



Probleme evidenzorientierter Unterrichtspraxis

Anregungen und Lösungsvorschläge

Theresa Wilkes  · Robin Stark

Eingegangen: 8. Januar 2021 / Überarbeitet: 8. Juni 2022 / Angenommen: 17. Juni 2022
© Der/die Autor(en) 2022

Zusammenfassung In diesem Beitrag soll anhand eines fiktiven Anwendungsbeispiels aus der schulischen Unterrichtspraxis exemplarisch dargestellt werden, mit welchen Barrieren sich Lehrkräfte konfrontiert sehen, wenn sie individuelle unterrichtsbezogene Entscheidungen evidenzbasiert zu treffen versuchen. Es werden Stärken und Schwächen des Konzeptes der Evidenzbasierung sowie unterschiedlicher Arten von Evidenz und Forschungsdesigns (zum Beispiel Meta-Metaanalysen, Metaanalysen, Fallstudien) reflektiert und diskutiert. Unter anderem wird ein Vorgehen vorgeschlagen, welches sich weniger an der gängigen Evidenzhierarchie orientiert, an deren Spitze sich als „Gold-Standard“ systematische Reviews, Metaanalysen und randomisierte kontrollierte Studien finden, sondern vielmehr an übergeordneten Wissensbeständen (d. h. insbesondere an wissenschaftlichen Theorien und Konzepten, aber auch an einschlägigen wissenschaftlichen Ansätzen, Prinzipien und Modellen), welche aus der bildungswissenschaftlichen Forschung hervorgegangen sind. Zudem werden exemplarisch Planungsprozesse skizziert, welche die Qualität der Unterrichtspraxis mit einfachen Mitteln unterstützen können, indem wissenschaftliches Wissen, Erfahrungswissen und Kontextwissen sowie praktisches Handlungswissen integriert werden. Abschließend werden als eine mögliche Förderkomponente evidenzorientierter Unterrichtspraxis verschiedene Möglichkeiten zum weiteren Ausbau von Clearing Houses vorgeschlagen. Zudem wird auf Forschungsdesiderate hinsichtlich der Integration verschiedener, möglicherweise inkohärenter Wissensbestände hingewiesen. Dieser Beitrag hat nicht zum Ziel, die aktuellen Entwicklungen hin zu einer evidenzbasierten Schul- und Unterrichtspraxis in Frage zu stellen. Vielmehr wird eine evidenzorientierte Praxis im Sinne einer Orientierung an *übergeordneten Wissensbeständen* ausdrücklich befürwortet.

Theresa Wilkes (✉) · Robin Stark
Bildungswissenschaften, Persönlichkeitsentwicklung und Erziehung, Campus, Universität des Saarlandes, Geb. A4 2, 66123 Saarbrücken, Deutschland
E-Mail: theresa.wilkes@uni-saarland.de

Schlüsselwörter Lehrer*innenbildung · Evidenz · Evidenzbasierte/
Evidenzorientierte (Unterrichts-)Praxis · Evidenzhierarchie · Theorie-Praxis-
Verknüpfung

Problems of an evidence-oriented educational practice

Suggestions and solution approaches

Abstract In this paper, a fictitious example from teaching practice at school will be used to illustrate the barriers a teacher faces when trying to make individual teaching-related decisions based on scientific evidence. The strengths and weaknesses of the concept of evidence-based educational practice as well as of different types of evidence and research designs (e.g., meta-meta-analyses, meta-analyses, case studies) will be considered and discussed. Among other things, a procedure will be proposed that is oriented less towards the common hierarchy of evidence, at the top of which systematic reviews, meta-analyses and randomised controlled trials are found as the “gold standard”. Instead, a procedure will be proposed that is oriented towards more general bodies of knowledge (i.e., especially scientific theories and concepts, but also relevant scientific approaches, principles, and models), which education research has put forth. Moreover, exemplary planning processes will be outlined that can support the quality of teaching practice in simple ways by integrating scientific knowledge, experiential knowledge, context knowledge and practical action knowledge. Finally, various possibilities for the further development of Clearing Houses are suggested as a component of fostering evidence-oriented teaching practice. Research desiderata concerning the integration of different, possibly incoherent bodies of knowledge will be pointed out. The paper does not aim to question current developments in evidence-based teaching practice. Rather, an evidence-oriented practice in the sense of an orientation to higher-level bodies of knowledge is explicitly advocated.

