



Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft

Lukas Begrich · Anna-Katharina Praetorius · Jasmin Decristan ·
Benjamin Fauth · Richard Göllner · Christian Herrmann ·
Marc Kleinknecht · Sandy Taut · Mareike Kunter

Eingegangen: 4. Oktober 2022 / Überarbeitet: 20. Januar 2023 / Angenommen: 21. Januar 2023 / Online
publiziert: 23. Februar 2023
© Der/die Autor(en) 2023

Zusammenfassung Die Unterrichtsqualitätsforschung hat in den vergangenen Jahrzehnten viel zum Verständnis der Merkmale beitragen, die einen qualitätsvollen Unterricht auszeichnen. Sie hat einflussreiche Modelle hervorgebracht, die eine gemeinsame Grundlage für die Erforschung des Unterrichts bereitstellen. Gleichzeitig bringen gesellschaftliche Veränderungsprozesse eine Neuorientierung des schulischen Unterrichts mit sich, und es lässt sich fragen, ob die Unterrichtsqualitätsforschung mit ihren bisherigen Ansätzen in der Lage ist, in Zukunft zum Gelingen von Unterricht beizutragen. Im Rahmen einer Zukunftswerkstatt hat das Leibniz-Netzwerk Unterrichtsqualitätsforschung daher einen längerfristig angelegten Prozess gestartet, um sich mit dem Unterricht der Zukunft zu beschäftigen und zu versuchen, drängende Fragen und Handlungsfelder für die Unterrichtsqualitätsforschung zu identifizieren. Der vorliegende Beitrag stellt die Ergebnisse dieses initialen Austauschs dar. Zudem wird ein Einblick in die gegenwärtige Arbeit im Netzwerk gegeben, die aufbauend auf den Ergebnissen der Zukunftswerkstatt eine Weiterentwicklung der Unterrichtsqualitätsforschung anstrebt.

Schlüsselwörter Unterrichtsqualität · Zukunft · Innovation · Reflexion

Der vorliegende Beitrag wurde von den Autor*innen verfasst, enthält jedoch auch Gedankengut folgender Personen: Christiane Bertram, Esther Brunner, Aiso Heinze, Benjamin Herbert, Miriam Hess, Doris Holzberger, Stefan Keller, Thilo Kleickmann, Eckhard Klieme, Carmen Köhler, Susanne Kuger, Linda Lenske, Anke Lindmeier, Alexander Naumann, Christine Pauli, Katrin Rakoczy, Patrick Schreyer, Mirjam Steffensky, Svenja Vieluf, Wolfgang Wagner, Iris Winkler.

Ausführliche Informationen zu den Autoren befinden sich auf der letzten Seite dieses Artikels.

What to do? Perspectives on teaching quality research for the future

Abstract Over the past decades, research on teaching quality has substantially contributed to the understanding of high quality teaching. It has brought forth influential models that provide a common ground for research on teaching. At the same time, societal change requires a reorientation of teaching in schools. It seems questionable whether the current approaches of teaching quality research can contribute to the success of teaching in the future. Therefore, based on a future workshop method, the Leibniz Network of Research on Teaching Quality has started a longer-term working process to consider future developments of teaching and to identify urgent questions and fields of action for research on teaching quality. The present paper outlines the results of this initial exchange. Furthermore, it gives insight into the current work within the network that aims at an advancement of research on teaching quality based on the results of the workshop.

Keywords Teaching quality · Future · Innovation · Reflection

1 Einleitung

Eine hohe Qualität schulischen Unterrichts hat eine zentrale Funktion für die Entwicklung von Kindern und Jugendlichen und ist daher von hoher gesellschaftlicher und volkswirtschaftlicher Relevanz. Dementsprechend gibt es ein großes Interesse daran, die Qualität von Unterricht wissenschaftlich zu fassen und zu verbessern. Während die wissenschaftliche Auseinandersetzung lange Zeit innerhalb unterschiedlicher Disziplinen wie beispielsweise der Allgemeinen Didaktik, der Lern- oder Instruktionspsychologie oder den jeweiligen Fachdidaktiken erfolgte, ist die Forschung der letzten Jahre zunehmend interdisziplinär geworden (Schalk und Stern 2018). Unabhängig von der disziplinärer Zuordnung haben sich über Jahrzehnte hinweg diverse Forschungsansätze entwickelt, die sich aus verschiedenen inhaltlichen und methodischen Perspektiven der Frage nach der Qualität von Unterricht widmen (Klieme 2019; Praetorius et al. 2021; Terhart 2002). Dabei wird der Begriff der Unterrichtsqualität hauptsächlich in der quantitativ-empirischen Unterrichtsforschung verwendet. In Arbeiten mit qualitativ-empirischem Einschlag wird er in der Regel nicht genutzt, obgleich Fragen der Qualität von Unterricht auch hier eine Rolle spielen (Praetorius et al. 2021). In der quantitativen Unterrichtsqualitätsforschung bestand das Hauptinteresse bislang darin, diejenigen Unterrichtsmerkmale zu identifizieren, zu beschreiben und bezogen auf ihre Wirkungen zu untersuchen, die für auf das Erreichen unterrichtlicher Ziele besonders effektiv sind und dabei gleichzeitig normativ-theoretischen Setzungen zu gutem Unterricht entsprechen (Berliner 2005; Fenstermacher und Richardson 2005). Die Beiträge der quantitativ-empirischen Unterrichtsqualitätsforschung konnten in den vergangenen Jahrzehnten viel zum Verständnis von wirksamen Unterricht beitragen. So wurden im deutschsprache-

chigen Raum¹ auf Grundlage groß angelegter Studien (z. B. COACTIV, Kunter und Voss 2013; TEDS-Unterricht, Blömeke et al. 2022; IGEL, Fauth et al. 2018) Modelle vorgeschlagen, die eine systematische und sparsame Betrachtung sowohl der Einflussfaktoren, die auf den Erfolg des Unterrichts Einfluss nehmen, als auch der Lehr-Lernprozesse, die im Unterricht ablaufen, ermöglichen. Solche Modelle stellen eine gemeinsame Grundlage für empirische Untersuchungen des Unterrichts dar und begünstigen damit einen kumulativen Erkenntnisgewinn. Mittlerweile existieren einige Übersichten über den entsprechenden Stand der Unterrichtsqualitätsforschung (Praetorius et al. 2020a; Klieme 2019; Kunter und Ewald 2016; Steffens und Messner 2019).

Gleichzeitig steht die Unterrichtsqualitätsforschung unter Innovationsdruck, denn viele ihrer Erkenntnisse beziehen sich auf Unterrichtswirklichkeiten, die sich in der Praxis komplexer darstellen, sich verändern oder bereits der Vergangenheit angehören. Da gesellschaftliche Rahmenbedingungen immer auch einen Einfluss auf den schulischen Unterricht ausüben, erscheint es von zentraler Bedeutung, dass die Unterrichtsqualitätsforschung die bereits sichtbaren Veränderungen in den Blick nimmt sowie zukünftige Veränderungen antizipiert. Ausgehend davon sollte sie ihren Fokus auf die Beantwortung der drängenden Fragen richten, denen die Unterrichtspraxis infolge dieser Veränderungen gegenüber steht. Nur so kann Unterrichtsqualitätsforschung in Zukunft zum Gelingen von Unterricht beitragen. Dazu gehört, die Forschungspraxis kritisch zu reflektieren und über neue Aspekte nachzudenken, etwa bei der Gewinnung von Fragestellungen. Die notwendigen Innovationen können nicht von Einzelnen identifiziert und auf den Weg gebracht werden. Vielmehr braucht es dazu einen gezielten Austausch zwischen Forschenden des Feldes, der den bisherigen Stand der Forschung einbezieht, aber auch deutlich darüber hinausgeht. Der vorliegende Beitrag stellt die Ergebnisse eines solchen Expert*innen-Diskurses dar, der im Leibniz-Netzwerk Unterrichtsforschung initiiert und gegenwärtig geführt wird. Das Leibniz-Netzwerk Unterrichtsforschung ist ein Verbund aus vorrangig empirisch-quantitativ arbeitenden Forschenden aus dem deutschsprachigen Raum, die sich mit Fragen der Unterrichtsqualität aus psychologischer, erziehungswissenschaftlicher und fachdidaktischer Perspektive beschäftigen (<https://unterrichtsforschung.dipf.de/de>). Ursprünglich wurde das Netzwerk gegründet, um eine umfassende Auswertung der TALIS-Videostudie Deutschland sicherzustellen (<https://www.dipf.de/de/forschung/projektarchiv/talis-videostudie-deutschland>). Bei der Gründung des Netzwerks wurde angestrebt, dass die Mitglieder einerseits ein hinreichend ähnliches Verständnis bezüglich Unterrichtsqualität und deren Erforschung haben, um gemeinsam an wichtigen Fragen arbeiten zu können. Andererseits sollte eine ausreichende Breite und Heterogenität der Zugänge sichergestellt werden. Das Netzwerk hat entsprechend nicht den Anspruch, das gesamte Feld der deutschsprachigen Unterrichtsqualitätsforschung zu repräsentieren. Es ist jedoch ein bislang

¹ Viele der im vorliegenden Beitrag angesprochenen Themen sind auch im internationalen Diskurs von Bedeutung. Allerdings sind die konzeptuell-theoretischen Auseinandersetzungen stark im deutschsprachigen Raum verankert. Um eine hinreichende Kohärenz herzustellen, fokussiert der Beitrag daher auf die Literatur aus dem deutschsprachigen Raum. Für Übersichten zur Forschung zu Unterrichtsqualität im internationalen Raum siehe z. B. Charalambous und Praetorius (2018), Darling-Hammond (2021) und Muijs et al. (2014).

einzigartiger Versuch im deutschsprachigen Raum, in einem größeren Verbund theoretische und methodische Grundfragen bezüglich der Unterrichtsqualität und ihrer Erfassung zu bearbeiten, um die Unterrichtsqualitätsforschung gemeinsam konzeptuell und methodisch voranzubringen (Themenheft in der Zeitschrift für Pädagogik; Praetorius et al. 2020a). Diesem Ziel haben sich auch der vorliegende Beitrag und die darin beschriebene Zukunftswerkstatt verschrieben.

Der vorliegende Beitrag stellt zunächst schlaglichtartig zentrale Erkenntnisse der Unterrichtsqualitätsforschung dar (Abschn. 2). Anschließend erfolgt der Bericht einer sogenannten Zukunftswerkstatt, die im Leibniz-Netzwerk Unterrichtsforschung durchgeführt wurde. In dieser Zukunftswerkstatt diskutierten wir erstens, welche möglichen Veränderungen des schulischen Unterrichts in den nächsten 15 Jahren zu erwarten sind und zweitens, was die Unterrichtsqualitätsforschung leisten muss, um den schulischen Unterricht angesichts dieser Veränderungen in Zukunft unterstützen zu können (Abschn. 3). Die Ergebnisse der Zukunftswerkstatt bilden den Ausgangspunkt für die gegenwärtige Arbeit des Netzwerks, in der wir uns mit konkreten Innovationsmöglichkeiten für die Unterrichtsqualitätsforschung auseinandersetzen (Abschn. 4). Wie das abschließende Gesamtfazit zeigt, gilt es, gemeinsam als wissenschaftliche Community sowohl theoretische als auch methodische Prämissen der Unterrichtsqualitätsforschung zu überdenken.

Der Anspruch dieses Beitrags ist es nicht, die zukünftigen Herausforderungen für den schulischen Unterricht umfassend identifiziert zu haben. Genauso wenig sollen fertige Lösungen für die angesprochenen Herausforderungen präsentiert werden. Stattdessen möchten wir Forschende, die sich mit Unterrichtsqualität beschäftigen, anregen, sich mit den aufgeworfenen Fragen zu beschäftigen sowie weitere Fragen zu entwickeln.

2 Errungenschaften und Herausforderungen der Unterrichtsqualitätsforschung im deutschsprachigen Raum

2.1 Errungenschaften der Forschung zu Unterrichtsqualität

Die deutschsprachige Unterrichtsqualitätsforschung hat in der Vergangenheit wichtige Erkenntnisse hervorgebracht und konnte somit vielfältige Impulse für Forschung und Praxis liefern. Sie steht aber gleichzeitig vor inhaltlichen und methodischen Herausforderungen, wie im Folgenden beschrieben wird. Bezüglich der Identifikation von Merkmalen und Gelingensbedingungen qualitativvollen Unterrichts sind in der deutschsprachigen Unterrichtsqualitätsforschung deutliche Fortschritte zu verzeichnen. Ein gemeinsamer theoretischer Nenner der Unterrichtsqualitätsforschung ist ein Verständnis von Unterricht im Sinne einer Angebots-Nutzungs-Konzeption (Fend 1981; Helmke 2003; Vieluf et al. 2020; Seidel 2014). Angebots-Nutzungs-Modelle betonen, dass der Erfolg von Unterricht nicht nur von einem qualitativvollen Unterrichtsangebot, sondern auch von der individuell variierenden Nutzung des Angebots durch die Lernenden abhängt. Die Qualität des Unterrichtsangebots bzw. die Qualität der Nutzung hängen wiederum von individuellen Voraussetzungen der Lehrenden und Lernenden sowie von Kontextfaktoren (Bildungssystem, Schule

etc.) ab. Die aktive und gestalterische Rolle, die den Lernenden in Hinblick auf ihre Lernprozesse zugesprochen wird, spiegelt dabei die Dominanz konstruktivistischer Vorstellungen vom Lernen wider, die in der deutschsprachigen Unterrichtsqualitätsforschung derzeit als geteilte normative Setzung betrachtet werden können (Lindmeier und Heinze 2020; Praetorius et al. 2021). Auch wenn sich die verschiedenen Angebots-Nutzungs-Modelle zum Teil deutlich unterscheiden (Abschn. 2.2) und die Frage nach einer angemessenen empirischen Untersuchung von Angebot und Nutzung noch nicht hinreichend geklärt ist (Alp Christ et al. 2022; Vieluf 2022), können sie im Sinne einer Weiterentwicklung des klassischen Prozess-Produkt-Paradigmas als Errungenschaft der deutschsprachigen Unterrichtsqualitätsforschung betrachtet werden.

Ein zweiter theoretischer Konsens in der deutschsprachigen Unterrichtsqualitätsforschung ist die Unterscheidung zwischen Oberflächen- und Tiefenmerkmalen des Unterrichts (Decristan et al. 2020; Oser und Baeriswyl 2001). Der Oberfläche werden in der Regel leicht beobachtbare Unterrichtsmerkmale wie z. B. die gewählte Sozialform zugeordnet, während sich Tiefenmerkmale auf die lernbezogenen Interaktionen zwischen Lehrkraft, Schüler*innen und Lerngegenstand, die sog. Lehr-Lernprozesse beziehen (Kunter und Ewald 2016). Die Unterscheidung ist vor allem als Heuristik sinnvoll und hilft, die konkreten Mechanismen, über die Unterrichtsmerkmale auf Lernprozesse Einfluss nehmen, mitzudenken (für eine eingehende Aufarbeitung des Themas, siehe Decristan et al. 2020; Pauli 2020). Ausgehend von diesen Überlegungen ist ein zentrales Anliegen der Unterrichtsqualitätsforschung, insbesondere die relevanten Tiefenmerkmale des Unterrichts zu identifizieren und messbar zu machen. Nachdem dieses Bemühen in den vergangenen Jahrzehnten zunächst zu mehr oder weniger umfangreichen Listen zentraler Merkmale qualitätsvollen Unterrichts führte (Einsiedler 1997; Helmke 2007; Lipowsky 2015; Steffens und Messner 2019), hat sich in der deutschsprachigen Unterrichtsqualitätsforschung mittlerweile eine sparsame Systematik etabliert, die zentrale Qualitätsmerkmale bündelt. Das Modell postuliert drei zentrale Dimensionen, die unabhängig von Fach, Schulform und Jahrgangsstufe entscheidend für die Qualität des Unterrichts sein und die kognitive sowie motivational-emotionale Entwicklung der Schüler*innen positiv beeinflussen sollen (Klieme 2006, 2019; Praetorius et al. 2020c). Dem Modell der drei Basisdimensionen zufolge lässt sich Unterrichtsqualität im Wesentlichen über das Gelingen der Klassenführung, einer erfolgreichen kognitiven Aktivierung der Lernenden sowie deren konstruktiver Unterstützung durch die Lehrkraft definieren (Klieme 2006; Praetorius et al. 2018). Obwohl zu diskutieren ist, ob das Modell der Basisdimensionen umfassend genug ist und der Fachspezifität von Unterrichtsqualität hinreichend Rechnung trägt (Abschn. 2.2 und 4.5), stellt es neben den Angebots-Nutzungs-Modellen und der Unterscheidung zwischen Oberflächen- und Tiefenmerkmalen eine dritte wichtige Errungenschaft der deutschsprachigen Unterrichtsqualitätsforschung dar. Viele Forschungsgruppen im deutschsprachigen Raum orientieren sich mittlerweile an diesen Dimensionen, sodass sich – zumindest zu einem gewissen Grad – eine gemeinsame Sprache und Struktur in der Erforschung von Unterrichtsqualität etabliert hat. Auch im internationalen Vergleich stellt das Modell der Basisdimensionen einen der wenigen Ansätze dar, die explizite Annahmen über die spezifische Wirkungsweise verschiedener Unterrichtsmerkmale formulieren und dabei auch die

Mediation dieser Wirkungsweisen im Sinne der Angebots-Nutzungs-Modelle berücksichtigen. Zudem wurden in den letzten Jahren Anstrengungen unternommen, das Modell der Basisdimensionen theoretisch weiterzuentwickeln, wenngleich das Modell noch nicht alle Kriterien einer wissenschaftlichen Theorie erfüllt (Praetorius et al. 2020c). Aufgrund zahlreicher internationaler Publikationen ist das Modell der Basisdimensionen der Unterrichtsqualität zudem auch in der internationalen Fachwelt bekannt (Decristan et al. 2015; Dorfner et al. 2018; Fauth et al. 2021a; Herbert et al. 2022; Klieme et al. 2009; Kunter et al. 2013; Lipowsky et al. 2009; Mu et al. 2022; Praetorius et al. 2018).

