetenzen angehende Lehrkräfte an Universitäten und Pädagogischen Hochschulen erwerben und wie deren Erwerb überprüft und sichergestellt wird. Angesichts der zunehmenden Diversifizierung der Zugangswege ins Lehramt – etwa durch duale Studiengänge, Quereinstiegsmaster und Ein-Fach-Lehrkräfte – wird es umso wichtiger, die Kompetenzen angehender Lehrkräfte systematisch zu überprüfen.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (https://doi.org/10.1007/s11618-026-01380-0) enthalten.

Förderung Diese Forschungsarbeit erhielt keine spezifischen Zuschüsse von öffentlichen, kommerziellen oder gemeinnützigen Fördereinrichtungen.

Funding Open Access funding enabled and organized by Projekt DEAL.

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.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786.

Autorengruppe Bildungsberichterstattung (2022). *Bildung in Deutschland 2022. Ein indikatorengestützter Bericht mit einer Analyse zum Bildungspersonal*. wbv. https://doi.org/10.3278/6001820hw.

Blömeke, S. (2010). Schulform- oder schulstufenspezifische Lehrerprofessionalität? In K.-H. Arnold, K. Hauenschild, B. Schmidt & B. Ziegenmeyer (Hrsg.), *Zwischen Fachdidaktik und Stufendidaktik. Perspektiven für die Grundschulpädagogik* (S. 33–44). Verlag für Sozialwissenschaften.

Blömeke, S., Kaiser, G., Döhrmann, M., & Lehmann, R. (2010). Mathematisches und mathematikdidaktisches Wissen angehender Sekundarstufen-I-Lehrkräfte im internationalen Vergleich. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten*

angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich (S. 197–238). Waxmann.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Hrsg.), *Handbook of research on teaching* (S. 171–246). Rand McNally.

Cochran, W. G., & Rubin, D. B. (1973). Controlling bias in observational studies: a review. *Sankhyā: The Indian Journal of Statistics*, *35*(4), 417–446.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification. Final report*. Mathematica Policy Research Reports.

Deutscher Philologenverband (2023). *Positionen des PhV BW zum Thema „Quer- und Seiteneinstieg ins gymnasiale Lehramt"*. https://www.dphv.de/2023/02/16/positionen-des-phv-bw-zum-thema-quer-und-seiteneinstieg-ins-gymnasiale-lehramt/

Dohmen, D. (2024). *Lehrkräftemangel! Und kein Ende in Sicht. FiBS-Forum Nr. 79. Forschungsinstitut für Bildungs- und Sozialökonomie. FiBS*. https://www.fibs.eu/fileadmin/user_upload/images/Leistungen/FiBS-Forum_79_Lehrkraeftebedarf_240301_final.pdf. Zugegriffen: 10. Okt. 2024.

Driesner, I., & Arndt, M. (2020). Die Qualifizierung von Quer- und Seiteneinsteiger*innen. Konzepte und Lerngelegenheiten im bundesweiten Überblick. *DDS – Die Deutsche Schule*, *2020*(4), 414–427. https://doi.org/10.31244/dds.2020.04.05.

Fauth, B. (2021). Schülerurteile zur Unterrichtsqualität in der Grundschule: Was messen wir da eigentlich? In K. Göbel, C. Weiss, K. Neuber & M. Raaflaub (Hrsg.), *Quo Vadis Forschung Schülerrückmeldungen zum Unterricht* (S. 49–63). Springer.

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001.

Fauth, B., Decristan, J., Decker, A., Büttner, G., Klieme, E., Hardy, I., & Kunter, M. (2019). Teachers' professional competence, teaching quality, and student outcomes in elementary science education. *Teaching and Teacher Education*, *86*, 102882. https://doi.org/10.1016/j.tate.2019.102882.

Ganzeboom, H. B. G. (2010). *A new international socio-economic index [ISEI] of occupational status for the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from the ISSP 2002–2007; with an analysis of quality of educational measurement in ISSP*. Presentation at the Annual Conference of International Social Survey Programme, Lissabon.