Keywords Teacher education · Evidence · Evidence-based/ -oriented teaching practice · Evidence hierarchy · Theory-practice linkage

1 Evidenzbasierung und -orientierung im Bildungsbereich

Die Kernaufgabe von Lehrkräften bildet gemäß den Standards für die Lehrerbildung (KMK 2014, S. 2) die „gezielte und nach wissenschaftlichen Erkenntnissen gestaltete Planung, Organisation und Reflexion von Lehr- und Lernprozessen sowie ihre individuelle Bewertung und systemische Evaluation“. Lehrkräfte sollen demnach unterrichtspraktische Entscheidungen nicht nur vor dem Hintergrund ihrer individuellen Erfahrungsbestände treffen, sondern auch auf der Basis wissenschaftlicher Theorien und empirischer Befunde. Damit definiert sich der Lehrberuf als eine evidenzbasierte Profession (Hartmann et al. 2016; Stark 2017).

Die Idee der Evidenzbasierung hat ihren Ursprung in der Medizin und deren Forderung, bei Entscheidungen über die Behandlung von Patient*innen neben der eigenen Expertise und den Präferenzen der Patient*innen die besten verfügbaren

empirischen Befunde zu berücksichtigen (Hastall und Lang 2019; Sackett et al. 1996). Spätestens seit dem sogenannten *PISA-Schock* hat das Konzept der Evidenzbasierung – mit der Hoffnung auf eine Verbesserung der Unterrichtspraxis – sowohl auf der Ebene politischer Steuerung als auch im Kontext individueller unterrichtsbezogener Entscheidungen Einzug in den Bildungsbereich gehalten (Bromme et al. 2014; Hartmann et al. 2016; Scharlau 2019).

Bezüglich der Frage, welche Wissensbestände seitens einer Lehrkraft vorhanden sein sollten, hat sich ausgehend von den theoretischen Arbeiten von Shulman (1987) die Einteilung in fachliches, fachdidaktisches und pädagogisches Wissen etabliert (z. B. Kunter et al. 2011). Auf Basis theoretischer und empirischer Arbeiten zu der Expertise und den Wissensbeständen von Lehrkräften, die für die Bewältigung schulischer und unterrichtlicher Anforderungen notwendig sind, wurden Modelle professioneller Handlungskompetenz generiert, die eben solche relevanten Voraussetzungen beschreiben (z. B. Baumert und Kunter 2006; vgl. König 2016). Die Forderung nach einer evidenzbasierten Unterrichtspraxis spiegelt sich hierbei insbesondere in denjenigen Komponenten des Professionswissens wider, die das Handeln von Lehrkräften direkt betreffen, also im fachdidaktischen und im pädagogischen Wissen.

Die Aufgabe der Lehrkräfteausbildung, angehende Lehrkräfte dazu zu befähigen, ihr Handeln auf Basis verallgemeinerbaren wissenschaftlichen Fachwissens sowie fachdidaktischen und pädagogischen Wissens zu planen, zu erklären und zu reflektieren, bringt es mit sich, dass hierfür relevante Grundsteine in der universitären Ausbildungsphase „forschungs- und nicht subjektiv-erfahrungsfundiert“ gelegt werden (Rothland 2016, S. 9). Dies impliziert, dass Wissen auf der Basis empirischer Befunde so zu vermitteln ist, dass es die ausgebildeten Lehrkräfte in der schulischen Praxis erfolgreich anwenden können (Bauer et al. 2015; Gröschner und Seidel 2015; Prenzel 2020).

Die Relevanz der Evidenzbasierung als leitendes Prinzip und bedeutsame Professionalisierungskomponente der Lehrkräfteausbildung (Gogolin et al. 2020) wird insbesondere im Licht solcher Studien deutlich, welche die Nutzung von Evidenzen durch angehende wie berufstätige Lehrkräfte im Kontext unterrichtspraktischer Fragen in den Blick nehmen (Csanadi et al. 2021; Hinzke et al. 2020; Kiemer und Kollar 2018; 2021; Rochnia und Trempler 2019; Thomm et al. 2021c) sowie mögliche Barrieren (Fischer 2021; Lysenko et al. 2014; Zlatkin-Troitschanskaia et al. 2016) und entsprechende Fördermöglichkeiten untersuchen (Wagner et al. 2018; Wekerle und Kollar 2021; Wenglein et al. 2015).