Auch im Hinblick auf die positive Einflussnahme auf die Unterrichtspraxis (Terhart 2020), sind erste Fortschritte zu verzeichnen. Im Rahmen von Interventionsstudien, durch die Entwicklung von Feedbackinstrumenten oder durch den Einsatz von Simulationen werden konkrete Versuche unternommen, die Unterrichtsqualität in der Praxis sowie die Ausbildung im Rahmen des Lehramtsstudiums zu verbessern (Decristan et al. 2015; Fauth et al. 2021b; Hardy et al. 2011; Kellermann et al. 2022; Huang et al. 2021; Klieme et al. 2010). Zudem sind die Erkenntnisse der Unterrichtsqualitätsforschung z. B. in Form von Lehrbüchern verstärkt in die Lehrkräfte(aus)bildung eingegangen (Helmke 2022; Kunter und Trautwein 2013; Meyer und Junghans 2022; siehe auch das Themenheft zur kognitiven Aktivierung in der Zeitschrift PÄDAGOGIK, Schnack 2021). Darüber hinaus lassen sich verstärkt Versuche verzeichnen, Forschungsergebnisse gezielt für Bildungspraktiker*innen aufzuarbeiten und in kompakter Form zugänglich zu machen (Fauth und Leuders 2018 oder <https://www.forschungsmonitor-schule.de/>).

2.2 Aktuelle Herausforderungen für die Unterrichtsqualitätsforschung

2.2.1 Konzeptuelle Herausforderungen

Obwohl Angebots-Nutzungs-Modelle und das Modell der drei Basisdimensionen in der deutschsprachigen Unterrichtsqualitätsforschung eine breite Rezeption erfahren, fällt bei genauerer Betrachtung auf, dass die Modellentwicklung ein laufender Prozess ist, in dem noch viele Fragen beantwortet und Annahmen empirisch geprüft werden müssen. Terhart (2020) stellt aus diesem Grund sogar die Frage, ob die Modelle zu früh ihren Weg in die Lehrbücher gefunden haben. So liegen Angebots-Nutzungs-Modelle in zahlreichen Varianten vor, die sich u. a. im Hinblick auf ihr Verständnis von Unterricht, ihre Komplexität und v. a. die Konzeptualisierung von Angebot und Nutzung und deren Zusammenhang deutlich unterscheiden (Fend 1981; Helmke 2003; Klieme et al. 2006; Kunter und Trautwein 2013; Lipowsky 2006; Meyer 2015; Seidel 2014). Diese Unterschiede können dazu führen, dass je nach zugrunde gelegter Modellvariante Studien anders konzeptualisiert, Forschungsergebnisse anders eingeordnet und andere Rückschlüsse bezüglich möglicher Verbesserungsstrategien des Unterrichts gezogen werden (Vieluf et al. 2020). Vieluf et al. (2020) schlagen daher ein integriertes Angebots-Nutzungs-Modell vor, in dem unter anderem eine reziproke Beziehung zwischen dem Unterrichtsangebot und dessen Nutzung angenommen wird. Das Unterrichtsangebot wird durch die Schüler*innen mitgestaltet, wobei deren Wahrnehmung und Interpretation des Unterrichts eine vermittelnde Rol-

le spielt. Das Potenzial solcher integrierten Modelle besteht darin, das Verständnis des Gegenstands in einem Forschungsfeld zu vereinheitlichen und kumulativen Erkenntnisgewinn zu begünstigen, indem sie eine Grundlage für die Einordnung einer großen Breite an empirischen Befunden bereitstellen (Bikner-Ahsbahs und Prediger 2010). Dies kann jedoch nur gelingen, wenn genügend Forschende den Vorschlag aufgreifen und weiterentwickeln. Ob dies für das integrierte Angebots-Nutzungsmodell von Vieluf und Kolleg*innen der Fall ist, wird die Zukunft zeigen (Vieluf et al. 2020).

Das Modell der drei Basisdimensionen besticht durch seine Sparsamkeit und seine generische Konzeption, also der postulierten Geltung über Fächer, Schulformen und Jahrgangsstufen hinweg. Gleichzeitig besteht bei seiner Anwendung in der Forschung die Gefahr einer Verengung. Obwohl das Modell im Anschluss an die Angebots-Nutzungsmodelle explizit die Annahme formuliert, dass das Unterrichtsangebot nur vermittelt über die Nutzung der Lernenden wirken kann (Praetorius et al. 2018), beschränken sich die meisten Forschenden auf die Erforschung des direkten Zusammenhangs zwischen Unterrichtsangebot und verschiedenen Schüler*innenoutcomes (Praetorius und Kleickmann 2022). Auf diese Weise läuft die Unterrichtsqualitätsforschung Gefahr, trotz ihrer theoretischen Kenntnis des komplexen Wirkungsgefüges, das über den Erfolg von Unterricht entscheidet, forschungspraktisch im Prozess-Produkt-Paradigma verhaftet zu bleiben. Ein Grund hierfür ist die heterogene theoretische Konzeptualisierung der Nutzung, wodurch ebenfalls unklar bleibt, wie diese in empirischen Studien erfasst werden kann (Alp Christ et al. 2022; Praetorius und Kleickmann 2022). Hinzu kommt, dass keine Einigkeit darüber herrscht, welche Unterrichtsprozesse der Angebotsseite des Unterrichts zuzuordnen sind (Vieluf 2022). So beschränken einige Forscher*innen den Bereich des Unterrichtsangebots auf das Lehrkrafthandeln (Brühwiler und Blatchford 2011; Lipowsky 2015), während andere Forscher*innen auch Schüler*innenbeiträge dem Unterrichtsangebot zuordnen (Helmke 2003). Des Weiteren kann das Modell der Basisdimensionen den fälschlichen Eindruck vermitteln, alle für die Qualität entscheidenden Aspekte des Unterrichts abzudecken. So sind einerseits weitere, in anderen Modellen enthaltene Merkmale wie zum Beispiel der Bereich der Unterstützung des Übens/Konsolidierens oder formative sowie summative Beurteilungen im Modell nicht abgedeckt (Praetorius et al. 2020c; Helmke 2022). Zum anderen ist davon auszugehen, dass es zusätzliche fachspezifische Dimensionen gibt, die entscheidend für die Effektivität des Fachunterrichts sind (Lipowsky et al. 2018). Es ist also fraglich, ob die Qualität von Unterricht anhand der genannten drei Basisdimensionen hinreichend konzeptualisiert werden kann. In den letzten Jahren wurden verschiedene Vorschläge zur Erweiterung des Modells formuliert, die sich sowohl auf fachspezifische (Brunner 2017; Drollinger-Vetter 2011; Lipowsky et al. 2018; Schlesinger und Jentsch 2016; Schlesinger et al. 2018; Szogs et al. 2016) als auch auf generische (Batzel et al. 2013; Decristan et al. 2022; Kleickmann et al. 2020; Rakoczy et al. 2007; Schlesinger et al. 2018; Taut und Rakoczy 2016; Vieluf 2013) Aspekte des Unterrichtsangebots beziehen (Praetorius et al. 2020b).

Die empirische Auseinandersetzung damit, was qualitativollen Unterricht ausmacht, ist immer von normativ-theoretischen Setzungen der Forschenden mitgeprägt, die Auswirkungen auf die gewählten Untersuchungsmethoden haben. So lei-

ten normative Annahmen und Vorstellungen sowohl die Auswahl der untersuchten Unterrichtsmerkmale als auch auf die Auswahl der Zielbereiche des Unterrichts, die als Kriterium für die Effektivität des Unterrichts gewertet werden (Praetorius et al. 2021). Etwa bei der Auswahl der untersuchten Unterrichtsmerkmale lassen sich die konstruktivistischen Überzeugungen der Forschenden deutlich erkennen (Terhart 2020). Das Vorhandensein normativer Setzungen selbst ist dabei nicht problematisch. Vielmehr sind sie notwendiger Bestandteil einer Beschreibung qualitativ hochwertigen Unterrichts (Berliner 2005). Während sich die Allgemeine Didaktik über Jahrzehnte mit Normativität auseinandergesetzt hat, fällt auf, dass die eigenen normativen Grundlagen von Forschenden in der Unterrichtsqualitätsforschung nur selten thematisiert werden (Lindmeier und Heinze 2020). Auf diese Weise erscheinen die Auswahl der untersuchten Unterrichtsmerkmale sowie die der Zielbereiche des Unterrichts häufig beliebig bzw. nicht immer klar begründet. Dabei ist gerade die Auswahl der untersuchten Outcome-Variablen alles andere als trivial. Schließlich sind Annahmen bezüglich der wünschenswerten Ergebnisse des Unterrichts von gesamtgesellschaftlicher Relevanz, da sie die Frage berühren, welche Kompetenzen zukünftig für die Teilhabe am gesellschaftlichen Leben gebraucht werden. Diese Frage können Forschende nicht allein beantworten. Vielmehr gehört sie in den bildungspolitischen Diskurs. Es wäre jedoch wünschenswert, dass Forschende die Auswahl von Outcome-Variablen gründlich reflektieren.

2.2.2 Methodische Herausforderungen

Die empirische Überprüfung der Angebot-Nutzungs-Modelle und des Modells der drei Basisdimensionen wird auch durch methodische Herausforderungen erschwert. Diese betreffen zum einen die Messung und Operationalisierung der Unterrichtsqualität und zum anderen Grenzen der bisher hauptsächlich angewandten Forschungsdesigns. Bezogen auf die Messung ist offensichtlich, dass sich eine Dimension wie z. B. die konstruktive Unterstützung nicht einfach feststellen oder quantifizieren lässt. Vielmehr ist man bei der Erfassung auf Einschätzungen angewiesen. Hierzu können Lehrkräfte und Schüler*innen befragt oder geschulte Beobachter*innen eingesetzt werden. Die Forschung zeigt jedoch, dass Einschätzungen des gleichen Unterrichts aus diesen Perspektiven häufig sehr unterschiedlich ausfallen und sich teilweise nicht auf die gleichen Konstrukte beziehen (Clausen 2002; Fauth et al. 2014; Kunter und Baumert 2006). Auch wenn eine Übereinstimmung der Perspektiven aus theoretischer Sicht kaum zu erwarten ist (Fauth et al. 2020), erschwert es die Perspektivenspezifität, Ergebnisse aufeinander zu beziehen und zu synthetisieren. Unabhängig von der gewählten Perspektive ist zudem auffällig, dass die drei Basisdimensionen in verschiedenen Studien sehr unterschiedlich operationalisiert werden. Dies reicht zum Teil soweit, dass in Forschungsarbeiten ähnliche Konzepte verschiedenen Dimensionen zugeordnet werden (Praetorius et al. 2018).

Da die in den Angebots-Nutzungs-Modellen und dem Modell der Basisdimensionen angenommenen Zusammenhänge häufig komplex und auf unterschiedlichen Ebenen (Schüler*innen, Klassen, Schulen) angesiedelt sind, werden große Stichproben und komplexe Forschungsdesigns mit mehreren Messzeitpunkten benötigt, um sie statistisch zu modellieren (Köhler et al. 2020). Die Datengewinnung in der Unter-

richtsqualitätsforschung ist entsprechend sehr aufwändig. So ist es nicht erstaunlich, dass aussagekräftige Längsschnittstudien mit mehr als zwei Messzeitpunkten oder experimentelle Studien mit Randomisierung auf (der häufig primär interessierenden) Klassenebene in der Unterrichtsqualitätsforschung vergleichsweise selten realisiert werden und sich deshalb in der Literatur für viele der in den Modellen postulierten kausalen Annahmen nur schwache empirische Belege finden lassen (Praetorius et al. 2018).

Eine weitere methodische, aber auch konzeptuelle Schwierigkeit ergibt sich aus einer unklaren Generalisierbarkeit der Befunde. So wurden die Basisdimensionen der Unterrichtsqualität über lange Zeit schwerpunktmäßig in den mathematisch-naturwissenschaftlichen Fächern in der Mittelstufe erforscht (Kunter und Ewald 2016) und die Übertragbarkeit auf andere Fächer blieb offen (für weitere Ausführungen siehe 4.5). Außerdem ist festzuhalten, dass die meisten Erkenntnisse der empirisch-quantitativen Unterrichtsqualitätsforschung in klassischen Lernsettings gewonnen wurden, in denen eine Lehrkraft in der Sekundarstufe I der Regelschule mit den Schüler*innen in einem Klassenraum in einem lehrkraftgeleiteten Unterricht zusammenarbeitet. Es ist demnach fraglich, ob diese Erkenntnisse auf immer häufiger anzutreffende andere Lernsettings wie beispielsweise offene Unterrichtsformate oder die Hochschulbildung übertragen werden können.

Zudem muss angemerkt werden, dass das Potenzial von mixed-method Ansätzen in der quantitativen Unterrichtsqualitätsforschung kaum ausgeschöpft wird (Gläser-Zikuda et al. 2012).

2.2.3 *Transfer in die Praxis*

Wie oben dargestellt, haben Erkenntnisse aus der Unterrichtsqualitätsforschung ihren Weg in die Lehramtsausbildung und somit auch ansatzweise in die Praxis gefunden. Bisher sind Unterrichtsforschende jedoch häufig von einem unidirektionalen Verständnis von Transfer geleitet, bei dem Transferaktivitäten vor allem darauf abzielen, Forschungsergebnisse praxistauglich aufzubereiten und neue Wege der Vermittlung zu erschließen. Wie stark die Forschungserkenntnisse wirklich die Unterrichtsgestaltung von Lehrkräften beeinflussen, ist jedoch eine offene Frage (Decristan und Hartmann 2016). Aktuelle Transferforschung weist darauf hin, dass die Lücke zwischen Forschung und Praxis nicht einfach dadurch zu schließen ist, dass Forschende ihre Forschung besser zu den Lehrkräften transportieren (Coburn und Stein 2010; Hartmann und Decristan 2018). Als vielversprechender werden ko-konstruktive Ansätze angesehen, die darauf abzielen, dass beide Welten – Wissenschaft und Praxis – die jeweils spezifischen Ziele und Logiken der jeweils anderen Welt anerkennen und Räume zur Begegnung und des Austausches schaffen (Coburn und Stein 2010; Farley-Ripple et al. 2018). Gelingender Transfer manifestiert sich demnach nicht nur in einer veränderten Unterrichtspraxis, sondern führt auch dazu, dass sich Forschung verändert, zum Beispiel durch ein gewandeltes Verstehen des Forschungsbereichs oder bei der Auswahl von Fragen (Farley-Ripple et al. 2018). Ein solches bi-direktionales Verständnis von Transfer hat in der Unterrichtsqualitätsforschung bisher noch kaum Einzug gehalten.