GEBF = Gesellschaft für Empirische Bildungsforschung (2023). *Alternative Qualifikationswege für Lehrkräfte ohne traditionelles Lehramtsstudium in Zeiten des Lehrkräftemangels*. https://www.gebf-ev.de/app/download/9408166676/GEBF_Stellungnahme_2023_07_24-1.pdf?t=1700415043

Geis-Thöne, W. (2022). *Lehrkräftebedarf und -angebot: Bis 2035 steigende Engpässe zu erwarten. Szenariorechnungen zum INSM-Bildungsmonitor*. IW.

Gerhard, K., Jäger-Biela, D., & König, J. (2023). Opportunities to learn, technological pedagogical knowledge, and personal factors of pre-service teachers: Understanding the link between teacher education program characteristics and student teacher learning outcomes in times of digitalization. *Zeitschrift für Erziehungswissenschaft*, *26*(3), 653–676. https://doi.org/10.1007/s11618-023-01162-y.

Hammel, L. (2011). *Selbstkonzepte fachfremd unterrichtender Musiklehrerinnen und Musiklehrer an Grundschulen. Eine Grounded-Theory-Studie*. LIT.

Hanfstingl, B., Gnambs, T., Porsch, R., & Jude, N. (2023). Exploring the association between non-specialised science teacher rates and student science literacy: an analysis of PISA data across 18 nations. *International Journal of Science Education*, *9*, 874–892. https://doi.org/10.1080/09500693.2023.2262729.

Hanke, P., & König, J. (2024). Auf die Lehrperson, ihr Wissen und ihren Unterricht kommt es an? Professionelles Wissen von Lehrkräften, Unterrichtsqualität und Lernfortschritte von Schüler:innen im basalen Lese- und Schreibunterricht (Projekt WibaLeS). In J. König, C. Hanisch, P. Hanke, T. Hennemann, K. Kaspar, M. Martens & S. Strauß (Hrsg.), *Auf die Lehrperson und ihren Unterricht kommt es an. Zehn Jahre empirische Professions- und Unterrichtsforschung im IZeF der Universität zu Köln* (S. 81–101). Waxmann. https://doi.org/10.31244/9783830998761.

Helbig, M., & Nikolai, R. (2019). *Bekommen die sozial benachteiligsten Schülerinnen die „besten" Schulen? Eine explorative Studie über den Zusammenhang von Schulqualität und sozialer Zusammensetzung von Schulen am Beispiel Berlins** (WZB Discussion Paper No. P 2019-002). Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts: Franz Emanuel Weinert gewidmet*. Klett.

Heppt, B., Olczyk, M., & Volodina, A. (2022). Number of books at home as an indicator of socioeconomic status: examining its extensions and their incremental validity for student achievement. *Social Psychology of Education*, *25*, 903–928. https://doi.org/10.1007/s11218-022-09704-8.

Hess, M., & Elting, C. (2024). Anfangsunterricht. In M. Götz, A. Hartinger, F. Heinzel, J. Kahlert, S. Miller & U. Sandfuchs (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (S. 501–510). Klinkhardt.

Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for teaching. *AERA Open*, *1*(4), 233285841561770. https://doi.org/10.1177/2332858415617703.

Hoffmann, L., & Richter, D. (2016). Aspekte der Aus- und Fortbildung von Deutsch- und Englischlehrkräften im Ländervergleich. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 481–507). Waxmann.

Hoffmann, L., Richter, D., & Weirich, S. (2023). Aspekte der Ausbildung von Deutsch- und Englisch-lehrkräften. In P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel & K. A. Sachse (Hrsg.), *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* (S. 389–407). Waxmann.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960. https://doi.org/10.2307/2289064.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615–631. https://doi.org/10.1016/j.econedurev.2007.05.005.

Keele, L., Pimentel, S., & Miratrix, L. (2020). Multilevel matching in cluster randomized trials with noncompliance. *Annals of Applied Statistics*, *14*(2), 783–800. https://doi.org/10.1214/19-AOAS1294.

Keele, L., Pimentel, S., & Miratrix, L. (2021). Matching methods for clustered data: reducing bias in clustered observational studies. *Journal of Educational and Behavioral Statistics*, *46*(5), 567–597. https://doi.org/10.3102/10769986211032411.

Keele, L., Lenard, M., Miratrix, L., & Page, L. (2023). A software tutorial for matching in clustered observational studies. *Observational Studies*, *9*(4), 73–96. https://doi.org/10.1353/obs.2023.a906624.