Potenzielle Barrieren wurden mit Blick auf kognitive (*skill*; Bauer et al. 2017) und affektive Dimensionen (*will*; Bauer et al. 2017) vor allem auf Seiten der Anwender*innen selbst, aber auch auf Seiten der Ausbilder*innen (Diery et al. 2020b; 2021) untersucht. Als kognitive Barrieren seitens der Anwender*innen wurden insbesondere defizitäre Wissensbestände (Neuweg 2014), Schwierigkeiten bei der Beurteilung (Thomm und Bromme 2016; Trempler et al. 2015) und bei der Anwendung wissenschaftlichen Wissens (Kiemer und Kollar 2018; Wagner et al. 2018; Wenglein et al. 2015) sowie Fehlkonzepte, Fehlinterpretationen und/oder fehlerhafte Anpassungen an subjektive Theorien (Asberger et al. 2020; Thomm et al. 2021a) identifiziert. Als affektive Barrieren wurden insbesondere die Motivation zur Beschäftigung mit Evidenz sowie wissenschaftsbezogene Einstellungen und Überzeugungen fokussiert

(Bleck und Lipowsky 2020; Egger und Groß Ophoff 2020; Kiemer und Kollar 2021; Rochnia und Trempler 2019). Weitere Barrieren wurden in beschränkten Ressourcen, vor allem in einem Mangel an Zeit und in fehlenden Zugangsmöglichkeiten gesehen (Thomm et al. 2021b).

Eine nicht zu vernachlässigende Barriere der Evidenzbasierung, die trotz vieler kritischer Perspektiven (zum Beispiel Hartmann et al. 2016) immer noch unzureichend reflektiert wird, gründet im Konzept der Evidenzbasierung sowie in Merkmalen von Evidenzen selbst. So herrscht immer noch keine Klarheit darüber, (1) *inwieweit* und (2) auf *welche Arten* von Evidenz sich Lehrkräfte in unterrichtsbezogenen Fragestellungen berufen sollten. Bei der Beantwortung dieser Fragen, sofern es überhaupt eine eindeutige Antwort geben kann, steht man zunächst vor der Schwierigkeit einer Begriffsdiffusität (Hinze et al. 2020).

So hat sich (1) neben einer strikteren Auffassung von *Evidenzbasierung* im Sinne einer „Ableitung praktischer Maßnahmen aus nachweisbaren empirischen Tatsachen“ (Scharlau 2019, S. 106) oder einer 1:1-Implementation evaluierter Programme und Ansätze (Slavin 2020; Slavin et al. 2021), in Anknüpfung an das sogenannte *pragmatistische Evidenz-Paradigma* (Reiss 2015), ein weiteres, moderateres Verständnis im Sinne einer *Evidenzorientierung* oder *-informiertheit* entwickelt. In diesem moderateren Verständnis kommt Evidenz eine handlungsunterstützende und weniger handlungsleitende Funktion zu, auf deren Basis individuelle Erfahrung und Expertise reflektiert, ergänzt und korrigiert werden soll (Bauer et al. 2015, 2017; Biesta 2011; Stark 2017). Zumindest der weitgehend an der Methodologie und Methodik der psychologischen Forschung orientierte bildungswissenschaftliche „Mainstream“, sofern sich ein solcher ausmachen lässt, scheint sich mit der moderateren Lesart angefreundet zu haben (Scharlau 2019; Stark 2017). Darüber hinaus kann (2) der Begriff der *Evidenz* in einem sehr weiten Sinne verstanden werden, der nicht nur wissenschaftliche Theorien und empirische Befunde, sondern auch kollegiale Hospitationen, Schüler*innenfeedbacks oder schulinterne Befragungen einschließt (van Ackeren et al. 2013). Auch bei einer engeren definitorischen Auslegung des Evidenzbegriffs – so merken Hinze und Kolleg*innen (2020, S. 1306) an – muss zwischen Evidenz im Sinne von „wissenschaftlichem Wissen“ und Evidenz im Sinne von „Wissen aus evidenzbasierten Steuerungsinstrumenten“ (Lernstandserhebungen, Schulinspektionen) differenziert werden. Fasst man den Evidenzbegriff nur im Sinne von „wissenschaftlichem Wissen“, ist damit noch nicht geklärt, welches Wissen bzw. welche Evidenzquellen Praktiker*innen als eine *verlässliche* und *belastbare* Entscheidungsgrundlage ansehen können.

Zur Klärung dieser Frage hat sich in der Medizin eine streng hierarchische Ordnung von Forschungsdesigns etabliert, welche auch für den Bildungsbereich diskutiert wird (Bromme et al. 2014; Gräsel 2019). An deren Spitze finden sich als sogenannter *Gold-Standard* systematische Reviews, Metaanalysen und randomisierte kontrollierte Studien (Hastall und Lang 2019), die möglichst darüber Auskunft geben sollen, *was funktioniert* (Joyce und Cartwright 2020). Um diese Evidenzen für Praktiker*innen anwendbarer zu machen, wurden sogenannte *Clearing Houses* eingerichtet, welche die beste einschlägige, aktuell verfügbare Evidenz zu verschiedensten Themengebieten kontinuierlich adressatengerecht aufbereiten (im Bildungsbereich vor allem das *Clearing House Unterricht* der TU München). Das *Clearing*