2.2.4 Herausforderungen durch gesellschaftliche Veränderungen

Unterricht hängt von gesellschaftlichen Rahmenbedingungen ab und ist deshalb starken Wandlungsprozessen unterworfen. Zu den absehbaren Veränderungen zählen die zunehmende Digitalisierung des Unterrichts, eine zunehmende Heterogenität der Schülerschaft, eine Zunahme inklusiver Bildungsangebote mit multiprofessionellen Teams im Klassenraum sowie ein auf Grund des Lehrkräftemangels steigender Anteil des Lehrpersonals, das kein Lehramtsstudium absolviert hat (Autorengruppe Bildungsberichterstattung 2020). Die deutschsprachige Unterrichtsqualitätsforschung hat bislang kaum Ergebnisse vorzuweisen, die Unterrichtsqualität vor dem Hintergrund dieser Veränderungen beleuchten. So wird etwa in dem vom Leibniz-Netzwerk Unterrichtsforschung gestalteten Beiheft der Zeitschrift für Pädagogik, das den Anspruch hat, den derzeitigen Stand der deutschsprachigen empirischen Unterrichtsqualitätsforschung abzubilden, das Thema Digitalisierung nur in einem einzigen Diskussionsbeitrag aufgegriffen (Terhart 2020). Angesichts der zunehmenden Heterogenität der Schülerschaft bei gleichzeitig fortschreitender Inklusion müsste die Frage der Unterrichtsqualität weniger im Sinne einer Effektivität über ganze Klassen hinweg gestellt werden. Vielmehr müssten differenzielle Effekte von Unterricht in den Blick genommen werden, um das Unterrichtsangebot an verschiedene Schüler*innengruppen, die immer häufiger gemeinsam unterrichtet werden, empirisch begründet anpassen zu können. Solche differenziellen Unterrichtseffekte werden bislang jedoch nur verhältnismäßig selten modelliert (Kunter und Ewald 2016; für Ausnahmen Decristan et al. 2022). Insgesamt stellt sich die Frage, ob die bisherigen Erkenntnisse der Unterrichtsqualitätsforschung einen Beitrag zum Gelingen eines sich stark verändernden Unterrichts leisten können. Soll dies gewährleistet werden, muss die Unterrichtsqualitätsforschung die erwartbaren gesellschaftlichen Veränderungen und ihre Auswirkungen auf den schulischen Unterricht antizipieren, um daraus die drängenden Fragen abzuleiten, die sie bearbeiten muss, um relevant zu bleiben. Zudem zeigt die Corona-Pandemie, dass die Frage der Unterrichtsqualität auch im Hinblick auf Situationen relevant ist, in denen der schulische Unterricht in gängigen Settings nicht mehr realisiert werden kann.

3 Unterricht im Jahr 2035 – Eine Zukunftswerkstatt

Es wird deutlich, dass das Ziel der Unterrichtsqualitätsforschung, praktisch relevante Erkenntnisse über die Merkmale und Gelingensbedingungen qualitätsvollen Unterrichts zu generieren, zukünftig nur dann erreicht werden kann, wenn es gelingt, zentrale Herausforderungen des Feldes zu adressieren. Um bei der Identifizierung der Herausforderungen über ein bloßes Literaturstudium hinauszugehen, hat das Leibniz-Netzwerk Unterrichtsforschung eine Zukunftswerkstatt durchgeführt. Ziel war es zum einen, zentrale Herausforderungen der Unterrichtsqualitätsforschung im Rahmen eines freien Austauschs zu sammeln und zu vertiefen. Zum anderen sollte mit der Zukunftswerkstatt ein Impuls gegeben werden, den Blick nach vorne zu richten, um zukunftsgerichtet Lösungsmöglichkeiten für die identifizierten Herausforderungen zu erarbeiten. Die Ergebnisse der Zukunftswerkstatt bilden den Aus-

gangspunkt für einen fortlaufenden Arbeitsprozess innerhalb des Leibniz-Netzwerks Unterrichtsrechnung, in dem fünf Arbeitsgruppen konkrete Entwicklungsmöglichkeiten für die Unterrichtsqualitätsforschung ausloten.

3.1 Die Methode „Zukunftswerkstatt“ und ihre Umsetzung im Leibniz-Netzwerk Unterrichtsrechnung

Die Zukunftswerkstatt ist eine Methode, die ursprünglich als „soziales Versuchslabor, in dem alternative Zukünfte von engagierten Bürgern entworfen und durchdacht werden“ (Jungk und Müllert 1989, S. 79) entwickelt wurde. Zentrales Anliegen war es, Mitglieder von Gemeinschaften und Organisationen eine aktive Teilnahme an Veränderungsprozessen zu ermöglichen und sie damit aus der Rolle der Betroffenen von Veränderung herauszuholen (Böttger 2006). Weinbrenner (2002) beschreibt die Methode als basisdemokratisch, integrativ, ganzheitlich, kreativ, kommunikativ und provokativ. Der Ablauf einer Zukunftswerkstatt gliedert sich neben Vor- und Nachbereitung in verschiedene Phasen, in denen jeweils rational-analytische und emotional-intuitive Prozesse ineinander greifen sollen (Jungk und Müllert 1989). Üblicherweise nimmt ein solches Verfahren mehrere Tage in Anspruch.

Das Leibniz-Netzwerk Unterrichtsrechnung hat sich an der Methode orientiert, um in einem kreativen und gleichzeitig strukturierten (jedoch wesentlich kürzeren) Austausch zwischen den Netzwerkmitgliedern Zukunftsszenarien für den Unterricht zu entwerfen und aus diesen Szenarien Handlungsfelder für die Unterrichtsqualitätsforschung abzuleiten. Die Zukunftswerkstatt fand im Oktober 2021 als Online-Veranstaltung unter Anleitung einer professionellen Moderation statt. Die Grundidee war es, sich die Unterrichtsrealität im Jahr 2035 vorzustellen und anhand dieser Zukunftsvision die zukünftigen Aufgaben der Unterrichtsrechnung abzuleiten. Der ca. dreistündige Austausch bestand aus zwei Phasen. Die erste Phase diente der Identifikation von Einflussfaktoren, die in den nächsten 15 Jahren aller Voraussicht nach relevant für das Gelingen von Unterricht sind. Durch Kombination der Einflussfaktoren wurden dann Zukunftsszenarien für den Unterricht entworfen. In der zweiten Phase wurden aus den Zukunftsszenarien Handlungsfelder abgeleitet, denen sich die Unterrichtsqualitätsforschung widmen muss, um zum Gelingen von Unterricht in Zukunft beitragen zu können.

Teilnehmende waren 13 Professor*innen und 4 Post-Doktorand*innen aus Deutschland und der Schweiz mit Forschungsschwerpunkten im Bereich der Unterrichtsqualitätsrechnung. Die Teilnehmenden waren Psycholog*innen bzw. Erziehungswissenschaftler*innen mit Professuren bzw. Stellen im Bereich der pädagogisch-psychologischen Lehr-Lernrechnung, der Allgemeinen Didaktik, der Fachdidaktik oder der Grundschuldidaktik.

Es ist offensichtlich, dass die Ergebnisse der Zukunftswerkstatt abhängig vom teilnehmenden Personenkreis sind und daher kein Anspruch auf allgemeine Gültigkeit und Vollständigkeit erhoben werden kann. Auch ist eine umfassende Erarbeitung der Themen im Rahmen eines solch zeitlich reduzierten Austausches nicht möglich. Aus diesem Grund sind die zusammengetragenen Gedanken als Ausgangspunkt für einen fortlaufenden Arbeitsprozess zu sehen, in den Abschn. 4 Einblick gibt.

3.2 Ergebnisse der Zukunftswerkstatt

Wir fassen nachfolgend die zentralen Ergebnisse der Zukunftswerkstatt in knapper Form zusammen. Sie beziehen sich auf die beiden leitenden Fragen:

- Welche Einflussfaktoren werden in den nächsten 15 Jahren aller Voraussicht nach relevant für das Gelingen von Unterricht?
- Was muss die Unterrichtsqualitätsforschung tun, um zum Gelingen von Unterricht in Zukunft beizutragen?

Heterogenität der Lernenden Die Teilnehmenden sahen die Zunahme der Heterogenität auf Seiten der Lernenden als klar zu erwartende Entwicklung an. Sie erwarten zunehmende Unterschiede zwischen den Lernenden im wirtschaftlichen, kulturellen und motivationalen Bereich sowie bezogen auf das Vorwissen. Eine notwendige Folge dieser Entwicklung wird ein sich verstärkender Trend hin zur stärkeren Differenzierung des Unterrichts sein. Um in Zukunft zum Gelingen von Unterricht beitragen zu können, muss die Unterrichtsqualitätsforschung daher aus Sicht der Teilnehmenden ihren Fokus auf die Erforschung adaptiver Unterrichtsformen und Lernunterstützung richten. Hier gilt es zum einen Qualitätsmerkmale von Adaptivität zu identifizieren und messbar zu machen, wodurch sich auch die Notwendigkeit der methodologischen Weiterentwicklung des Feldes ergibt. Zum anderen stellt sich die Frage, ob sich die Unterrichtsqualitätsforschung stärker als bislang an der Entwicklung von Unterrichtskonzepten und Methoden beteiligen sollte, um der Praxis empirisch erprobte Best-Practice Ansätze für die Differenzierung im Unterricht an die Hand geben zu können. Zudem werden andere Prüfungs- und Diagnostikformate benötigt, die dazu geeignet sind, individuelle Lernprozesse zu begleiten und zu unterstützen.

Digitalisierung Die Digitalisierung wird aus Sicht der Teilnehmenden ebenfalls mit Sicherheit und in zunehmendem Maße den schulischen Unterricht prägen. Sie wird sowohl als Herausforderung als auch als Chance gesehen. Eine Herausforderung stellt dabei sowohl die Entwicklung geeigneter digitaler Tools für den Unterricht dar als auch ihre sinnvolle Implementation, die auf Seiten der Lehrkräfte professionelle Kompetenzen voraussetzt. Eine Chance sehen die Teilnehmenden in den neuen Möglichkeiten, die digitale Medien (potenziell) im Bereich des adaptiven Unterrichts und der formativen Diagnostik bieten. Eine weitere Aufgabe für die Unterrichtsqualitätsforschung wird daher in der Begleitung der Digitalisierung des Unterrichts gesehen. Zum einen muss sie ihr Wissen in die Entwicklung digitaler Lehr-Lernsysteme einbringen, um deren Qualität und Lernwirksamkeit sicherzustellen. Zum anderen muss sie weiter Forschung zu der Frage betreiben, wie digitale Medien im Unterricht sinnvoll implementiert werden können und wo die Grenzen ihrer Nutzbarkeit liegen.

Kontext und Strukturen In Folge der Notwendigkeit, zunehmend individualisierter und mit Unterstützung digitaler Tools zu unterrichten, könnte schulisches Lernen zukünftig in anderen Kontexten und weniger ortsgebunden stattfinden als bislang. Strukturen wie Klassen, Jahrgangsstufen oder der Fachunterricht könnten anderen

Formaten weichen, wodurch ganz andere Ansätze der Unterrichtsgestaltung mit evtl. anderen Qualitätsmerkmalen entstehen werden. Im Angesicht dieser Veränderungen muss aus Sicht der Teilnehmenden mehr Wissen über die Bedeutung des Kontextes (sowohl räumlich als auch sozial) für das schulische Lernen generiert werden. Dies sei wichtig, um in einer Situation des Umbruchs Empfehlungen bezüglich geeigneter Strategien bei der Gestaltung von Lernumwelten geben zu können.

Professionalisierung Die Art und Weise, wie sich Heterogenität und Digitalisierung im Unterricht auswirken und welche Risiken und Potenziale mit diesen und anderen Veränderungen verbunden sind, hängt aus Sicht der Teilnehmenden davon ab, inwieweit die Professionalisierung von Lehrkräften gelingt. Ein nach einheitlichen Prinzipien und auf Grundlage gesicherter wissenschaftlicher Erkenntnisse ausgebildetes Lehrpersonal sei unverzichtbar, um die Möglichkeiten, die etwa die Digitalisierung potenziell bietet (s. oben) nutzen zu können. Die Professionalisierung der Lehrkräfte fällt zwar nicht direkt in den Forschungskontext der Unterrichtsqualitätsforschung. Es wird jedoch deutlich, dass eine enge Überschneidung der beiden Themengebiete besteht.

Bildungsgerechtigkeit Die Teilnehmenden diskutierten an verschiedenen Punkten der Zukunftswerkstatt, wie gesellschaftliche und politische Rahmenbedingungen auf Schule und Unterricht Einfluss nehmen. So hängt etwa die Frage nach Wohlstand und seiner Verteilung mit der Frage zusammen, inwieweit eine Volkswirtschaft ein inklusives Bildungsangebot und Förderung für leistungsschwache Lernende zur Verfügung stellt. Die Unterrichtsforschung kann zu dieser Frage nicht unmittelbar beitragen. Es wurde jedoch deutlich, dass wie oben aufgeführt auch die Unterrichtsforschung sich einer Diskussion über Werte und Normen nicht entziehen kann.

4 Wie können die Herausforderungen zukünftig bearbeitet werden?

In Abschn. 2 wurden auf Grundlage der Literatur zentrale Errungenschaften und Herausforderungen der Unterrichtsqualitätsforschung skizziert. Die Ergebnisse des in Abschn. 3 dargestellte Expert*innendiskurs bestätigen und ergänzen die Liste der Herausforderungen. Zudem wurde in einem freien Gedankenaustausch die Zukunft des schulischen Unterrichts in den Blick genommen und Möglichkeiten für die Unterrichtsqualitätsforschung diskutiert, auf diese Veränderungen zu reagieren. In einem dritten Schritt bringen zurzeit fünf Arbeitsgruppen beide Ansätze zusammen, in dem sie ausloten, wie die Unterrichtsqualitätsforschung ausgehend von ihren bisherigen Erkenntnissen auch weiterhin und verstärkt wissenschaftliche Impulse für die Umsetzung erfolgreichen Unterrichts liefern kann. Folgende Fragestellungen wurden hierbei spezifiziert:

1. Welche neuen Aufgaben und Ziele ergeben sich für die die Unterrichtsqualitätsforschung im Lichte sich verändernder gesellschaftlicher Rahmenbedingungen?
2. Wie verändert sich Unterricht im Kontext sich auflösender Strukturen (z. B. bezogen auf Klasse, Fach, Jahrgangsstufe)?

3. Welche (neuen) Formen der Diagnostik tragen zu einer Verbesserung des Unterrichts bei?
4. Wie können innovative und belastbare Forschungsdesigns der Unterrichtsqualitätsforschung aussehen?
5. Sollte die Unterrichtsqualitätsforschung fachspezifischer werden?

Da eine zunehmende Heterogenität der Lernenden im Unterricht sowie eine zunehmende Digitalisierung des Unterrichts von allen Teilnehmenden als zu erwartende Entwicklungen angesehen wurden, wurden diese Aspekte als Querschnittsthemen in den Überlegungen aller Arbeitsgruppen berücksichtigt. Die Arbeitsgruppen hatten die Möglichkeit, sich ihrer Fragestellung theoretisch-konzeptionell zu nähern oder an Ideen für konkrete Forschungsprojekte zu arbeiten. Die Arbeit an diesen Themen dauert derzeit noch an. Es folgen Werkstattberichte aus den Gruppen, die Einblick in die laufende Arbeit geben, aber noch nicht als abschließende Ergebnisse zu verstehen sind. Dementsprechend spiegeln die einzelnen Berichte den jeweils spezifischen Diskussionsstand der Arbeitsgruppen wider. Die Arbeitsgruppen wurden gebeten, den Beitrag zu ihrer spezifischen Frage nach folgendem Schema zu strukturieren: 1. Was hat die Unterrichtsqualitätsforschung bereits anzubieten? 2. Wo liegen die Herausforderungen? 3. Erste Ideen, wie die Forschung diesen Herausforderungen begegnen könnte.