Kleickmann, T., & Anders, Y. (2011). Lernen an der Universität. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 305–316). Waxmann.

Klemm, K. (2022a). *Entwicklung von Lehrkräftebedarf und -angebot in Deutschland bis 2035. Aktualisierte Expertise mit Bezug auf die von der Kultusministerkonferenz (KMK) am 14. März 2022 veröffentlichte Berechnung „Lehrkräfteeinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2021–2035". Im Auftrag des Verbandes Bildung und Erziehung (VBE).* Essen.

Klemm, K. (2022b). Zum Mangel an Grundschullehrkräften. In I. Mammes & C. Rotter (Hrsg.), *Professionalisierung von Grundschullehrkräften. Kontext, Bedingungen und Herausforderungen* (S. 304–314). Klinkhardt. https://doi.org/10.25656/01:24635.

Klemm, K., & Zorn, D. (2019). *Steigende Schülerzahlen im Primarbereich: Lehrkräftemangel deutlich stärker als von der KMK erwartet*. Bertelsmann.

Klemm, K., & Zorn, D. (2024). *Weniger Geburten, mehr Lehrkräfte. Spielraum für die Grundschulentwicklung.* Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/user_upload/2024_Studie_Weniger_Geburten_-_mehr_Lehrkraefte.pdf

Klieme, E. (2019). Unterrichtsqualität. In M. Harring, C. Rohlfs & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 393–408). Waxmann.

Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur" und Unterrichtsgestaltung im internationalen Vergleich. In E. Klieme & J. Baumert (Hrsg.), *TIMSS – Impulse für Schule und Unterricht* (S. 43–57). BMBF.

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Waxmann.

Klusmann, U., Richter, D., & Lüdtke, O. (2016). Teachers' emotional exhaustion is negatively related to students' achievement: Evidence from a large-scale assessment study. *Journal of Educational Psychology*, *108*(8), 1193–1203. https://doi.org/10.1037/edu0000125.

KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2013). *Gestaltung von Sondermaßnahmen zur Gewinnung von Lehrkräften zur Unterrichtsversorgung (Beschluss der Kultusministerkonferenz vom 05.12.2013)*. Berlin: KMK. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/2013_12_05-Gestaltung-von-Sondermassnahmen-Lehrkraefte.pdf. Zugegriffen: 19. Sep. 2024

KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2023). *Einstellung von Lehrkräften 2022*. KMK.
KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2024). *Lehrkräfteeinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2023–2035. Zusammengefasste Modellrechnungen der Länder*. KMK.
KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2025). *Einstellung von Lehrkräften 2024*. KMK.
König, J., Ligtvoet, R., Klemenz, S., & Rothland, M. (2017). Effects of opportunities to learn in teacher preparation on future teachers' general pedagogical knowledge: analyzing program characteristics and outcomes. *Studies in Educational Evaluation*, *53*, 122–133. https://doi.org/10.1016/j.stueduc.2017.03.001.
Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM*, *40*(5), 873–892. https://doi.org/10.1007/s11858-008-0141-9.
Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. UTB.
Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse aus dem Forschungsprogramm COACTIV* (S. 55–68). Waxmann.
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.
Künzli David, C., Bertschy, F., Leonhard, T., & Müller, C. (2020). Universaldilettant* innen, defizitäre Generalist* innen? Herausforderungen für die Primarstufenausbildung. *journal für lehrerInnenbildung jlb*, *20*(3), 86–93.
Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: student achievement and motivation in middle school. *Education Finance and Policy*, *12*(2), 241–279.
Landesverband Sachsen-Anhalt, G. E. W. (2024). *Seiteneinstieg: Unverzichtbar & Reformbedürftig. Neun Forderungen zur Verbesserung des Seiteneinstiegs in Sachsen-Anhalt*. Gewerkschaft Erziehung und Wissenschaft.
Lenske, G. (2016). *Schülerfeedback in der Grundschule: Untersuchung zur Validität*. Waxmann.
Liegmann, A. B., Kötter-Mathes, S., & van Ackeren, I. (2022). Position des Grundschullehramts im Bildungssystem und Konsequenzen für die Ausbildung von Grundschullehrkräften. In I. Mammes & C. Rotter (Hrsg.), *Professionalisierung von Grundschullehrkräften. Kontext, Bedingungen und Herausforderungen* (S. 36–52). Klinkhardt.
Lucksnat, C., Richter, E., Klusmann, U., Kunter, M., & Richter, D. (2020). Unterschiedliche Wege ins Lehramt – unterschiedliche Kompetenzen? Ein Vergleich von Quereinsteigern und traditionell ausgebildeten Lehramtsanwärtern im Vorbereitungsdienst. *Zeitschrift für pädagogische Psychologie*, *36*(4), 263–278. https://doi.org/10.1024/1010-0652/a000280.
Lucksnat, C., Richter, E., Schipolowski, S., Hoffmann, L., & Richter, D. (2022). How do traditionally and alternatively certified teachers differ? A comparison of their motives for teaching, their well-being, and their intention to stay in the profession. *Teaching and Teacher Education*, *117*, 103784. https://doi.org/10.1016/j.tate.2022.103784.
Lucksnat, C., Richter, E., Henschel, S., Hoffmann, L., Schipolowski, S., & Richter, D. (2024). Comparing the teaching quality of alternatively certified teachers and traditionally certified teachers: findings from a large-scale study. *Education Assessment Evaluation Accountability*, *36*, 75–106. https://doi.org/10.1007/s11092-023-09426-1.
Marsh, H., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A., Abduljabbar, A., & Köller, O. (2012). Classroom climate and contextual effects: conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, *47*(2), 106–124. https://doi.org/10.1080/00461520.2012.670488.
McCarty, W. L., & Dietz, D. (2011). Alternative teacher certification: the case for transition to teaching. *Journal of Applied Learning in Higher Education*, *3*, 45–57.
Melzer, W., Pospiech, G., & Gehrmann, A. (2014). *QUER – Qualifikationsprogramm für Akademiker zum Einstieg in den Lehrerberuf. Abschlussbericht 2014*. Technische Universität Dresden.
MSW = Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2012). *Allgemeine Dienstordnung für Lehrerinnen und Lehrer, Schulleiterinnen und Schulleiter an öffentlichen Schulen (ADO)*. MSW. RdErl. d. Ministeriums für Schule und Weiterbildung v. 18.06.2012 (ABl. NRW. S. 384). Stand 01.06.2025