House Unterricht bietet zudem ein Qualitätsrating, anhand dessen Rückschlüsse auf die Verlässlichkeit der Quellen gezogen werden können (Diery et al. 2020a; Seidel et al. 2020). Darüber hinaus wird am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) aktuell an breit angelegten systematischen Reviews gearbeitet¹. Legte man allerdings die oben angeführte Evidenz-Hierarchie auch im Bildungsbereich als Maßstab zur Beurteilung von Evidenz konsequent an, resultierte daraus, dass ein Großteil der empirischen Forschung aus den noch vergleichsweise jungen Feldern der Erziehungs- und Bildungswissenschaften sowie der Fachdidaktiken strenggenommen als nur wenig *verlässlich* und *belastbar* gelten dürfte (Bromme et al. 2014; Hedges 2018; Malouf und Taymans 2016; Prenzel 2020). Auch wenn man dem *Gold-Standard* eine geringere Relevanz beimisst und stattdessen einem breiteren Verständnis von Evidenzbasierung bzw. -orientierung folgt, ist die Forderung nach einer evidenzorientierten Unterrichtspraxis mit einer Reihe von Problemen behaftet, denen sich der vorliegende Beitrag widmet. Am Beispiel einer fiktiven pädagogischen Anwendungssituation soll zunächst verdeutlicht werden, dass im Zuge aktueller Entwicklungen teilweise unrealistische Anforderungen an Praktiker*innen gestellt werden.

2 Evidenzorientierte Unterrichtspraxis

2.1 Ein fiktives Beispiel: Kooperatives Lernen mit Tablets, ja oder nein?

Als fiktive Praktikerin führen wir eine Referendarin ein, die im Rahmen ihres Prüfungsunterrichts im Fach Englisch in der siebten Klasse eine Tablet-gestützte Unterrichtsreihe zum Thema *Digital Media* durchführen möchte:

Die Referendarin unterhält sich im Lehrer*innenzimmer mit einem erfahrenen Englischlehrer über ihre Ideen. Der Kollege berichtet ihr von seinen negativen Erfahrungen mit den angeschafften Tablets und rät ihr von deren Nutzung ab. Er warnt sie, dass besonders die Lernenden der Unter- und Mittelstufe im Tablet-gestützten Unterricht sehr unruhig würden, sich weniger aktiv am Unterricht beteiligten und sich nicht mehr auf den eigentlichen Inhalt konzentrierten. Außerdem gibt er zu bedenken, dass kooperatives Lernen schwierig umzusetzen sei, da sich jede*r auf das Tablet konzentriere und kaum noch miteinander kommuniziert werde. Diskussionen kämen kaum zustande, weder innerhalb von Kleingruppen noch in Plenumsphasen. Dies sei ihm zufolge gerade im Fach Englisch fatal, da für das Erlernen einer Sprache kooperative Lernformen besonders relevant seien.

Nach dieser Unterhaltung ist die junge Lehrerin verunsichert: Die vorgebrachten Argumente erscheinen ihr durchaus plausibel, dennoch möchte sie ihren Unterricht, und insbesondere auch ihren Prüfungsunterricht, innovativ mithilfe von Tablets gestalten, gerade wenn es inhaltlich um *Digital Media* geht. Sie

¹ Weiterführende Informationen zum Fortschritt des Projekts finden sich online unter: <https://www.dipf.de/de/forschung/forschungsthemen/systematic-reviews>. Letzter Zugriff am 18.05.2022.

unterhält sich mit einem befreundeten Referendar, der die Tablets im Mathematikunterricht bereits erfolgreich genutzt habe, aber bisher nur in Einzelarbeit mit adaptiven Lernprogrammen. Er rät ihr, den kooperativen Einsatz der Tablets einfach *auszuprobieren*. Die Referendarin möchte sich allerdings nicht nur auf die Erfahrungen ihrer Kollegen verlassen oder *ausprobieren*, sondern auch fundierte Argumente für die Nutzung von Tablets in ihrem Unterricht finden. Sie beschließt daher, sich auf die Suche nach geeigneten Quellen zu machen.