4.1 Welche neuen Aufgaben und Ziele ergeben sich für die Unterrichtsqualitätsforschung im Lichte sich verändernder gesellschaftlicher Rahmenbedingungen?²

Sich verändernde gesellschaftliche Rahmenbedingungen führen nicht nur zu der Frage, ob die Ziele und Aufgaben der Unterrichtsqualitätsforschung adjustiert werden sollten, sondern sie stellen auch die bisherige Vorgehensweise der Unterrichtsqualitätsforschung auf den Prüfstand. Weiter oben wurde bereits auf die bisher eher unbefriedigende Verzahnung von Forschung und Praxis aufmerksam gemacht, die unter anderem auf ein unidirektionales Transferverständnis aus der Forschung hinein in die Schulpraxis zurückzuführen ist. Dabei finden die unterschiedlichen Denk- und Handlungslogiken von Forschung und Praxis nicht ausreichend Berücksichtigung (Hartmann et al. 2016; Hartmann und Kunter 2022). Bisher konnte die Forschung nicht hinreichend auf zeitkritische, drängende Herausforderungen in der Unterrichtspraxis (z. B. obligatorischer Fernunterricht, Integration von geflüchteten Kindern, Umgang mit einer zunehmend heterogenen Schülerschaft, Umsetzung eines inklusiven Unterrichts) reagieren und überzeugende Antworten liefern, die aus der Praxis heraus wiederum aufgegriffen und umgesetzt worden wären (Brühwiler und Leutwyler 2020). Dies ist unter anderem der Tatsache geschuldet, dass wissenschaftlicher Erkenntnisgewinn als langfristiger Prozess angelegt werden muss. Gleichwohl sind die bisherigen Strategien der Erkenntnisgewinnung auch grundsätzlich zu hinterfragen: Wie kann die Unterrichtsqualitätsforschung vorgehen, um

² Mitglieder der Arbeitsgruppe: Lukas Begrich, Doris Holzberger, Thilo Kleickmann, Anna-Katharina Praetorius, Sandy Taut, Svenja Vieluf.

zukunftsgerichtet und proaktiv Wissen zu generieren, das als kumulative Grundlage bedeutsamer, zeitgemäßer Veränderungen in der Schulpraxis nützlich ist?

Obwohl aus Sicht der Arbeitsgruppe alle Schattierungen auf dem Kontinuum zwischen Grundlagenforschung und Anwendungsforschung notwendig sind, sollte die Unterrichtsqualitätsforschung insgesamt noch mehr tun, um sich stärker mit der Praxis zu vernetzen, wie dies in sog. *research practice partnerships* in anderen Ländern schon länger Tradition hat (Coburn und Penuel 2016; Donovan et al. 2021). Dafür muss die Perspektive von Praktiker*innen schon bei der Entwicklung gemeinsamer Forschungsfragen – und dann während des gesamten Forschungsprozesses – berücksichtigt werden. Jones et al. (2018) sprechen bei einer gewinnbringenden Zusammenarbeit von Forschung und Praxis davon, zunächst Veränderungstheorien (*theories of change*) zu entwickeln, die beschreiben, was passieren muss, damit ein gemeinsam definiertes Problem gelöst werden kann. Diese subjektiven Theorien machen die zugrundeliegenden Annahmen aller Beteiligtegruppen explizit und dienen als Ausgangspunkt für die Generierung und anschließende Überprüfung von Lösungsvorschlägen in Form von Interventionen. In Feldtests (*inquiry cycles*) können diese Interventionen ausprobiert und die Veränderungstheorie kann – auch unter Berücksichtigung von Variationen in der Implementierung – aufgrund der Ergebnisse dieser Feldtests präzisiert werden. Zudem können Best Practice Ansätze aus der Praxis in standardisierten Settings systematisch erprobt und optimiert werden. Ein weiterer positiver Nebeneffekt einer solchen Vernetzung von Forschung und Praxis wäre ein längerfristig erleichterter Zugang zum Feld. Nach Meinung der Gruppe sollten sich nicht nur Forschung und Praxis stärker vernetzen, auch mehrere Forschungsstandorte sollten zusammenarbeiten, um die Bearbeitung von Fragestellungen abzustimmen und Daten entsprechend ressourcenschonend zu erheben, damit sie mehreren Zwecken dienen können.

Die Ansprüche an die Kompetenzen von Kindern und Jugendlichen, bspw. im Bereich des Lernens mit und über digitale Medien, interkultureller Kompetenzen oder von kollaborativem Lernen, verändern sich aktuell rasant. Diese Veränderungen müssen sich auch in den untersuchten Outcome-Variablen der Unterrichtsqualitätsforschung adäquat widerspiegeln, wenn sie weiterhin relevantes Wissen für die Unterrichtspraxis generieren möchte. Es wäre zu untersuchen, inwiefern die Outcome-Variablen bisheriger Unterrichtsforschungsprojekte die aktuell in Deutschland geltenden fachspezifischen Bildungsstandards sowie insbesondere auch überfachliche Bildungsziele und -kompetenzen bereits adressieren, und wie letztere in zukünftigen Forschungsprojekten noch stärker einbezogen werden können. Hier stellt auch die valide Messung dieser immer relevanteren sozialen, personalen und lernmethodischen Kompetenzen eine Herausforderung dar. Als weiteres Entwicklungsfeld wurde konstatiert, dass Forschungsbefunde meist in klassischen Unterrichtsettings gewonnen werden, während in der Praxis eine Veränderung hin zu anderen (teilweise stark adaptiven, individualisierten und digitalisierten) Unterrichtsformen stattfindet (Abschn. 4.2).

Die Arbeitsgruppe diskutierte schließlich zwei konkrete Handlungsschritte, um die hier skizzierten Probleme anzugehen. Erstens scheint es gewinnbringend, ein systematisches Review zur theoretischen Fassung und Operationalisierung von in der Unterrichtsqualitätsforschung berücksichtigten Outcome-Variablen und deren Bezug

zu den Bildungsstandards sowie überfachlichen Kompetenzen durchzuführen, um einen systematischen Überblick über zukünftig stärker zu adressierende Outcome-Variablen zu erhalten. Zweitens wurde die Durchführung eines standortübergreifenden Pilotprojekts als *research practice partnership* in den Blick genommen, das einen nicht oder kaum beachteten Zielbereich bearbeiten soll. Hierfür müssten zunächst ausgewiesene Praxisexpert*innen identifiziert und in den Gestaltungs- und Umsetzungsprozess des Projektes einbezogen werden. Die konkreten Schritte würden die gemeinsame Definition eines Problems, die Erstellung einer Veränderungstheorie und die Entwicklung praxistauglicher Interventionen umfassen, die dann in Feldstudien auf ihre Wirksamkeit zu prüfen wären. Die Komplexität der Problemstellung und der Interventionen sollte überschaubar, aber trotzdem aussagekräftig sein. Methodisch wäre je nach Forschungsstand der Einbezug qualitativer und quantitativer Ansätze wünschenswert, wie beispielsweise quasiexperimentelle Designs oder *collaborative action research* (Somekh 2010; Johnson 2011; Altrichter et al. 2018).

4.2 Wie verändert sich Unterricht im Kontext sich auflösender Strukturen (z. B. bezogen auf Klasse, Fach, Jahrgangsstufe)?³

Wie in Abschn. 2.2 erwähnt, liegt eine der Herausforderungen der Unterrichtsqualitätsforschung darin, dass Unterricht immer vom jeweiligen Kontext abhängt. Entsprechend beeinflussen auch gesellschaftliche Veränderungsprozesse, wie eine zunehmend heterogene Schüler*innenschaft, den Unterricht. Die damit einhergehende Forderung einer individuellen Förderung (Klieme und Warwas 2011) ist inzwischen in allen Schulgesetzen der Bundesländer in Deutschland verankert. Auch in einigen Schulen wird ein Stück weit von der tradierten *Grammar of schooling* (Tycak und Tobin 1994) Abstand genommen und den Akteur*innen werden zunehmend mehr Freiheitsgrade und Flexibilität eingeräumt. In Schulen mit einem hohen Grad an Individualisierung zeigt sich dieses in den strukturell-organisatorischen Unterrichtsabläufen (z. B. durch Freiarbeitszeiten und weniger starke Taktungen statt fixer Stundenpläne), in den Sozialformen (z. B. durch flexible Lerngruppen und freie Wahl der Lernpartner*innen statt einem festen Klassensetting) und bei den Inhalten (z. B. durch flexible Ausgestaltung des Curriculums und Verknüpfung verschiedener Fachinhalte statt strikt nach Fächern getrenntem Unterricht). Damit einher gehen ein verändertes Rollenverständnis sowohl der Lehrkräfte (im Sinne von Supervisor*innen und Moderator*innen des Lernens; Pauli und Reusser 2000) als auch der Schüler*innen, bei denen verstärkt auf eigenverantwortliches („choice and voice of learning“, Mötteli et al. 2022) und selbstreguliertes Lernen gesetzt wird. Ebenso werden im Zuge der Schulentwicklungsprozesse nicht nur fachliche Leistungen der Schüler*innen in den Blick genommen, sondern Bildungs- und Entwicklungsziele ganzheitlicher betrachtet (siehe auch Abschn. 4.1). Erste Einblicke in innovative Schulen auf Basis empirischer Studien, wie Ada*Q (Adaptivität und Unterrichtsqualität im individualisierten Unterricht; Decristan et al. 2017) und Per-

³ Mitglieder der Arbeitsgruppe: Jasmin Decristan, Doris Holzberger, Eckhard Klieme, Susanne Kuger, Christine Pauli, Katrin Rakoczy.

Len (Personalisiertes Lernen in heterogenen Lerngruppen; Stebler et al. 2018) legen nahe, dass dieser aus den Schulen heraus stattgefundene bottom-up-Prozess in einem Blumenstrauß an mehr oder weniger kreativen Konzepten mündet.

Zur Einordnung dieser Konzepte aus Perspektive der Unterrichtsqualitätsforschung kann auf die in Abschn. 2.1 und 2.2 angesprochenen Errungenschaften und Herausforderungen zurückgegriffen werden. Entsprechend der Unterscheidung zwischen Oberflächen- und Tiefenmerkmalen sind Maßnahmen der Differenzierung und Individualisierung an der Oberfläche verortet und nicht per se lernförderlich (z. B. Gehrler und Nusser 2020). Vielmehr gilt es, vor allem die Tiefenebene zu betrachten und zu untersuchen, wie qualitativ hochwertig der Unterricht ist und wie Oberflächen- und Tiefenmerkmale zusammenspielen (Decristan et al. 2020). Dabei legen Befunde der Unterrichtsqualitätsforschung auch nahe, dass Unterrichtsqualitätsmerkmale sowohl allgemein als auch differenziell wirksam sein können: Besonders leistungsschwächere Schüler*innen profitieren von einer hohen Unterrichtsqualität, im Sinne der Tiefenmerkmale (Decristan et al. 2016, 2022).

Zum Unterricht an Schulen mit hohem Grad an Individualisierung liegt jedoch bislang wenig Empirie vor. Entsprechend gilt es nun dafür zu sorgen, dass die Konzepte aus der Praxis Eingang in die Unterrichtsqualitätsforschung finden (also umgekehrt zu der in Abschn. 2.1 formulierten Errungenschaft; weiterführende Diskurse zum Verhältnis von Forschung und Praxis finden sich in Abschn. 4.1). Vor diesem Hintergrund wird es eine zentrale Aufgabe und Herausforderung der zukünftigen Unterrichtsqualitätsforschung sein, diese veränderten Lehr-Lernsettings empirisch zu erforschen und dabei folgende zentrale Fragen zu beantworten: Was charakterisiert Unterricht im Kontext sich auflösender Strukturen? Welche empirischen Wirkungen lassen sich hierzu aufzeigen – sowohl mit Blick auf multiple Ziele von Schule und Unterricht als auch hinsichtlich der verschiedenen Lernbedarfe der Schüler*innen? Und wie greifen die in den Schulen implementierten verschiedenen Maßnahmen ineinander?

Um diese zentralen Fragen zu adressieren, gilt es zunächst zu explizieren, wie diese „neue Adaptivität“ mit ihren ganz unterschiedlichen Begrifflichkeiten (z. B. Personalisiertes Lernen, Neue Lernkultur) an bisherige Begriffe und Konzepte der Erziehungswissenschaft und Pädagogischen Psychologie anknüpft (Dumont 2019) und mit welchen theoretischen Grundannahmen diese jeweils in Beziehung gesetzt werden können. Ebenso bedarf es vertiefter Überlegungen, was qualitätsvollen Unterricht unter den veränderten Strukturen charakterisiert. Es lässt sich erwarten, dass Unterrichtsqualität weiterhin eine zentrale Rolle für fachliches Lernen und Motivation einnimmt, dass Unterrichtsqualitätsmerkmale jedoch in ihren Konzeptualisierungen und Operationalisierungen stärker auf die individualisierten Angebotsstrukturen angepasst werden müssen.

Diese theoretisch-konzeptionellen Überlegungen leiten zu zentralen methodischen Herausforderungen der zukünftigen Unterrichtsqualitätsforschung über. Es bedarf einer Ergänzung, Anpassung und teils auch innovativen Neu-Entwicklung von Erhebungsinstrumenten, um die Flexibilität, die unterrichtlichen Interaktionen und die Individualisierung des Lernens zu erfassen. Multiple Kamerasysteme (Troll et al. 2022) und mobile Tischkameras sind hierfür erste Beispiele. Fragebögen für Schüler*innen müssen um multiple Ziele ergänzt werden, sodass zusätzlich zu fach-

lichen Leistungen auch sozio-emotionale und motivationale Ziele, aber auch Persönlichkeitsentwicklung, Demokratiebildung und Kompetenzen im Umgang mit gesellschaftlichen und ökologischen Veränderungen (Digitalisierung, Vielfalt, Klima) einbezogen werden (Abschn. 4.1). Hierbei kann auch auf bestehende Instrumente verschiedener Disziplinen zurückgegriffen werden (z. B. zu interkultureller Kompetenz Göbel und Buchwald 2017). Um die momentanen Lernbedarfe und das situative Erleben der Schüler*innen abhängig von dem gegenwärtigen Lernkontext besser zu erfassen, bedarf es entsprechend sensitiver Instrumente mit kleinschrittigeren Erhebungsintervallen. Experience-Sampling Methoden (z. B. Rakoczy et al. 2022) stellen hierfür einen möglichen Zugang dar, weitere Ausführungen zu der für eine Förderung nötigen (lernbegleitenden) Diagnostik finden sich in Abschn. 4.3. Zudem gilt es das soziale Gefüge (z. B. über Netzwerkanalyse wie in Troll et al. 2022; Zander et al. 2017) eingehender in den Blick zu nehmen, auch um bspw. das Spannungsfeld aus Individualisierung und Gemeinschaft näher zu beleuchten. Um differenzielle Wirkungen abhängig von den Lernvoraussetzungen zu erfassen (vgl. die in Abschn. 2.2 angesprochenen Herausforderungen), kann auf die ursprüngliche Idee der Aptitude-Treatment-Interaktionen (Corno und Snow 1986) zurückgegriffen und diese weiterentwickelt werden. Für valide Aussagen sollten quantifizierende Methoden mit einer dichteren Beschreibung der unterrichtlichen Interaktionen durch qualitative Verfahren (z. B. Breidenstein und Rademacher 2017) angereichert werden. Grundsätzlich wird es mit Blick auf Forschungs- und Erhebungsdesigns (nähere Ausführungen in Abschn. 4.4) in der zukünftigen Unterrichtsqualitätsforschung auch darum gehen, genauer zu eruieren, welche Untersuchungseinheit für die Erforschung der veränderten Strukturen angemessen ist – die einzelne Schule, ein Jahrgang, oder wechselnde Kleingruppen, statt wie bisher der Klassenverbund? Gleichmaßen ist zu hinterfragen, ob und welche Vergleichsgruppen für die Erforschung der veränderten Strukturen geeignet wären, vor allem vor dem Hintergrund der Bandbreite der konzeptionellen Ausgestaltungen der „neuen Schulen“.

Zusammen genommen wird es darum gehen, unter Einbezug verschiedener auf den situativen Kontext angepasster sowie innovativer Instrumente und Auswertungsmethoden, die individualisierten Angebotsstrukturen und die unterrichtlichen Interaktionen unter Berücksichtigung sich vielfach verändernder Sozialformen und Kleingruppen besser nachzuzeichnen sowie deren allgemeine und schüler*innenspezifische Wirkungen mit Blick auf das (simultane) Erreichen multipler Ziele von Schule und Unterricht zu erforschen. Schulen haben sich bereits auf den Weg gemacht, um der zunehmenden Heterogenität der Schüler*innenschaft gerecht zu werden. Nun braucht es Mut, Innovationskraft und Expertise, damit die Unterrichtsqualitätsforschung diese Entwicklung besser greifen und empirisch fundierte Aussagen hierzu treffen kann.