Muthén, L. K., & Muthén, B. (2024). *Mplus user's guide*. Muthén & Muthén.
Oettinghaus, L., Lamprecht, J., & Korneck, F. (2014). Analyse der professionellen Kompetenz von Referendaren. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht* (Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung, München, S. 135–137). IPN.
Olbrisch, M. (2024). *Warum viele Quereinsteiger schnell wieder aufgeben. DER SPIEGEL*. https://www.spiegel.de/panorama/bildung/lehrermangel-warum-viele-quereinsteiger-an-schulen-schnell-wieder-aufgeben-a-d96f82f1-efe6-4876-985c-a13e0abd6a8f. Zugegriffen: 27. März 2024.
Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS). Dimensions overview, Pre-K-3*. Paul H. Brookes.
Pimentel, M. S. (2023). *Package 'matchMulti'*.
Pimentel, S. D., & Keele, L. (2018). Optimal multilevel matching using a network algorithm. *Journal of the American Statistical Association*, *113*(521), 35–47. https://doi.org/10.1080/01621459.2017.1395362.
Porsch, R. (2016). Fachfremd unterrichten in Deutschland: Definition–Verbreitung–Auswirkungen. *DDS–Die Deutsche Schule*, *108*(1), 9–32.
Porsch, R., & Wendt, H. (2015). Welche Rolle spielt der Studienschwerpunkt von Sachunterrichtslehrkräften für ihre Selbstwirksamkeit und die Leistungen ihrer Schülerinnen und Schüler? In H. Wendt, T. Stubbe, K. Schwippert & W. Bos (Hrsg.), *IGLU & TIMSS. 10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011* (S. 161–183). Waxmann.
Porsch, R., & Wendt, H. (2016). Aus- und Fortbildung von Mathematik- und Sachunterrichtslehrkräften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland* (S. 189–204). Waxmann.
Porsch, R., & Whanell, R. (2019). Out-of-field teaching affecting students and learning: what is known and unknown. In L. Hobbs & G. Törner (Hrsg.), *Examining the Phenomenon of "Teaching out-of-field": International perspectives on teaching as a non-specialist* (S. 179–191). Springer.
Porsch, R., & Wilden, E. (2022). Teaching english out-of-field in primary school: differences in professional characteristics and effects on self-assessed instructional quality. In L. Hobbs & R. Porsch (Hrsg.), *Out-of-field teaching across teaching disciplines and contexts* (S. 117–134). Springer.
Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM*, *50*(3), 535–553. https://doi.org/10.1007/s11858-018-0946-0.
Praetorius, A.-K., & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, *49*, 167–188. https://doi.org/10.1007/s42010-021-00119-6.
Praetorius, A.-K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Taut, S., & Charalambous, C. (2020). Towards developing a theory of generic teaching quality. Origin, current status, and necessary next steps regarding the Three Basic Dimensions Model. *Zeitschrift für Pädagogik*, *66*(1), 15–36.
R Core Team (2020). *R: A language and environment for statistical computing. In Version 4.0.2*. R Foundation for Statistical Computing.
Redding, C., & Smith, T. M. (2016). Easy in, easy out: are alternatively certified teachers turning over at increased rates? *American Educational Research Journal*, *53*(4), 1086–1125. https://doi.org/10.3102/0002831216653206.
Richter, D., & Marx, A. (2019). Quereinsteigende und grundständig ausgebildete Lehrkräfte im Vorbereitungsdienst in Berlin: Eine vergleichende Analyse ihres Einsatzortes. *Zeitschrift für Erziehungswissenschaft*, *22*(6), 1385–1395. https://doi.org/10.1007/s11618-019-00915-y.
Richter, D., Kuhl, P., Reimers, H., & Pant, H. A. (2012). Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 237–250). Waxmann.
Richter, D., Kuhl, P., Haag, N., & Pant, H. A. (2013). Aspekte der Aus- und Fortbildung von Mathematik- und Naturwissenschaftslehrkräften im Ländervergleich. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012* (S. 367–390). Waxmann.
Richter, D., Marx, A., & Zorn, D. (2018). *Lehrkräfte im Quereinstieg: sozial ungleich verteilt?* Bertelsmann. https://doi.org/10.11586/2018048.
Richter, D., Becker, B., Hoffmann, L., Busse, J., & Stanat, P. (2019). Aspekte der Aus- und Fortbildung von Lehrkräften im Fach Mathematik und in den naturwissenschaftlichen Fächern. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018. Mathematische*

und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe im zweiten Ländervergleich (S. 385–410). Waxmann.

Richter, D., Huang, Y., & Richter, E. (2024). Ungleichheiten in der Lehrkräfteversorgung: Eine Analyse zur Verteilung qualifizierten Lehrpersonals auf Schulen mit unterschiedlicher Schülerschaft und verschiedenen sozio-ökonomischen Kontexten. *Zeitschrift für Erziehungswissenschaft*, *27*(6), 1491–1517. https://doi.org/10.1007/s11618-024-01271-2.

Richter, E., Lucksnat, C., Redding, C., & Richter, D. (2022). Retention intention and job satisfaction of alternatively certified teachers in their first year of teaching. *Teaching and Teacher Education*, *114*, 103704. https://doi.org/10.1016/j.tate.2022.103704.

Richter, E., Klusmann, U., & Richter, D. (2023). Berufliche Herausforderungen und Beanspruchungserleben von Lehrkräften im Seiteneinstieg: *Die Rolle der Mentorinnen und Mentoren*. *Erziehung und Unterricht*, *173*(7/8), 605–615.

Rjosk, C., Hoffmann, L., Richter, D., Marx, A., & Gresch, C. (2017). Qualifikation von Lehrkräften und Einschätzungen zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016.Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 335–353). Waxmann.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.2307/2335942.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38. https://doi.org/10.1080/00031305.1985.10479383.

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.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350.

Sachse, K. A., Weirich, S., Becker, B., Hafiz, N. J., Schneider, R., & Schipolowski, S. (2022). Methodische Grundlagen: Anlage, Durchführung und Auswertung des IQB-Bildungstrends 2021. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 233–258). Waxmann.

Scholz, L. A., Beese, C., Jentsch, A., Sierau, J., & Steffensky, M. (2024). Unterrichten in der Grundschule: Unterrichtsqualität und Lehrkräftefortbildung. In K. Schwippert, D. Kasper, B. Eickelmann, F. Goldhammer, O. Köller, C. Selter & M. Steffensky (Hrsg.), *TIMSS 2023. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 305–326). Waxmann. https://doi.org/10.31244/9783830999591.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for genralized causal inference*. Houghton Mifflin Company.