Sie erinnert sich, dass vor einiger Zeit eine Rangliste verschiedener Einflussfaktoren auf den Lernerfolg propagiert wurde, die prominente *Hattie-Studie* (Hattie 2009). Da die Rangliste im Internet unter <https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement/> veröffentlicht ist und offenbar laufend aktualisiert wird, kann sie schnell prüfen, ob in der Meta-Metaanalyse auch (kooperatives) Lernen mit digitalen Medien thematisiert wurde. Für „Technology in small groups“ wird auf Platz 189 von 252 eine Effektstärke von $d=0,21$ angegeben (Stand Mai 2022). Wie im Studium gelernt, legt sie zur Interpretation die Konventionen nach Cohen (1988) an, nach welchen der Effekt zwar klein, aber praktisch bedeutsam ist. Jedoch wundert sie sich über die eher *schlechte* Platzierung; „Classroom discussion“, die bei ihrem Englischkollegen aufgrund des Tableteinsatzes ausgeblieben sei, liegen immerhin auf Platz 15 (Stand Mai 2022).

Problematisch ist, dass die Referendarin weder mit (1) aktuellen Diskursen bezüglich der Interpretation von Effektstärken (z. B. Kraft 2020) noch (2) der speziellen methodischen Problematik der *Hattie-Studie* (Renkl 2022; Wecker et al. 2017) vertraut ist. Auch wenn die Konventionen nach Cohen aktuell noch breite Anwendung finden (auch in den Kurzreviews des *Clearing House Unterricht*; beispielsweise Knogler et al. 2018a), ist es umstritten, in Entscheidungsfragen zur Implementation bestimmter Maßnahmen pauschale Schlussfolgerungen ausschließlich auf Basis von berichteten Effektstärken abzuleiten (Kraft 2020). Effektstärken bedürfen einer umfassenden Kontextualisierung, um angemessen interpretiert werden zu können. Beispielsweise ist es wichtig zu wissen, welche Art von Studiendesign der Effektstärkenberechnung zu Grunde lag, wie distal bzw. proximal die verwendeten Outcome-Maße in Bezug auf die Intervention konzeptualisiert waren oder wie reliabel und in welchem zeitlichen Abstand zur Intervention sie gemessen wurden. Zudem sind verschiedene Zusatzinformationen erforderlich, etwa über Stichprobencharakteristika, über mathematische Details der angewandten Standardisierungsprozedur, über das Ausmaß der Unterschiedlichkeit der Interventionen für die Kontroll- und Experimentalgruppe sowie über die Art der Interventionsdurchführung. Auch spielen Kosten- bzw. Effizienzfragen in Bezug auf die umgesetzte Intervention eine Rolle. Schließlich stellen sich auch diverse Fragen hinsichtlich der Replizierbarkeit der berichteten Effekte (Kraft 2020).

Die *Hattie-Studie* stieß zwar in den öffentlichen Medien auf große Resonanz, dennoch ist ein Großteil der Befunde begründeten Zweifeln ausgesetzt (Renkl 2022; Wecker et al. 2017). Eine Aggregation von Studien über die Meta-Ebene hinaus ist bereits mit verschiedenen Problemen behaftet. Bei Hattie (2009) ist sowohl die Datengrundlage als auch sein methodisches Vorgehen kritisch zu sehen. Außerdem

suggeriert die Rangliste, dass bei der Unterrichtsgestaltung möglichst die wirksamste Handlungsoption ausgewählt werden sollte, obwohl nicht nur Handlungsoptionen als Einflussfaktoren aufgelistet sind (z. B. Self-efficacy der Lernenden, Glaubwürdigkeit der Lehrkraft), und die Kontrollbedingungen der Primärstudien keinen Vergleichsmaßstab bieten (Wecker et al. 2017). Zudem stellt sich die Frage, wie es sich mit den Effektstärken verhält, wenn verschiedene Einflussfaktoren gleichzeitig wirksam werden.

Da sich die junge Lehrerin nicht mit dem bisherigen Ergebnis ihrer Recherche zufriedengeben möchte, sucht sie weiter. Dabei merkt sie, dass sie Schwierigkeiten hat, die im Internet zu findenden Quellen hinsichtlich ihrer Verlässlichkeit und Belastbarkeit zu beurteilen. Sie erinnert sich an das *Clearing House Unterricht* der TU München, welches – wenn auch primär für Lehrkräfteauszubildende – Kurzreviews zu einschlägigen Studien, vor allem Metaanalysen, zu verschiedensten Themengebieten anbietet.