4.3 Welche (neuen) Formen der Diagnostik tragen zu einer Verbesserung des Unterrichts bei?⁴

Die in Abschn. 4.1 skizzierten gesellschaftlichen Veränderungen bringen auch neue Ansätze und Praktiken der Unterrichtsgestaltung mit sich. Wie in Abschn. 4.2 ausgeführt, zeigt sich das beispielsweise in einer immer öfter systematisch herbeigeführten Flexibilisierung des Unterrichts auf unterschiedlichen Dimensionen (z. B. sozial, räumlich, zeitlich; Dumont 2019). Neue Lernformen des Hybridunterrichts, zusammen mit neuen Möglichkeiten bei der Nutzung digitaler Medien für die Unterrichtsgestaltung, verstärken die Entwicklungen zusätzlich. Egal ob Flexibilisierungen im Unterricht nun absichtlich herbeigeführt werden, oder eine notwendige Reaktion auf sich verändernde schulische Rahmenbedingungen und eine zunehmende Heterogenität der Schüler*innen sind – sie stellen die Lehrkräfte vor neue Herausforderungen (Abschn. 4.2). Nimmt die Unterrichtsqualitätsforschung ihren Anspruch auf Praxisrelevanz ernst, so wird sie sich dieser Herausforderungen aus einer wissenschaftlichen Perspektive annehmen müssen und evidenzbasierte, praxistaugliche Lösungen anbieten müssen.

In Anbetracht von immer offensichtlicher werdender Heterogenität in den Lernvoraussetzungen von Schüler*innen und einer damit einhergehenden zunehmend wichtigeren Adaptivität von Lehr-Lern-Arrangements (Corno 2008) wird eine gute unterrichtsintegrierte Diagnostik in Zukunft eine entscheidende Rolle spielen (siehe auch das aktuelle SWK-Gutachten zur Grundschule; Köller et al. 2022). Für die unterrichtenden Lehrkräfte dürfte die Herausforderung steigen, die individuellen Lernstände und motivationalen Voraussetzungen ihrer Schülerinnen und Schüler einzuschätzen, um ihnen passgenaue Angebote unterbreiten zu können.

Diese zunehmende Bedeutung der Diagnostik im weiteren Sinne macht es nötig, dass theoretische Ansätze und empirische Analysen, die bislang zumeist in relativ disparaten Forschungsfeldern diskutiert wurden, in Zukunft stärker miteinander in Beziehung zu setzen und sie zum Teil neu zu denken.

Aus unserer Sicht sind es vor allem drei Forschungsfelder, die sich mit der Rolle von Diagnostik im Unterricht befasst haben: Die Forschung zu pädagogisch-psychologischen Diagnostik, die empirische Unterrichtsqualitätsforschung und die allgemein- und fachdidaktische Forschung.

Pädagogisch-psychologische Diagnostik In der Forschung zur pädagogisch-psychologischen Diagnostik haben sich Forschende seit jeher mit der Frage beschäftigt, wie sich psychometrisch orientierte Verfahren im Unterricht und für den Unterricht nutzen lassen (Hesse und Latzko 2017). Vor allem in der Forschung zu *curriculum-based measurements* und *formative assessments* steht explizit die Frage im Mittelpunkt, wie sich Diagnostik, z. T. sehr weit gefasst auch als *on-the-fly assessment*, für die individuelle Förderung einzelner Schüler*innen für die Gestaltung des Unterrichts nutzen lässt (Black und Wiliam 1998). Instrumente der Lernverlaufsdiagnostik wie beispielsweise quop (Salaschek et al. 2014) oder die in der Lernverlaufs-

⁴ Mitglieder der Arbeitsgruppe: Esther Brunner, Benjamin Fauth, Marc Kleinknecht, Anke Lindmeier, Alexander Naumann.

Infrastruktur Levumi (Gebhardt et al. 2016) entstehenden Tests bieten die Möglichkeit, individuelle Lernverläufe ökonomisch zu erfassen, grafisch abzubilden und bei auffälligen Entwicklungen individuell intervenieren zu können (Souvignier et al. 2020). Die dafür notwendigen aufwändigen Testentwicklungsprozesse wurden bisher jedoch nur für sehr ausgewählte Themen und Populationen durchgeführt, so dass die Frage der praktischen Anwendbarkeit etwa in der Sekundarstufe des Regelunterrichts noch offen ist.

Unterrichtsqualitätsforschung Auch in der empirischen Unterrichtsqualitätsforschung wird immer wieder die Bedeutung von Diagnostik für die Unterrichtsgestaltung hervorgehoben (Klieme und Rakoczy 2008; Charalambous und Praetorius 2020). Im Modell der drei Basisdimensionen ist Diagnostik keine eigene Dimension, aber ein zentrales Thema innerhalb der Dimensionen. Überall dort, wo die Konzeptualisierung von Unterrichtsqualität mehr oder weniger explizit beinhaltet, dass das Unterrichtsangebot passend sein muss zu den individuellen Lernvoraussetzungen der Schüler*innen, benötigt die Lehrkraft diagnostische Informationen, um didaktisch fundiert Entscheidungen treffen zu können. Das wird besonders relevant bei der kognitiven Aktivierung und der individuellen (fachlichen) Unterstützung (Sliwka et al. 2022) – zum Beispiel wenn es darum geht, für einzelne und für bestimmte Gruppen von Schüler*innen zu einem bestimmten Zeitpunkt die Aufgaben auszuwählen, die sie im Hinblick auf ihr aktuelles Verständnis angemessen zum Weiterlernen herausfordern. Dasselbe gilt für alle Situationen, in denen Lehrkräfte ihren Schüler*innen Feedback zu ihrem Lernen geben (Käfer et al. 2021). Gutes, lernwirksames Feedback braucht immer eine diagnostische Information zum Vorwissen und möglichen Wissenslücken des individuell Lernenden (Hattie und Timperley 2007).

Allgemein- und fachdidaktische Forschung Schließlich sind Fragen der Nutzung diagnostischer Informationen für die Unterrichtsgestaltung auch immer schon selbstverständlicher Bestandteil allgemeindidaktischer (Gläser-Zikuda et al. 2018) und fachdidaktischer Forschung gewesen (Leuders et al. 2022). Diese Ansätze sind unverzichtbar, weil sie – anders als die ersten beiden Forschungsrichtungen – für Unterrichtssituationen fachübergreifend und fachspezifisch explizieren können, was das relevante deklarative, prozedurale und konzeptuelle Wissen ist (Ehmke et al. 2018) und daran anknüpfend Erkenntnisse erlangen können, wie es sich entwickelt und welche individuellen Alltags- und Fehlkonzepte das Lernen beeinflussen (Enenkiel et al. 2022; Kleickmann et al. 2016; Prediger 2011, Schadl 2020). Die Entwicklung diagnostischer Instrumente ist von diesen spezifischen Wissensbeständen sowie von Erwerbs- und Fördermodellen abhängig. Seitens unterschiedlicher Fachdidaktiken wurde in der Vergangenheit darauf hingewiesen, dass diese Aspekte in der pädagogisch-psychologischen Forschung zum Teil nicht hinreichend berücksichtigt wurden (Leuders et al. 2022). Solche Verbindungen sind ein Desiderat künftiger Forschung.

Zum Teil gibt es Überschneidungen zwischen diesen drei Forschungsfeldern (pädagogisch-psychologische Diagnostik, Unterrichtsqualitätsforschung und allgemein- und fachdidaktische Forschung) und interdisziplinäre Kollaborationen, oftmals stehen die jeweiligen Überlegungen jedoch noch unverbunden nebeneinander, obgleich

zum Teil ähnliche Problemstellungen adressiert werden. Zudem gibt es nach wie vor eine große Lücke zwischen wissenschaftlich erforschten Ansätzen und deren konkreter Nutzung in der Unterrichtspraxis auf breiter Basis (theory-practice gap). Es ist daher eine Aufgabe für die Unterrichtsforschung – aus pädagogisch-psychologischer, allgemein- und fachdidaktischer Perspektive – zu eruieren, wie sich die oben genannten Forschungsstränge zusammenführen ließen, um theoretische Überlegungen und empirische Befunde für die konkrete Praxis der Lehrkräftebildung und der Unterrichtsentwicklung fruchtbar zu machen (Abschn. 4.1 und 4.5).

Es stellt sich die Frage, ob und inwiefern die Unterrichtsqualitätsforschung auch ein *change agent* für die Verbesserung des Unterrichts in der Praxis sein kann und will. Neben den Folgen für das professionelle Selbstverständnis von Bildungsforscher*innen, hätte das auch Implikationen für die Forschungspraxis. Wird die Erforschung von Unterricht auch im Hinblick auf eine Verbesserung der Unterrichtspraxis betrieben, so werden vermutlich Forschungsansätze relevant, die bislang noch nicht im Fokus der deutschsprachigen Unterrichtsqualitätsforschung standen. So scheint es beispielsweise lohnenswert, von den konkreten, alltäglichen Praktiken und Situationen (*core practices*; Grossman 2018) auszugehen, in denen diagnostische Vorgehensweisen relevant werden, etwa wenn Lehrkräfte Unterrichtsgespräche führen oder Schüler*innen beim Erschließen von Fachtexten anleiten (Kleinknecht et al. 2022). Dies knüpft an allgemeindidaktische Konzepte an, die davon ausgehen, dass sich aus lernpsychologischen Überlegungen Grundformen des Lehrens ableiten lassen, die auch fachdidaktisch anschlussfähig sind (Aebli 1983; Reusser 2018). Von einer partiellen Synthese der oben genannten theoretischen Ansätze könnten Vorschläge ausgehen, wie solche konkreten Unterrichtspraktiken weiterentwickelt und verbessert werden könnten.

4.4 Wie können innovative und belastbare Forschungsdesigns der Unterrichtsqualitätsforschung aussehen?⁵

Im Zuge der Forderung nach evidenzorientierter Praxis im Bildungsbereich (Bauer et al. 2015) ist auch die Unterrichtsqualitätsforschung gefordert, empirisch belastbare Nachweise zur Wirksamkeit eines qualitätsvollen Unterrichts zu liefern und Wissen über konkrete Handlungsweisen für Lehrkräfte bereitzustellen (Pant 2014; Wilkes und Stark 2022). Die besonderen methodischen Herausforderungen des Forschungsfeldes wie die uneinheitliche Operationalisierung von Qualitätsdimensionen des Unterrichts, die Perspektivenspezifität oder die Notwendigkeit, große Stichproben zu akquirieren, wurden einführend bereits benannt (Abschn. 2.2). Weitere Herausforderungen wie die Entwicklung geeigneter Instrumente für die Erfassung von Unterrichtsqualität in individualisierten Lernsettings wurden ebenfalls bereits adressiert (Abschn. 4.3). Darüber hinaus besteht eine grundlegende methodische Herausforderung für die Unterrichtsqualitätsforschung darin, die Wirksamkeit des Unterrichts für die Lern- und Leistungsentwicklung von Schüler*innen empirisch zu prüfen. Naturgemäß bedarf es in der empirischen Forschung hierzu eines For-

⁵ Mitglieder der Arbeitsgruppe: Richard Göllner, Aiso Heinze, Miriam Hess, Carmen Köhler, Wolfgang Wagner.

schungsdesigns, das sowohl die Identifikation der spezifischen Einflussfaktoren erlaubt und eine Ergebnisinterpretation im Sinne einer Ursache-Wirkungsbeziehung im besten Fall möglich macht. Um diese Herausforderung in Zukunft noch besser zu adressieren, bedarf es einer Weiterentwicklung von Forschungsdesigns in der Unterrichtsqualitätsforschung. Drei Ansätze erscheinen dabei besonders vielversprechend und sollen im Folgenden im Überblick dargestellt werden:

Erweiterung längsschnittlicher Untersuchungsdesigns Die empirische Forschung zu Unterrichtseffekten verfolgt nach wie vor häufig einen querschnittlichen Untersuchungsansatz oder betrachtet im Rahmen von Prä-Posttest-Designs lediglich einen kurzen Untersuchungszeitraum. Während die erstgenannten Ansätze nur wenig Aussagekraft im Hinblick auf die Schätzung von Unterrichtseffekten besitzen, zeigen sich für die gängigen Prä-Posttest-Designs häufig eine hohe Zeitstabilität der untersuchten Zielkriterien. Inspiriert durch Längsschnittdesigns aus der entwicklungs- und persönlichkeitspsychologischen Forschung stellt die stärkere Berücksichtigung von sogenannten Transitionsphasen, wie etwa Klassen- oder Lehrkraftwechsel, einen ersten wichtigen Ansatzpunkt dar. Beispielsweise könnte die Untersuchung eines Unterrichtseffektes (z. B. Klassenführung) auf die Mathematikkompetenzentwicklung von Schüler*innen im Laufe eines Schuljahres interessant sein. Insofern das resultierende Prä-Posttest-Design um einen vorgeschalteten Messzeitpunkt im vorausgegangenen Jahr ergänzt würde, wäre eine Trennung lehrkraftspezifischer und lehrkraftunspezifischer Effekte und damit eine sensitivere Testung der angenommenen Unterrichtseffekte möglich. Die analytischen Strategien zur Auswertung solcher Datenstrukturen sind inzwischen gut etabliert. Im Kern trennen die entsprechenden Verfahren situationsspezifische („occasion“) und situationsunspezifische („trait“) Messanteile und erlauben die Modellierung von Carry-Over-Effekten (für einen Überblick siehe Eid et al. 2017). Die Unterscheidung zwischen trait- und occasion-spezifischen Effekten spielt auch eine zentrale Rolle bei der vieldiskutierten Frage der Bevorzugung des Change-Score- versus des ANCOVA-Ansatzes zur Schätzung kausaler (Unterrichts-)Effekte, die abhängig von den jeweiligen Annahmen von trait- und occasion-spezifischen Effekten auf das Outcome und den Prädiktor (z. B. Unterrichtsqualität) getroffen werden sollte (Köhler et al. 2021). Auch zur Untersuchung von Veränderungstheorien im Kontext von inquiry cycles der Forschungs-Praxis-Kooperationen (Abschn. 4.1) wären solche sich über viele Messzeitpunkte erstreckende Längsschnitterhebungen sinnvoll – insbesondere bei relativ kleinen Stichprobengrößen. Damit würden sich Interventionseffekte deutlicher von „Störeinflüssen“ wie Lernvoraussetzungen der Schüler*innen trennen lassen.

Einsatz von Unterrichtsskripts in der experimentellen Forschung Gemeinhin werden experimentelle Studien als Königsweg zur Prüfung von Ursache-Wirkbeziehungen erachtet. Entsprechend existieren auch im Bereich der Unterrichtsqualitätsforschung inzwischen Studien, die experimentelle Variationen des Unterrichts, wie etwa unterschiedliche Formen eines Schülerexperimentes im naturwissenschaftlichen Unterricht (Wirth et al. 2008), umsetzen und bezüglich etwaiger Lerneffekte empirisch prüfen. Wenngleich derartige Studien äußert wertvolle Erkenntnisse zur

Gestaltung eines lernwirksamen Unterrichts liefern, betreffen die experimentellen Variationen häufig nur einen vergleichsweise kleinen Ausschnitt des Unterrichts und werden damit dem Unterricht einer Unterrichtsstunde oder einer Unterrichtseinheit nur bedingt gerecht. Vor diesem Hintergrund ist der Einsatz von Unterrichtsskripts ein vielversprechender Ansatz. Allgemein haben Skripts ihren Ursprung in der Kognitionsforschung (Schank und Abelson 1977). Sie werden als Muster bzw. Regeln definiert, die eine bestimmte Handlung beschreiben und die das Verhalten von Menschen lenken. Skripts werden gleichzeitig als die zentralen Gedächtnisbausteine zur Repräsentation von sich wiederholenden Handlungsabläufen erachtet. Unterrichtsskripts umfassen demnach Handlungsmuster, an denen sich Lehrkräfte in ihrem Unterrichtshandeln orientieren und in denen sie sich erheblich unterscheiden (Blömeke et al. 2003). Ein Überblickartikel von Wolff et al. (2021) zeigt, dass Lehrkräfte unterschiedliche Skripts bezüglich der Wahrnehmung und Interpretation von Unterrichtsstörungen besitzen und entsprechend unterschiedliche Handlungsalternativen zur Vermeidung bzw. Beendigung der Störung im konkreten Unterricht ableiten. Vor diesem Hintergrund sind Skripts eine naheliegende Möglichkeit um Unterricht experimentell zu variieren und etwaige Auswirkungen auf das Lernen von Schülerinnen und Schülern zu prüfen (z.B. Cohen et al. 2003; Kollar et al. 2006). Zu den Vorteilen gehört, dass Skriptvariationen beispielsweise in Form von videografierten Anschauungsbeispielen einfach zu kommunizieren sind, sich auf Handlungsmuster beziehen, die an der mentalen Repräsentation alltäglicher Handlungen ansetzen und für Lehrkräfte zugleich eine hohe Praxistauglichkeit aufweisen. Zudem lassen sich mit Skripts Variationen der Unterrichtsorchestrierung und damit des Zusammenspiels von fachdidaktischen und generischen Perspektiven auf Unterrichtsqualität erreichen. Insofern bieten Skripts zugleich eine integrative Perspektive auf die Erreichung eines lernwirksamen Unterrichts. Schließlich ist davon auszugehen, dass mit Hilfe von Skripts eine praktikable Möglichkeit geschaffen werden kann, um Unterrichtskonzepte zu realisieren und empirisch zu prüfen, die nicht länger herkömmlichen Strukturen (siehe Abschn. 4.2) folgen.