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann.

SWK = Ständige Wissenschaftliche Kommission der Kultusministerkonferenz (SWK) (2023). *Empfehlungen zum Umgang mit dem akuten Lehrkräftemangel. Stellungnahme der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz (SWK)*. https://doi.org/10.25656/01:26372.

Tatto, M. T., & Senk, S. (2011). The mathematics education of future primary and secondary teachers: methods and findings from the teacher education and development study in mathematics. *Journal of Teacher Education*, *62*(2), 121–137. https://doi.org/10.1177/0022487110391807.

Valtin, R. (2020). Schreibenlernen erfordert mehr als „phonologische Bewusstheit". Eine Längsschnittstudie zur Entwicklung sprachanalytischer Fähigkeiten von Schulanfängern. In R. Valtin & I. Naegele (Hrsg.), *„Schreiben ist wichtig!". Grundlagen und Beispiele für kommunikatives Schreiben(lernen)* (4. Aufl., S. 23–53). Arbeitskreis Grundschule e. V..

Van Buuren, S. (2018). *Flexible imputation of missing data* (2. Aufl.). CRC Press.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03.

Van Den Ham, A. K., & Heinze, A. (2018). Does the textbook matter? Longitudinal effects of textbook choice on primary school students' achievement in mathematics. *Studies in Educational Evaluation*, *59*, 133–140. https://doi.org/10.1016/j.stueduc.2018.07.005.

Van Overschelde, J. P. (2022). Value-lost: The hidden cost of teacher misassignment. In L. Hobbs & R. Porsch (Hrsg.), *Out-of-field teaching across teaching disciplines and contexts* (S. 49–70). Springer.

Voss, T., & Wittwer, J. (2020). Unterricht in Zeiten von Corona: Ein Blick auf die Herausforderungen aus der Sicht von Unterrichts- und Instruktionsforschung. *Unterrichtswissenschaft*, *48*, 601–627. https://doi.org/10.1007/s42010-020-00088-2.

Weinhold, S. (2009). Effekte fachdidaktischer Ansätze auf den Schriftspracherwerb in der Grundschule. Lese-und Rechtschreibleistungen in den Jahrgangsstufen 1–4. *Didaktik Deutsch: Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur*, *14*(27), 53–75.

Westphal, A., Schulze, A., Schlesier, J., & Lohse-Bossenz, H. (2024). Who thinks about dropping out and why? Cognitive and affective-motivational profiles of student teachers explain differences in their intention to quit their teaching degree. *Teaching and Teacher Education*, *150*, 104718. https://doi.org/10.1016/j.tate.2024.104718.

Zaruba, N., Richter, E., Porsch, R., & Richter, D. (2023). Was trägt zu einem erfolgreichen Berufseinstieg bei Seiteneinsteiger:innen bei? Eine Studie zur Relevanz pädagogischer Vorerfahrungen und sozialer Unterstützung im ersten Berufsjahr. In R. Porsch & P. Gollub (Hrsg.), *Professionalisierung von Lehrkräften im Beruf. Stand und Perspektiven der Lehrkräftebildung und Professionsforschung* (S. 345–365). Waxmann.

Ziegler, C., & Richter, D. (2017). Der Einfluss fachfremden Unterrichts auf die Schülerleistung: Können Unterschiede in der Klassenzusammensetzung zur Erklärung beitragen? *Unterrichtswissenschaft*, *45*(2), 136–155.

Ziegler, C., Richter, D., & Hartung-Beck, V. (2019). Entwicklung des Anteils fachfremden Unterrichts an Berliner Schulen: Eine Untersuchung zur Identifizierung verschiedener Verlaufsmuster. *Die Deutsche Schule*, *14*, 121–139. https://doi.org/10.25656/01:17793.

Ziegler, C., Richter, D., & Hartung-Beck, V. (2022). Die Relevanz von Quer- und Seiteneinsteigenden für den Lernerfolg von SchülerInnen: Eine empirische Analyse auf Basis des IQB-Ländervergleiches. *Zeitschrift für Pädagogik*, *68*(5), 587–608.

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