Dass Praktiker*innen von sich aus regelmäßig auf Metaanalysen (als Originalarbeiten) zurückgreifen, kann als eher unwahrscheinlich angesehen werden. Metaanalysen sind von Nicht-Fachleuten nicht einfach zu verstehen und ihre Beurteilung sowie die Übertragbarkeit der gewonnenen Erkenntnisse auf die Praxis ist in der Regel sehr anspruchsvoll (Stark 2017). Clearing Houses stellen einen Ansatz dar, um insbesondere Evidenz aus Metaanalysen greifbarer zu machen (Diery et al. 2020a; Seidel et al. 2020). Das *Clearing House Unterricht* richtet sein Angebot aktuell an Lehrkräftebildende und konzentriert sich dabei vor allem auf wissenschaftliche Erkenntnisse zu effektivem Unterricht in den *MINT*-Fächern. Neben einem kurzen Überblick über die jeweilige Metaanalyse werden Informationen zur durchschnittlichen Effektstärke, zur Zielgruppe, zu weiteren Befunden und zu den Hintergründen der Studie berichtet. Außerdem, und dies sollte für potenzielle Anwender*innen besonders hilfreich sein, wird ein Fazit für die Unterrichtspraxis gezogen. Zudem werden ein Studienbeispiel sowie weitere Referenzen und Links vorgeschlagen. Dazu wird ein begleitendes *Rating-Sheet* zur Verfügung gestellt, welches zu jeweils vier Qualitätsdimensionen (Recherche, Auswahl, Erfassung, statistische Analyse) einen Einzelscore sowie einen Gesamtscore zur Qualität der Metaanalyse (Prozentangaben) präsentiert. Diese Bewertungen soll dabei unterstützen, die methodische Verlässlichkeit, die Einhaltung aktueller Standards sowie die Transparenz und Begründung von Entscheidungen besser einschätzen zu können (Diery et al. 2020a; Seidel et al. 2020).

Tatsächlich findet die Referendarin ein Kurzreview, das ihr auf den ersten Blick passend erscheint: ein Review zur Wirksamkeit der Kombination aus kollaborativem Lernen und mobilen digitalen Geräten (Knogler et al. 2018a), in welchem eine Metaanalyse von Sung et al. (2017) aufbereitet wurde. In der Metaanalyse wurde der Einfluss mobiler digitaler Geräte beim kollaborativen Lernen in kleinen Gruppen auf Lernleistung, Einstellungen und Qualität der Zusammenarbeit untersucht. Zusätzlich wurde mittels Moderatoranalysen geprüft, ob die Wirksamkeit der Lernform von Faktoren wie Schulfach, Alter, Gruppengröße, Gruppenzusammensetzung oder Interventionsdauer beeinflusst wird. Viele der

sehr differenziert dargelegten Befunde erscheinen der Lehrerin für ihre Fragestellung relevant: Auf der Basis von 163 Effektstärken ergab sich ein signifikant positiver, mittelgroßer Gesamteffekt ($g = 0,52$); Lernende zeigten beim kollaborativen Lernen mit mobilen Geräten vergleichsweise höhere Leistungen, positivere Einstellungen zum Lernen und eine produktivere Zusammenarbeit. Außerdem war diese Lernform effektiver als individuelles Lernen mit mobiler Technologie oder kollaboratives Lernen ohne digitale Unterstützung. Insbesondere Schüler*innen aus der Sekundarstufe profitierten. Dazu bietet die Metaanalyse Informationen für die konkrete Ausgestaltung: So profitierten vorwiegend (hinsichtlich Persönlichkeitseigenschaften und Leistungsvoraussetzungen) homogene und vierköpfige Gruppen. Kollaborative Ansätze erbrachten größere Effekte als kompetitive (zitiert nach Knogler et al. 2018a).

Die vom *Clearing House Unterricht* vorgenommene Einschätzung der Befunde ruft bei der Referendarin jedoch Bedenken hervor: Es sei davon auszugehen, dass methodische Schwächen und die große Zeitspanne der eingegangenen Primärstudien mitunter zu einer Überschätzung der Effektivität führten. Zwar deuteten die Moderatoranalysen darauf hin, dass kollaboratives Lernen mit mobilen digitalen Geräten unter unterschiedlichen Bedingungen lernwirksam sein könnte, jedoch sei die Anzahl der Studien für die einzelnen Moderatorstufen gering, sodass die Ergebnisse eher als vorläufige Trends, nicht aber als belastbare Befunde betrachtet werden könnten (Knogler et al. 2018a). Das Fazit für die Unterrichtspraxis schließt wie folgt:

„Die Erwartung, dass eine optimale Nutzung der Möglichkeiten mobiler Geräte und ein förderliches Arrangement der Zusammenarbeit zu besseren Lernprozessen und Resultaten im Unterricht führt, bleibt durch die Befunde dieser Metaanalyse bestehen, kann aber durch sie nicht verlässlich belegt werden.“ (Knogler et al. 2018a, S. 5).