Digitalisierte Unterrichtsforschung zur Prozessbeschreibung Schließlich kann sich die empirische Unterrichtsqualitätsforschung nicht ausschließlich auf die Ermittlung von Lern- und Leistungseffekten beschränken, sondern sollte die relevanten Lernprozesse als solche enger in den Blick nehmen. In diesem Zusammenhang ermöglichen digital-basierte Methoden neue Möglichkeiten zur Beschreibung lernwirksamer Unterrichtsprozesse. Zu nennen sind vor allem Verfahren zur Erfassung von Blickbewegungen im Unterricht, die bereits zur Erfassung von lehrkraftseitigen Kompetenzen seit einigen Jahren eingesetzt werden (Cortina et al. 2015). Die Kopplung solcher Daten mit den Blickbewegungsdaten von Lernenden im Unterricht ist eine konsequente Weiterentwicklung dieser Forschungsansätze, um etwa Fragen der Aufmerksamkeitssteuerung im Unterricht prozessnah adressieren zu können (Goldberg et al. 2021). Die Auswertung der daraus resultierenden Daten erfordert mit hoher Wahrscheinlichkeit die Anwendung von informationstechnologischen Verfahren, die weit über die gängigen Datenauswertungsstrategien in der Unterrichtsqualitätsforschung hinausgehen. An erster Stelle sind Verfahren des maschinellen Lernens zu nennen (Hilbert et al. 2021), um etwa im genannten Beispiel dyadische

Blickbewegungsmuster zwischen der Lehrkraft und den Lernenden zu identifizieren. Im Zusammenhang mit einer stärkeren Digitalisierung der Unterrichtsqualitätsforschung ist zudem mit der Weiterentwicklung von Forschungsmethoden zu rechnen, die in vielen Fällen auch eine Wiederverwendung von bereits existierendem Datenmaterial möglich macht. Insbesondere die Nutzung von Unterrichtsvideographien bietet inzwischen einen überaus reichhaltigen Fundus, um Unterrichtsprozesse zu beschreiben. Hierbei wird die Forschung in einigen Jahren nicht mehr ausschließlich auf zeitaufwändige Kodierungen angewiesen sein, sondern mittels KI-gestützter Automatisierungstechniken auf leistungsstarke Ressourcen ergänzend zurückgreifen können.

4.5 Sollte die Unterrichtsqualitätsforschung fachspezifischer werden?⁶

Die Inszenierung und Durchführung guten Fachunterrichts steht im Zentrum der Unterrichtsqualitätsforschung und bildet ein zentrales Ziel der Aus- und Weiterbildung von Lehrkräften (Baumert und Kunter 2006). Die in der Unterrichtsqualitätsforschung entwickelten Konzeptualisierungen beschreiben Unterrichtsqualität dabei oft fachübergreifend (im Überblick Charalambous und Praetorius 2018). Während der Fokus auf eine fachübergreifende und damit einheitliche Betrachtung von Unterrichtsqualität zur Etablierung des Feldes sinnvoll erschien, muss man sich mittlerweile kritisch die Frage stellen, ob diese fachübergreifenden Modelle auch für eine zukünftige Erforschung von Unterrichtsqualität hinreichend geeignet sind, da sich die Lernziele, -gegenstände und -prozesse zwischen den Fächern teils stark unterscheiden. In der Allgemeinen Fachdidaktik (Bayrhuber et al. 2017) werden diese Fragen zum fachlichen Lernen zwar bereits diskutiert, jedoch beziehen sich die didaktischen Argumentationen vornehmlich auf die Strukturierung der Inhalte und die Gestaltung von Lernangeboten. Die Spezifika der unterschiedlichen fachlichen Gegenstände bleiben meist unberücksichtigt.

Dieses Forschungsdesiderat wurde auch in aktuellen Übersichtsarbeiten der Unterrichtsqualitätsforschung deutlich. Es wird attestiert, dass die verschiedenen Subdimensionen der Unterrichtsqualität insbesondere auf Ebene der beobachtbaren Indikatoren fachspezifisch ausdifferenziert werden müssen, um die Unterrichtsprozesse in den verschiedenen Fächern mit ihren unterschiedlichen Lernzielen und Lerngegenständen adäquat beschreiben und operationalisieren zu können (Praetorius et al. 2020b). Damit rückt eine fachdidaktische Unterrichtsqualitätsforschung in den Fokus, die ausgehend von den (1) *fachlichen Zugängen* (z. B. modellieren, bewegen, experimentieren, argumentieren) (2) *Lernziele und Lerngegenstände*⁷ ableitet (z. B. Verstehen von quadratischen Gleichungen, Bewegungen an Geräten lernen) und zugehörige (3) *Lehr- und Lernprozesse* beschreibt (z. B. mathematisches Modellieren, motorisches Lernen). Aus diesen spezifischen Lehr- und Lernprozessen können notwendige Anpassungen bei den Subdimensionen der bestehenden Unterrichts-

⁶ Mitglieder der Arbeitsgruppe: Christiane Bertram, Christian Herrmann, Stefan Keller, Anke Lindmeier, Anna-Katharina Praetorius, Patrick Schreyer, Mirjam Steffensky, Sandy Taut, Iris Winkler.

⁷ In diesem Zusammenhang verstehen wir unter dem fachlichen Lerngegenstand die das Lernziel konstituierenden Inhalte und Methoden/Arbeitsweisen.

qualitätsmodellen gefolgt und in konkrete beobachtbare Indikatoren guten Fachunterrichts operationalisiert werden. Was beispielsweise eine gelingende kognitive Aktivierung ausmacht, wird in und zwischen den Fächern sehr kontrovers diskutiert. Zwar besitzt die kognitive Aktivierung in jedem Fach eine hohe Bedeutung, jedoch werden in Abhängigkeit der didaktischen und lerntheoretischen Fundierung der Fächer unterschiedliche Kognitionen als bedeutsam erachtet, um fachbezogene Lernprozesse bei Schüler*innen zu unterstützen. Entsprechend wird kognitive Aktivierung in den Fachdidaktiken aufgrund der unterschiedlichen Lerngegenstände unterschiedlich definiert und erweitert. So wird im Fach Sport (auch) eine kognitiv-motorische bzw. ästhetische Aktivierung (Herrmann, Gogoll und Gerlach 2020; Laging 2022) und im Literaturunterricht eine kognitiv-emotionale Aktivierung (Hesse und Winkler 2022) diskutiert.

Aus diesen Gründen wird die Frage nach der spezifischen Fachlichkeit der verschiedenen Schulfächer vermehrt in den Fokus der Unterrichtsqualitätsforschung gerückt: Was ist der fachliche Gegenstand auf der Grundlage von welchen Fachvorstellungen? Was macht diesen Gegenstand fachspezifisch und welchen spezifischen fachlichen Beitrag leistet das Fach zum besseren Verstehen der Welt? Wie kann dieser Fachgegenstand qualitativ gelehrt und gelernt werden und welche Inhalte und Aufgabenformate sind hierfür zentral?

Zur Beantwortung dieser Fragen kann sicherlich die Allgemeine Fachdidaktik einen Beitrag leisten, indem sie Merkmale und Unterschiede der Fächer herausarbeitet und sich um eine *fachübergreifende* Theorie des fachlichen Lernens bemüht (Rothgangel et al. 2020). *Fachspezifische* Theorien des fachlichen Lernens (z. B. motorisches Lernen) werden dabei jedoch weitestgehend ausgeklammert. Um das Eigenständige der jeweiligen Fächer zu erhalten, müssen die fachspezifischen Diskurse um (kognitive) Aktivierung adäquat berücksichtigt werden (Laging 2022) und von den jeweiligen Fachdidaktiken als Subdimensionen und Indikatoren der Unterrichtsqualität verortet und begründet werden.

Diese Heterogenität von fachspezifischen und fachübergreifenden Modellen der Unterrichtsqualität stellt eine Hürde für den Diskurs über die Schulfächer hinweg dar, erschwert die Aus- und Weiterbildung der Lehrkräfte und verhindert für die Unterrichtspraxis klare Orientierungslinien.

Zielführend wäre es daher, ein Rahmenmodell guten *Fachunterrichts* zu erarbeiten, bei dem die fachübergreifenden Dimensionen der Unterrichtsqualität ein Gerüst bilden, um lerngegenstandsspezifische Anpassungen für die jeweiligen Fächer vorzunehmen. Als Basis können weitestgehend fachübergreifende Dimensionen der Unterrichtsqualität angenommen werden, wie sie beispielsweise im Modell der drei Basisdimensionen bestimmt wurden (Praetorius et al. 2018). So sind gewisse Dimensionen des Unterrichts (z. B. Klassenführung) in allen Fächern wichtig, haben jedoch in den Fächern unterschiedliche Relevanz und differenzieren sich auf Ebene der Subdimensionen unterschiedlich aus (z. B. Sicherheit im Fach Sport).

Insbesondere die Dimension der kognitiven Aktivierung bedarf einer fachdidaktischen Ausdifferenzierung, da unterschiedliche Lerngegenstände auch unterschiedliche Arten der Aktivierung implizieren oder unterschiedlich gewichten. Beispielsweise wird für den Literaturunterricht eine rein kognitiv-analytische Auseinandersetzung mit literarischen Texten nicht als ausreichend erachtet, sondern die Relevanz

des Wechselspiels von subjektiv-emotionaler Involviertheit und genauer Textwahrnehmung betont, weshalb Hesse und Winkler (2022) aus literaturdidaktischer Sicht eine Differenzierung von “kognitiv-emotionaler Aktivierung” vorschlagen.

Dabei stellt sich die Frage, wie die kognitive Aktivierung angelehnt an den fachdidaktischen Diskursen in den Schulfächern durch weitere lerngegenstandsorientierte Aktivierungsformen ergänzt werden kann (z. B. motorische, ästhetische und emotionale Aktivierung, Herrmann et al. 2020; Laging 2022; Hesse und Winkler 2022).

Das Rahmenmodell guten *Fachunterrichts* müsste also über eine lerngegenstandsspezifische Ausdifferenzierung verfügen und jene (Sub-)Dimensionen der Unterrichtsqualität Raum geben, die sich auf den Aufbau, die Integration, Umstrukturierung, Erweiterung und Differenzierung fach- und konzeptspezifischer Lernprozesse richten (Reusser und Pauli 2021).

In der Verknüpfung der bislang vorrangig fachübergreifend ausgerichteten Unterrichtsqualitätsforschung und Allgemeinen (Fach-)Didaktik mit den fachdidaktischen Diskursen in den Schulfächern liegt ein zentrales zukünftiges Handlungsfeld der Unterrichtsqualitätsforschung auf theoretisch-konzeptioneller wie auch auf empirischer Ebene. Der Erfolg wird sich jedoch an der Implementation in die Unterrichtspraxis messen lassen müssen. Entsprechend eröffnet sich daraus ein weiteres Handlungsfeld, welches in der Nutzung der konzeptionellen und empirischen Erkenntnisse in der Aus- und Weiterbildung der Lehrkräfte liegt. So könnte das angezielte Rahmenmodell in den fachübergreifenden Lehrveranstaltungen theoretisch eingeführt werden, in den fachdidaktischen Lehrveranstaltungen vertieft und auch zu anderen Fächern in Beziehung gesetzt werden sowie in der Fachpraxis als Orientierungsrahmen für Feedback und Beurteilung genutzt werden.

5 Fazit

Wenn Forschungsfelder im Zuge ihrer Weiterentwicklung eine gemeinsame Sprache und geteilte konzeptuelle Bezugspunkte in Form von Theorien und Modellen hervorbringen, ist die Voraussetzung für wissenschaftlichen Fortschritt und kumulativen Erkenntnisgewinn gegeben. Die langjährige Auseinandersetzung mit den gleichen theoretischen Grundkonzepten hat in der Unterrichtsqualitätsforschung dazu geführt, dass an vielen Stellen empirisch gut abgesicherte Erkenntnisse gewonnen werden konnten und ein kumulativer Erkenntnisgewinn sichtbar wird. Wenn Forschende sich auf bestehende Befunde beziehen können, erhöht dies die Wahrscheinlichkeit, in der akademischen Welt wahrgenommen und akzeptiert zu werden. Gleichzeitig besteht die Gefahr, dass ein selbst-referentielles akademisches System entsteht, das den unmittelbaren Kontakt zum realen Untersuchungsgegenstand sowie dessen Veränderungen über die Zeit verliert. Diese Gefahr besteht auch für die Unterrichtsqualitätsforschung. Es erscheint wichtig, dass sich die Forschenden neben der Orientierung am eigenen konzeptuellen und methodischen Repertoire schon bei der Gewinnung ihrer Fragestellungen an der Unterrichtswirklichkeit orientieren und Impulse aus der Praxis aufnehmen. Bedenkt man, dass die Erkenntnisse der Unterrichtsqualitätsforschung in weiten Teilen auf Untersuchungen klassischer Unterrichtsettings im Mathematik- und naturwissenschaftlichen Unterricht vergangener Jahrzehnte beru-

hen, wie wenig die Verallgemeinerung der empirischen Befunde gelingt und zudem die Geschwindigkeit, mit der sich Unterricht im Zuge gesellschaftlicher Veränderungen wandelt, scheint das Feld vor der Frage zu stehen, ob es Forschung für die Vergangenheit oder für die Zukunft betreiben möchte. Es scheint erforderlich, gesellschaftliche Wandlungsprozesse, die Einfluss auf die Unterrichtswirklichkeit nehmen, zu antizipieren und daraus in einem iterativen Prozess und aufbauend auf den eigenen Grundlagen notwendige Innovationen für die Unterrichtsqualitätsforschung abzuleiten.

Der vorliegende Beitrag gibt Einblicke in den Versuch eines solchen Arbeitsprozesses im Leibniz-Netzwerk Unterrichtsforschung. Wir stellen nicht den Anspruch auf Vollständigkeit und präsentieren bewusst keine fertigen Ergebnisse. Vielmehr wollen wir Forschende, die sich mit Unterrichtsqualität beschäftigen, zum Nachdenken darüber anregen, was notwendig ist, um die eigene Forschung zukunftsfähig zu machen. Die Zwischenergebnisse der Arbeitsgruppen zeigen deutlich, dass die anvisierten Innovationen kaum von einzelnen Forschenden geleistet werden können. Vielmehr wird es breiter Kooperationen bedürfen: zum einen zwischen Forschenden des Feldes bei der Umsetzung komplexer Studiendesigns, zum anderen zwischen Forschung und Praxis z. B. zur Identifikation entscheidender Problembereiche oder bei der Entwicklung und Ausgestaltung von Forschungsfragen. Wir wünschen uns daher, dass auch andere Forschende die im vorliegenden Beitrag angeschnittenen Themen und Fragen weiterentwickeln und ergänzen. Zudem sehen wir es als hoch bedeutsam an, dass wir alle als in der Unterrichtsqualitätsforschung tätigen Personen die Herausforderungen ernst nehmen und diese verstärkt in unserer zukünftigen Forschung versuchen zu adressieren. Angesichts der großen Veränderungen und Herausforderungen, denen der schulische Unterricht gegenübersteht, geht es in unseren Bemühungen um nichts Geringeres als darum, gemeinsam die Relevanz der Unterrichtsqualitätsforschung für die Zukunft sicherzustellen.

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

- Aebli, H. (1983). *Zwölf Grundformen des Lehrens*. Stuttgart: Klett-Cotta.
- Alp Christ, A., Capon-Sieber, V., Grob, U., & Praetorius, A.-K. (2022). Learning processes and their mediating role between teaching quality and student achievement: A systematic review. *Studies in Educational Evaluation*. <https://doi.org/10.1016/j.stueduc.2022.101209>.
- Altrichter, H., Posch, P., & Spann, H. (2018). *Lehrerinnen und Lehrer erforschen ihren Unterricht: Unterrichtsentwicklung und Unterrichtsevaluation durch Aktionsforschung*. UTB.
- Autorengruppe Bildungsberichterstattung (2020). *Bildung in Deutschland 2020: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. wbv Media. <https://doi.org/10.3278/6001820gw>.
- Batzel, A., Bohl, T., Kleinknecht, M., Leuders, T., & Ehret, C. (2013). Kognitive Aktivierung im Unterricht mit leistungsschwächeren Schülerinnen und Schülern. Theoretische Grundlagen, methodisches Vorgehen und erste Ergebnisse. In U. Riegel & K. Macha (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (S. 97–113). Münster: Waxmann.
- Bauer, J., Prenzel, M., & Renkl, A. (2015). Evidenzbasierte Praxis – im Lehrberuf?! Einführung in den Thementeil. *Unterrichtswissenschaft*, 43(3), 188–192. <https://doi.org/10.3262/UW1503188>.
- Baumert, J., & Kunter, M. (2006). Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/S11618-006-0165-2>.
- Bayrhuber, H., Abraham, U., Frederking, V., Jank, W., Rothgangel, M., & Vollmer, H.J. (2017). *Auf dem Weg zu einer allgemeinen Fachdidaktik (Fachdidaktische Forschungen)*. Münster: Waxmann.
- Berliner, D.C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56(3), 205–213. <https://doi.org/10.1177/0022487105275904>.
- Bikner-Ahsbals, A., & Prediger, S. (2010). Networking of theories—An approach for exploiting the diversity of theoretical approaches. In *Theories of mathematics education* (S. 483–506). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-00742-2_46.
- Black, P.J., & William, D. (1998). Assessment and classroom learning. Assessment in education. *Principles Policy and Practice*, 5(1), 7–73. <https://doi.org/10.1080/0969595980050102>.
- Blömeke, S., Eichler, D., & Müller, C. (2003). Rekonstruktion kognitiver Strukturen von Lehrpersonen als Herausforderung für die empirische Unterrichtsforschung. Theoretische und methodologische Überlegungen zu Chancen und Grenzen von Videostudien. *Unterrichtswissenschaft*, 31, 103–121. <https://doi.org/10.25656/01:6774>.
- Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2022.101600>.
- Böttger, I. (2006). *Beteiligung fördern durch Zukunftswerkstätten & Zukunftskonferenzen*. Berlin: BLK. <https://doi.org/10.25656/01:141>.
- Breidenstein, G., & Rademacher, S. (2017). *Individualisierung und Kontrolle. Empirische Studien zum geöffneten Unterricht in der Grundschule*. Wiesbaden: Springer.
- Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and instruction*, 21(1), 95–108. <https://doi.org/10.1016/j.learninstruc.2009.11.004>.
- Brühwiler, C., & Leutwyler, B. (2020). Praxisrelevanz von Forschung als gemeinsame Aufgabe von Wissenschaft und Praxis: Entwurf eines Angebots-Nutzungs-Modells. *Beiträge Zur Lehrerinnen-Und Lehrerbildung*, 38(1), 21–36. <https://doi.org/10.25656/01:21772>.
- Brunner, E. (2017). Qualität von Mathematikunterricht. Eine Frage der Perspektive. *Journal für Mathematik-Didaktik*. <https://doi.org/10.1007/s13138-017-0122-z>.
- Charalambous, C. Y., & Praetorius, A.-K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM Mathematics Education*, 50(3), 355–366. <https://doi.org/10.1007/s11858-018-0914-8>.
- Charalambous, C., & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*. <https://doi.org/10.1016/j.stueduc.2020.100894>.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Coburn, C.E., & Penuel, W.R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational researcher*, 45(1), 48–54. <https://doi.org/10.3102/0013189X16631750>.

- Coburn, C. E., & Stein, M. K. (2010). *Research and practice in education: Building alliances, bridging the divide*. Rowman & Littlefield.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 119–142. <https://doi.org/10.3102/01623737025002119>.
- Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, 43, 161–173. <https://doi.org/10.1080/00461520802178466>.
- Corno, L., & Snow, R. (1986). Adaptive teaching to individual differences among learners. In M. Wittrock (Hrsg.), *Handbook of research on teaching* (S. 605–629). New York: Macmillan.
- Cortina, K., Miller, K., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science & Mathematics Education*, 13, 389–403. <https://doi.org/10.1007/s10763-014-9610-5>.
- Darling-Hammond, L. (2021). Defining teaching quality around the world. *European Journal of Teacher Education*, 44(3), 295–308. <https://doi.org/10.1080/02619768.2021.1919080>.
- Decristan, J., & Hartmann, U. (2016). Das Projekt „wissenschaft(f)tpraxis“. Guter Unterricht im Dialog zwischen Bildungsforschung und Bildungspraxis. *Lernende Schule*, 19(73), 17–19. <https://doi.org/10.25656/01:14937>.
- Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., & Hardy, I. (2015). Embedded formative assessment and classroom process quality: How do they interact in promoting science understanding? *American Educational Research Journal*, 52(6), 1133–1159. <https://doi.org/10.3102/0002831215596412>.
- Decristan, J., Kunter, M., Fauth, B., Büttner, G., Hardy, I., & Hertel, S. (2016). What role does instructional quality play for elementary school children's science competence? A focus on students at risk. *Journal for Educational Research Online*, 8(1), 66–89. <https://doi.org/10.25656/01:12032>
- Decristan, J., Dumont, H., & Fauth, B. (2017). *Adaptivität und Unterrichtsqualität im individualisierten Unterricht*. Unveröffentlichter Projektantrag an die Robert Bosch Stiftung und die Deutsche Schulkademie
- Decristan, J., Hess, M., Holzberger, D., & Praetorius, A. K. (2020). Oberflächen- und Tiefenmerkmale: eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. *Zeitschrift für Pädagogik. Beiheft*, 66(1), 102–116. <https://doi.org/10.3262/ZPB2001102>.
- Decristan, J., Fauth, B., & Kunter, M. (2022). Die Bedeutung individueller Merkmale und konstruktiver Unterstützung der Lehrkraft für die soziale Integration von Schülerinnen und Schülern im Mathematikunterricht der Sekundarstufe. *Zeitschrift für pädagogische Psychologie*, 36(1–2), 85–100. <https://doi.org/10.1024/1010-0652/a000329>.
- Donovan, S., Snow, C., & Huyghe, A. (2021). Differentiating research-practice partnerships: Affordances, constraints, criteria, and strategies for achieving success. *Studies in Educational Evaluation*. <https://doi.org/10.1016/j.stueduc.2021.101083>.
- Dorfner, T., Förtsch, C., & Neuhaus, B. J. (2018). Effects of three basic dimensions of instructional quality on students' situational interest in sixth-grade biology instruction. *Learning and Instruction*, 56, 42–53. <https://doi.org/10.1016/j.learninstruc.2018.03.001>.
- Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit*. Münster: Waxmann.
- Dumont, H. (2019). Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 22(2), 249–277. <https://doi.org/10.1007/s11618-018-0840-0>.
- Ehmke, T., Hammer, S., Köker, A., Ohm, U., & Koch-Priewe, B. (Hrsg.). (2018). *Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache*. Münster: Waxmann.
- Eid, M., Holtmann, J., Santangelo, P., & Ebner-Priemer, U. (2017). On the definition of latent-state-trait models with autoregressive effects: Insights from LST-R theory. *European Journal of Psychological Assessment*, 33(4), 285–295. <https://doi.org/10.1027/1015-5759/a000435>.
- Einsiedler, W. (1997). Unterrichtsqualität in der Grundschule. Empirische Grundlagen und Programmatik. In E. Glumpler & S. Luchtenberg (Hrsg.), *Jahrbuch Grundschulforschung* Bd. 1.
- Enenkiel, P., Bartel, M.-E., Walz, M., & Roth, J. (2022). Diagnostische Fähigkeiten mit der videobasierten Lernumgebung ViviAn fördern. *Journal für Mathematik-Didaktik*, 43(1), 67–99. <https://doi.org/10.1007/s13138-022-00204-y>.
- Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher*, 47(4), 235–245. <https://doi.org/10.3102/0013189X18761042>.
- Fauth, B., & Leuders, T. (2018). *Kognitive Aktivierung im Unterricht*. Landesinstitut für Schulentwicklung Stuttgart (LS).

- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28, 127–137. <https://doi.org/10.1024/1010-0652/a000129>.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2018). Exploring teacher popularity: associations with teacher characteristics and student outcomes in primary school. *Social psychology of education*, 21(5), 1225–1249. <https://doi.org/10.1007/s11218-018-9462-x>.
- Fauth, B., Göllner, R., Lenske, L., Praetorius, A., & Wagner, W. (2020). Who sees what? Theoretical considerations on the measurement of teaching quality from different perspectives. *Zeitschrift für Pädagogik*, 66, 138–155. <https://doi.org/10.3262/ZPB2001138>.
- Fauth, B., Atlay, C., Dumont, H., & Decristan, J. (2021a). Does what you get depend on who you are with? Effects of student composition on teaching quality. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2020.101355>.
- Fauth, B., Herbein, E., & Maier, J. L. (2021b). *Beobachtungsmニュアル zum Unterrichtsfeedbackbogen Tiefenstrukturen*. Baden-Württemberg (Institut für Bildungsanalysen).
- Fend, H. (1981). *Theorie der Schule*. München: Urban & Schwarzenberg.
- Fenstermacher, G. D., & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record*, 107(1), 186–213. <https://doi.org/10.1111/j.1467-9620.2005.00462.x>.
- Gebhardt, M., Diehl, K., & Mühlhng, A. (2016). Online Lernverlaufsmessung für alle SchülerInnen in inklusiven Klassen. *Zeitschrift für Heilpädagogik*, 67(10), 444–454. <https://doi.org/10.5283/epub.43877>.
- Geher, K., & Nusser, L. (2020). Binnendifferenzierender Deutschunterricht und dessen Einfluss auf die Lesekompetenzentwicklung in der Sekundarstufe I. *Journal for educational research online*, 12(2), 166–189. <https://doi.org/10.25656/01:20976>.
- Gläser-Zikuda, M., Seidel, T., Rohlf, C., Gröschner, A., & Ziegelbauer, S. (Hrsg.). (2012). *Mixed Methods in der empirischen Bildungsforschung*. Münster: Waxmann.
- Gläser-Zikuda, M., Hagenauer, G., Hofmann, F., & Wolf, N. (2018). Reflexion in Lehr-Lernprozessen. In M. Harring, C. Rohlf, & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 516–528). Münster: Waxmann.
- Göbel, K., & Buchwald, P. (2017). *Interkulturalität und Schule: Migration – Heterogenität – Bildung*. UTB.
- Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., Kasneci, E., & Trautwein, U. (2021). Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educational Psychology Review*, 33, 27–49. <https://doi.org/10.1007/s10648-019-09514-z>.
- Grossman, P. (Hrsg.). (2018). *Teaching core practices in teacher education*. Cambridge: Harvard Education Press.
- Hardy, I., Hertel, S., Kunter, M., Klieme, E., Warwas, J., & Büttner, G. (2011). Adaptive Lerngelegenheiten in der Grundschule: Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 57(6), 819–833. <https://doi.org/10.25656/01:8783>.
- Hartmann, U., & Decristan, J. (2018). Brokering activities and learning mechanisms at the boundary of educational research and school practice. *Teaching and Teacher Education*, 74, 114–124. <https://doi.org/10.1016/j.tate.2018.04.016>.
- Hartmann, U., & Kunter, M. (2022). Mehr Praxis in der Bildungsforschung? Eine Studie zu Praxisperspektiven in Forschungsprojekten. *Bildungsforschung*. <https://doi.org/10.25656/01:25469>.
- Hartmann, U., Decristan, J., & Klieme, E. (2016). Unterricht als Feld evidenzbasierter Bildungspraxis? Herausforderungen und Potenziale für einen wechselseitigen Austausch von Wissenschaft und Schulpraxis. *Zeitschrift für Erziehungswissenschaft*, 19(2), 179–199. https://doi.org/10.1007/978-3-658-13785-4_11.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Helmke, A. (2003). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern*. Seelze: Kallmeyersche Verlagsbuchhandlung.
- Helmke, A. (2007). Lernprozesse anregen und steuern. Was wissen wir über Klarheit und Strukturiertheit? *Pädagogik*, 59(6), 44–47.
- Helmke, A. (2022). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.
- Herbert, B., Fischer, J., & Klieme, E. (2022). How valid are student perceptions of teaching quality across education systems? *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2022.101652>.

- Herrmann, C., Gogoll, A., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport. In E. Balz, C. Krieger, W.-D. Miethling & P. Wolters (Hrsg.), *Empirie des Schulsports* 3. Aufl. Edition Schulsport, (Bd. 20, S. 61–81). Meyer & Meyer.
- Hesse, F., & Winkler, I. (2022). Fachliche Qualität im Literaturunterricht. *Zeitschrift für Sprachlich-Literarisches Lernen und Deutschdidaktik*. <https://doi.org/10.46586/SLLD.Z.2022.9605>.
- Hesse, I., & Latzko, B. (2017). *Diagnostik für Lehrkräfte*. Utb.
- Hilbert, S., Coors, S., Kraus, E., Bischl, B., Lindl, A., Frei, M., Wild, J., Krauss, S., Goretzko, D., & Stachl, C. (2021). Machine learning for the educational sciences. *Review of Education*, 9, 1–39. <https://doi.org/10.1002/rev3.3310>.
- Huang, Y., Richter, E., Kleickmann, T., Wipke, A., & Richter, D. (2021). Classroom complexity affects student teachers' behavior in a VR classroom. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2020.104100>.
- Johnson, A. (2011). *A short guide to action research* (4. Aufl.). London: Pearson.
- Jones, S., Farrington, C., Jagers, R., & Brackett, M. (2018). The National Commission on Social, Emotional and Cognitive Development. A research agenda for the next generation. The Aspen Institute. <https://www.nationalcommissionresources.com/product/research-agenda>. Zugegriffen: 25. Juni 2022
- Jungk, R., & Müllert, N. R. (1989). *Die Zukunftswerkstatt*. München: Heyne.
- Käfer, J., Herbein, E., & Fauth, B. (2021). *Formatives Feedback im Unterricht*. Stuttgart: Institut für Bildungsanalysen Baden-Württemberg (IBBW).
- Kellermann, C., Nachbauer, N., Gärtner, H., & Thiel, F. (2022). Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften – Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*. <https://doi.org/10.1007/s11618-022-01087-y>.
- Kleickmann, T., Tröbst, S., Jonen, A., Vehmeyer, J., & Möller, K. (2016). The effects of expert scaffolding in elementary science professional development on teachers' beliefs and motivations, instructional practices, and student achievement. *Journal of Educational Psychology*, 108, 21–42. <https://doi.org/10.1037/edu0000041>.
- Kleickmann, T., Steffensky, M., & Praetorius, A.-K. (2020). Quality of teaching in science education: more than three basic dimensions? *Zeitschrift für Pädagogik, Beiheft*, 66, 37–55. <https://doi.org/10.3262/ZPB2001037>.
- Kleinknecht, M., Krammer, J., & Seel, A. (Hrsg.). (2022). CORE-Practices in der Lehrer*innenbildung [Themenheft]. *Journal für Lehrer*innenbildung*, 22(3). <https://doi.org/10.35468/jlb-03-2022-ed1>
- Klieme, E. (2006). Empirische Unterrichtsforschung: aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52(6), 765–773. <https://doi.org/10.25656/01:4487>.
- Klieme, E. (2019). Unterrichtsqualität. In M. Harring, C. Rohlf's & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 393–408). Münster: Waxmann.
- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237. <https://doi.org/10.25656/01:4348>.
- Klieme, E., & Warwas, J. (2011). Konzepte der individuellen Förderung. *Zeitschrift für Pädagogik*, 57(6), 805–818. <https://doi.org/10.25656/01:8782>.
- Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts Pythagoras. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 128–146). Münster: Waxmann.
- 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 & P. Najvar (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster: Waxmann.
- Klieme, E., Bürgermeister, A., Harks, B., Blum, W., Leiß, D., & Rakoczy, K. (2010). Leistungsbeurteilung und Kompetenzmodellierung im Mathematikunterricht. In *Projekt Co2CA* (S. 64–74). <https://doi.org/10.25656/01:3361>.
- Köhler, C., Kuger, S., Naumann, A., & Hartig, J. (2020). Multilevel models for evaluating the effectiveness of teaching: Conceptual and methodological considerations. *Zeitschrift für Pädagogik. Beiheft*, 66, 197–209. <https://doi.org/10.3262/ZPB2001197>.
- Köhler, C., Hartig, J., & Naumann, A. (2021). Detecting instruction effects—Deciding between covariance analytical and change-score approach. *Educational Psychology Review*, 33(3), 1191–1211. <https://doi.org/10.1007/s10648-020-09590-6>.