Verunsichert hofft die Referendarin durch das Rating-Sheet weitere Informationen über die Verlässlichkeit der Studie erhalten zu können (Knogler et al. 2018b). Jedoch weiß sie nicht, was sie aus den angegebenen Prozentangaben, insbesondere aus dem Gesamturteil von 60%, schließen soll. Zwar gibt das Rating-Sheet Auskunft darüber, wie sich die einzelnen Kategorien definieren, die Lehrerin erhält jedoch keine genaueren Informationen, wie die konkreten Werte für diese Metaanalyse zustande gekommen sind. Sind die Befunde nun verlässlich und belastbar? Ermöglichen sie eine fundierte Entscheidung?

Ob eine bestimmte Maßnahme als wirksam gelten kann, ist insbesondere daran zu beurteilen, ob diese mehrfach im Rahmen experimenteller, kontrollierter Designs untersucht wurde, ob die Effekte sowohl innerhalb und zwischen den Studien konsistent sind, ob keine negativen Effekte überwiegen, ob sich die Effekte zwischen Untergruppen oder unter verschiedenen Bedingungen unterscheiden, und ob die Effekte nur unmittelbar oder auch nachhaltig sind (Taylor et al. 2021). Die Ursachenfindung ist gerade bei uneindeutigen oder widersprüchlichen Befunden in einem umschriebenen Feld sicher eine der Stärken, die Metaanalysen zugeschrieben werden können (Beelmann 2014). Neben weiteren Vorteilen gegenüber Einzelstudien können auf der Grundlage von Moderatoranalysen, die zum Standardrepertoire me-

thodisch anspruchsvollerer Metaanalysen zählen (Beelmann 2014), Ansatzpunkte für die Erklärung unterschiedlicher Ergebnisse der integrierten Einzelstudien gefunden werden. Wie das fiktive Beispiel zeigt, kann aber auch eine anwender*innengerechte Aufbereitung die Schwierigkeiten, die mit der Methodik einer Metaanalyse selbst verbunden sind, nur begrenzt überwinden (Renkl 2022). Aufgenommene Studien unterscheiden sich teilweise in methodischen Details, beispielsweise in ihren Untersuchungsdesigns, einbezogenen Variablen sowie deren Konzeptualisierungen und Operationalisierungen. Zudem finden sich vielfältige Unterschiede in den Untersuchungskontexten und -situationen sowie der Umsetzung und Reichweite der Interventionen, den Stichprobengrößen und deren Zusammensetzung, zum Teil auch in den verwendeten statistischen Analysemethoden. Auch wenn Unterschiede zwischen Vergleichsstudien durch die Analyse von Moderatoreffekten teilweise berücksichtigt werden (Beelmann 2014; Connolly et al. 2018), ist diese grundlegende und komplexe Problematik einer vieldimensionalen Unterschiedlichkeit der einbezogenen Studien kaum zu überwinden und von Leser*innen, denen das entsprechende methodische Hintergrundwissen fehlt, auch kaum zu durchdringen (zu Problemen bei der Anwendung von Metaanalysen im Kontext evidenzorientierter Unterrichtspraxis vgl. weiterführend Renkl 2022).

Während die Lehrerin nun in einzelnen Interventionsstudien nach konkreteren Umsetzungsbeispielen sucht, bemerkt sie, dass sie in vielen Fällen keinen Zugriff auf die Publikationen hat. Jedoch sollten sich wichtige Hauptergebnisse in den Abstracts finden. Passend erscheint ihr eine Studie von Kirkpatrick et al. (2018), die den Einfluss von iPads im Sprachen- und Mathematikunterricht auf den Lernerfolg (Zeugnisnote) kanadischer 7. Klässler*innen in drei Kohorten untersucht hat. Das Abstract fasst die Befunde wie folgt zusammen:

„The effects were mixed in that the impact was sometimes positive, sometimes negative, sometimes negative and then positive, and sometimes neutral.“ (Kirkpatrick et al. 2018)

Die Lehrerin fragt sich, wie sie mit diesem Befund weiter verfahren soll. Wurde in dieser Studie auch kooperativ gelernt? Ist eine siebte Klasse in Kanada vergleichbar mit einer siebten Klasse in Deutschland? Lässt sich eine lang andauernde Intervention mit einer Unterrichtsreihe vergleichen, die vielleicht zehn bis zwölf Unterrichtsstunden umfasst? Ist die Zeugnisnote ein angemessenes Kriterium, um Aussagen über die Effektivität der Tablets zu treffen? Abgesehen von den bekannten Objektivitäts-, Reliabilitäts- und Validitätsproblemen von Zensuren allgemein, dürften Zeugnisnoten von mannigfaltigen Faktoren beeinflusst sein. Und überhaupt: Wenn sich die einzelnen Befunde in einer Studie in unterschiedlichen Kohorten so drastisch unterscheiden, kann dann eine Vorhersage für den Erfolg einer Implementation in ihrer eigenen Klasse getroffen werden? Darüber hinaus wird nicht klar, inwieweit sich die Wirksamkeit von Tablets individuell von Schüler*in zu Schüler*in unterscheidet. Auch wie es sich mit der Wirksamkeit bei Schüler*innen aus ihrer eigenen Klasse verhält, ist ungewiss.