- Kollar, I., Fischer, F., & Hesse, F. (2006). Computer-supported cooperation scripts—A conceptual analysis. *Educational Psychology Review*, 18, 159–185. <https://doi.org/10.1007/s10648-006-9007-2>.
- Köller, O., Thiel, F., Ackeren, I. V., Anders, Y., Becker-Mrotzek, M., Cress, U., & Lewalter, D. (2022). *Basale Kompetenzen vermitteln – Bildungschancen sichern. Perspektiven für die Grundschule. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz (SWK)*. <https://doi.org/10.25656/01.25542>.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251. <https://doi.org/10.1007/s10984-006-9015-7>.
- Kunter, M., & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Münster: Waxmann.
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Bd. 3895. Stuttgart: UTB.
- Kunter, M., & Voss, T. (2013). The model of instructional quality in COACTIV: A multicriteria analysis. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (S. 97–124). New York: Springer. https://doi.org/10.1007/978-1-4614-5149-5_6.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>.
- Laging, R. (2022). Bewegung als Aufgabe. *ZSF Zeitschrift für sportpädagogische Forschung*, 10(1), 28–51. <https://doi.org/10.5771/2196-5218-2022-1>.
- Leuders, T., Loibl, K., Sommerhoff, D., Herppich, S., & Praetorius, A.-K. (2022). Toward an overarching framework for systematizing research perspectives on diagnostic thinking and practice. *Journal Für Mathematik-Didaktik*, 43(1), 13–38. <https://doi.org/10.1007/s13138-022-00199-6>.
- Lindmeier, A., & Heinze, A. (2020). Die fachdidaktische Perspektive in der Unterrichtsqualitätsforschung: (bisher) ignoriert, implizit enthalten oder nicht relevant? *Zeitschrift für Pädagogik. Beiheft*, 66(1), 255–268.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. AllemannGhionda & E. Terhart (Hrsg.), *Kompetenz und Kompetenzentwicklung von Lehrerinnen und Lehrern*. Beiheft der Zeitschrift für Pädagogik, (Bd. 51, S. 47–70). Weinheim, Basel: Beltz.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie*. Heidelberg: Springer.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>.
- Lipowsky, F., Drollinger-Vetter, B., Klieme, E., Pauli, C., & Reusser, K. (2018). Generische und fachdidaktische Dimensionen von Unterrichtsqualität – Zwei Seiten einer Medaille? In M. Martens, K. Rabenstein, K. Bräu, M. Fetzer, H. Gresch, I. Hardy & C. Schelle (Hrsg.), *Konstruktionen von Fachlichkeit* (S. 183–202). Bad Heilbrunn: Klinkhardt.
- Meyer, H. (2015). *Unterrichtsentwicklung*. Berlin: Cornelsen.
- Meyer, H., & Junghans, C. (2022). *Unterrichtsmethoden. Theorieband* (20. Aufl.). Berlin: Cornelsen.
- Mötteli, C., Grob, U., Pauli, C., Reusser, K., & Stebler, R. (2022). „Choice and voice“ in Schulen mit personalisierten Lernkonzepten aus Sicht der Schülerinnen und Schüler. *Unterrichtswissenschaft*, 50(2), 287–308. <https://doi.org/10.1007/s42010-021-00122-x>.
- Mu, J., Bayrak, A., & Ufer, S. (2022). Conceptualizing and measuring instructional quality in mathematics education: A systematic literature review. *Frontiers in Education*. <https://doi.org/10.3389/educ.2022.994739>.
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School effectiveness and school improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>.
- Oser, F., & Baeriswyl, F. (2001). Choreographies of teaching. Bridging instruction to learning. In V. Richardson (Hrsg.), *AERA's handbook of research on teaching*. Washington: American Educational Research Association.
- Pant, H. A. (2014). Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 17, 79–99. <https://doi.org/10.1007/s11618-014-0510-9>.

- Pauli, C. (2020). Kommentar zum Themenblock „Oberflächen- und Tiefenstruktur des Unterrichts“. *Zeitschrift für Pädagogik. Beiheft*, 66(1), 132–137. <https://doi.org/10.3262/ZPB2001132>.
- Pauli, C., & Reusser, K. (2000). Zur Rolle der Lehrperson beim kooperativen Lernen. *Schweizerische Zeitschrift für Bildungswissenschaften*, 22(3), 421–442. <https://doi.org/10.25656/01:3747>.
- Praetorius, A.-K., & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188. <https://doi.org/10.1007/s42010-021-00119-6>.
- Praetorius, A.-K., & Kleickmann, T. (2022). Nutzung von Lerngelegenheiten im Unterricht: Konzeptuelle und methodische Zugänge und Herausforderungen – Einführung in den Thementeil. *Unterrichtswissenschaft*, 50(2), 149–155. <https://doi.org/10.1007/s42010-022-00145-y>.
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM Mathematics Education*, 50, 407–426. <https://doi.org/10.1007/s11858-018-0918-4>.
- Praetorius, A.-K., Grünkorn, J., & Klieme, E. (2020a). Empirische Forschung zu Unterrichtsqualität: theoretische Grundfragen und quantitative Modellierungen. *Zeitschrift für Pädagogik. Beiheft*. <https://doi.org/10.3262/ZPB2001009>.
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B., & Nehring, A. (2020b). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>.
- Praetorius, A.-K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Taut, S., & Charalambous, C. (2020c). 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(Beiheft 1/20), 15–36. <https://doi.org/10.3262/ZPB2001015>.
- Praetorius, A.-K., Martens, M., & Brinkmann, M. (2021). Unterrichtsqualität aus Sicht der quantitativen und qualitativen Unterrichtsforschung. In T. Hascher, T.-S. Idel & W. Helsper (Hrsg.), *Handbuch Schulforschung* (S. 1–20). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-24734-8_40-1.
- Prediger, S. (2011). Vorstellungsentwicklungsprozesse initiieren und untersuchen. *Der Mathematikunterricht*, 57(3), 5–14.
- Rakoczy, K., Klieme, E., Drollinger-Vetter, B., Lipowsky, F., Pauli, C., & Reusser, K. (2007). Structure as a quality feature in mathematics instruction of the learning environment vs. a structured presentation of learning content. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (S. 101–120). Münster: Waxmann.
- Rakoczy, K., Frick, U., Weiß-Wittstadt, S., Tallon, M., & Wagner, E. (2022). Einmal begeistert, immer begeistert? Eine Experience-Sampling Studie zur wahrgenommen Unterrichtsqualität und Motivation von Schülerinnen und Schülern im Kunstunterricht. *Unterrichtswissenschaft*, 50(2), 211–236. <https://doi.org/10.1007/s42010-022-00147-w>.
- Reusser, K. (2018). Allgemeine Didaktik – quo vadis? *Beiträge zur Lehrerinnen- und Lehrerbildung*, 36(3), 311–328. <https://doi.org/10.25656/01:18852>.
- Reusser, K., & Pauli, C. (2021). Unterrichtsqualität ist immer generisch und fachspezifisch. Ein Kommentar aus kognitions- und lehr-lerntheoretischer Sicht. *Unterrichtswissenschaft*, 49(2), 189–202. <https://doi.org/10.1007/s42010-021-00117-8>.
- Rothgangel, M., Abraham, U., Bayrhuber, H., Frederking, V., Jank, W., & Vollmer, H.J. (Hrsg.). (2020). *Lernen im Fach und über das Fach hinaus: Bestandsaufnahmen und Forschungsperspektiven aus 17 Fachdidaktiken im Vergleich*. Allgemeine Fachdidaktik, Bd. 2. Waxmann.
- Salaschek, M., Zeuch, N., & Souvignier, E. (2014). Mathematics growth trajectories in first grade: Cumulative vs. compensatory patterns and the role of number sense. *Learning and Individual Differences*, 35, 103–112. <https://doi.org/10.1016/j.lindif.2014.06.009>.
- Schadl, C. (2020). *Individuelle Lernvoraussetzungen für den Erwerb des Bruchzahlkonzepts: Strukturanalysen und Untersuchung der längsschnittlichen Prädiktivität*. Waxmann.
- Schalk, L., & Stern, E. (2018). Wer erklärt eigentlich was wie? Gedanken zum Disziplinenschwengel in der Lehrerinnen- und Lehrerbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 36(3), 398–406. <https://doi.org/10.25656/01:18860>.
- Shank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale: Erlbaum.
- Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM Mathematics Education*, 48(1–2), 29–40. <https://doi.org/10.1007/s11858-016-0765-0>.
- Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM Mathematics Education*, 50(3), 475–490. <https://doi.org/10.1007/s11858-018-0917-5>.

- Schnack, J. (2021). Kognitiv Aktivieren. *Pädagogik*, 11, 202.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie: Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60(6), 850–866. <https://doi.org/10.1007/s11858-018-0917-5>.
- Sliwka, A., Klopsch, B., & Dumont, H. (2022). *Konstruktive Unterstützung im Unterricht*. Reihe Wirksamer Unterricht, Bd. 3. Stuttgart: Institut für Bildungsanalysen Baden-Württemberg.
- Somekh, B. (2010). The Collaborative Action Research Network: 30 years of agency in developing educational action research. *Educational Action Research*, 18(1), 103–121. <https://doi.org/10.1080/09650790903484566>.
- Souvignier, E., Förster, N., Hebbecker, K., Peters, M., & Salaschek, M. (2020). Online-Begleitung von Lernprozessen mit quop. Lernverlaufsdiagnostik und Rückmeldung an die Lernenden. *Schulmanagement-Online*, 2020(4), 11–13.
- Stebler, R., Pauli, C., & Reusser, K. (2018). Personalisiertes Lernen. Zur Analyse eines Bildungsschlagwortes und erste Ergebnisse aus der perLen-Studie. *Zeitschrift für Pädagogik*, 64(2), 159–178. <https://doi.org/10.25656/01:21816>.
- Steffens, U., & Messner, R. (Hrsg.). (2019). *Unterrichtsqualität: Konzepte und Bilanzen gelingenden Lehrens und Lernens*. Grundlagen der Qualität von Schule, Bd. 3. Waxmann.
- Szogs, M., Korneck, F., Krüger, M., Oettinghaus, L., & Kunter, M. (2016). Kognitive Aktivierung in standardisierten Unterrichtsminiaturen. In C. Maurer (Hrsg.), *Authentizität und Lernen – das Fach in der Fachdidaktik* (S. 605–610). Regensburg: Gesellschaft für Didaktik der Chemie und Physik.
- Taut, S., & Rakoczy, K. (2016). Observing instructional quality in the context of school evaluation. *Learning and Instruction*, 46, 45–60. <https://doi.org/10.1016/j.learninstruc.2016.08.003>.
- Terhart, E. (2002). Fremde Schwestern: Zum Verhältnis von Allgemeiner Didaktik und empirischer Lehr-Lern-Forschung. *Zeitschrift für Pädagogische Psychologie*, 16, 77–86.
- Terhart, E. (2020). Unterrichtsqualität zwischen Theorie und Empirie. *Zeitschrift für Pädagogik. Beiheft*, 66(1), 223–234.
- Troll, B., Heil, C., Pietsch, M., & Besser, M. (2022). Netzwerkbasierte Betrachtung von ko-konstruktiven Interaktionsprozessen im Unterricht – Ein Ansatz zur Beschreibung und Analyse von Angebot und Nutzung. *Unterrichtswissenschaft*, 50(2), 237–263. <https://doi.org/10.1007/s42010-022-00142-1>.
- Tyack, D., & Tobin, W. (1994). The “grammar” of schooling: Why has it been so hard to change? *American Educational Research Journal*, 31(3), 453–479. <https://doi.org/10.3102/00028312031003453>.
- Vieluf, S. (2013). *ÖGRETMEK OR ENSEÑAR? A cross-national study of teacher's professional competence and teaching quality*. Dissertation. Frankfurt: Goethe-Universität
- Vieluf, S. (2022). Wie, wann und warum nutzen Schüler*innen Lerngelegenheiten im Unterricht? Eine übergreifende Diskussion der Beiträge zum Thementeil. *Unterrichtswissenschaft*. <https://doi.org/10.1007/s42010-022-00144-z>.
- Vieluf, S., Praetorius, A. K., Rakoczy, K., Kleinknecht, M., & Pietsch, M. (2020). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts: ein kritischer Vergleich verschiedener Modellvarianten. *Zeitschrift für Pädagogik. Beiheft*, 66(1), 63–80. <https://doi.org/10.3262/ZPB2001063>.
- Weinbrenner, P. (2002). *Zur Theorie und Praxis von Zukunftswerkstätten. Dokumentation einer Zukunftswerkstatt zum „Lebenslanges Lernen“ in Ahle*
- Wilkes, T., & Stark, R. (2022). Probleme evidenzorientierter Unterrichtspraxis. *Unterrichtswissenschaft*. <https://doi.org/10.1007/s42010-022-00150-1>.
- Wirth, J., Thillmann, H., Künsting, J., Fischer, H.E., & Leutner, D. (2008). Das Schülerexperiment im naturwissenschaftlichen Unterricht. Bedingungen der Lernförderlichkeit einer verbreiteten Lehrmethode aus instruktionspsychologischer Sicht. *Zeitschrift für Pädagogik*, 54, 361–375. <https://doi.org/10.25656/01:4356>.
- Wolff, C.E., Jarodzka, H., & Boshuizen, H.P.A. (2021). Classroom management scripts: A theoretical model contrasting expert and novice teachers' knowledge and awareness of classroom events. *Educational Psychology Review*, 33, 131–148. <https://doi.org/10.1007/s10648-020-09542-0>.
- Zander, L., Kreutzmann, M., & Hannover, B. (2017). Peerbeziehungen im Klassenzimmer. *Zeitschrift für Erziehungswissenschaft*, 20(3), 353–386. <https://doi.org/10.1007/s11618-017-0768-9>.

Affiliations

✉ Lukas Begrich · Mareike Kunter

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main, Deutschland

E-Mail: l.begrich@dipf.de

Mareike Kunter

E-Mail: m.kunter@dipf.de

Anna-Katharina Praetorius

Universität Zürich, Zürich, Schweiz

E-Mail: anna.praetorius@ife.uzh.ch

Jasmin Decristan

Bergische Universität Wuppertal, Wuppertal, Deutschland

E-Mail: decristan@uni-wuppertal.de

Benjamin Fauth

Institut für Bildungsanalysen Baden-Württemberg, Stuttgart, Deutschland

E-Mail: benjamin.fauth@ibbw.kv.bwl.de

Richard Göllner

Universität Tübingen, Tübingen, Deutschland

E-Mail: richard.goellner@uni-tuebingen.de

Christian Herrmann

Pädagogische Hochschule Zürich, Zürich, Schweiz

E-Mail: christian.herrmann@phzh.ch

Marc Kleinknecht

Leuphana Universität Lüneburg, Lüneburg, Deutschland

E-Mail: marc.kleinknecht@leuphana.de

Sandy Taut

Bayrisches Landesamt für Schule, Gunzenhausen, Deutschland

E-Mail: Sandy.Taut@las.bayern.de