Diese letztere Frage betrifft die Generalisierbarkeit der Studienergebnisse bzw. die externe Validität, welche bei der Verwendung eines experimentellen Untersu-

chungsdesigns und durch die damit verbundene Künstlichkeit der Situation – wenn auch zugunsten der internen Validität – verringert wird (Stark und Mandl 2007). Die Beantwortung der Frage nach der Generalisierbarkeit der Befunde macht mindestens eine Analyse von lernrelevanten Merkmalen der Stichprobe in Verbindung mit einer Reihe theoretischer Überlegungen sowie eine Analyse weiterer Befunde dieser Studie (z. B. potenzielle Moderatoren) und auch weitere Studien notwendig.

Die Referendarin beschließt, das konsultierte Abstract nicht in ihre Entscheidungsfindung einfließen zu lassen und ihre Recherche zu beenden. Daraufhin erhält sie eine E-Mail von ihrem befreundeten Referendarkollegen, in deren Anhang sie eine Fallstudie von Baumann (2013) zu der Einführung von Tablets bei 25 Siebtklässler*innen an der Integrierten Gesamtschule Lehrte findet. Nach einer Auflistung von Vor- und Nachteilen sowie einer konkreten Beschreibung der Innovation zieht Baumann (2013, S. 3) folgendes Zwischenfazit:

„In vielen Lehrplänen sind Teamfähigkeit und kooperative Kompetenzen als Ziele aufgeführt. Beiden Zielen wird durch das Tablet-Projekt Rechnung getragen, denn die gegenseitige Unterstützungsbereitschaft der Schüler ist stark gestiegen. [...] Dieser Aspekt wirkt sich in besonderem Maße auf die Zusammenarbeit in der Gruppe aus, da wichtige Grundlagen zur Gruppenarbeit gefordert und gefördert werden. So entsteht bspw. eine positive Abhängigkeit und es findet eine unterstützende Interaktion statt. Das kooperative Lernen erlangt durch den Tablet-Einsatz ebenfalls einen neuen Aspekt, denn gerade das kollaborative Arbeiten [...] wird gefördert. [...]“

Die Fallstudie zeigt zumindest ein Beispiel dafür, dass auch im Tablet-gestützten Unterricht kooperativ gelernt werden kann.

In der oben genannten Evidenzhierarchie nehmen Fallstudien einen der unteren Plätze ein (Bromme et al. 2014). Aus einer forschungsmethodischen Perspektive ist der wissenschaftliche Erkenntnisgewinn dieser Art von deskriptiven Studien mehr als fraglich (Rost 2007), zumal massive Probleme der internen Validität gegeben und bei diesem methodischen Vorgehen auch nicht zu überwinden sind. Dennoch bieten Fallstudien insbesondere in Ermangelung von Evidenzen, die in der Evidenzhierarchie höher platziert sind, erste Anhaltspunkte, auf denen in Forschung und Praxis weiter aufgebaut werden kann, wenngleich sie für beide Felder mehr Fragen aufwerfen als beantworten dürfen. Aber in einem neuen Forschungsgebiet ist auch diese eher heuristische Funktion positiv zu bewerten.

Die negativen Erfahrungen von ihrem Kollegen, so überlegt die Referendarin, könnten auch von vielen anderen Faktoren, zum Beispiel dessen Unterrichtsdidaktik und -methodik, den individuellen Eigenschaften der Schüler*innen oder dem Lernklima in der Klasse abhängen und müssen nicht oder nicht primär durch die Tablet-Nutzung bedingt sein. Letztlich ist es an ihr gelegen, die Tablets in ihren Unterricht didaktisch und methodisch so einzubetten, dass ihre Schüler*innen von deren Einsatz profitieren werden. Vielleicht hat ihr befreundeter Referendarkollege recht und sie muss einfach selbst *ausprobieren, was funktioniert und was nicht*.

2.2 Problematik einer evidenzbasierten Argumentation und Entscheidungsfindung

Auch wenn es zweifellos wünschenswert wäre, dass sich Lehrkräfte bei der Bewältigung von unterrichtsbezogenen Fragestellungen an (bildungs-)wissenschaftlicher Evidenz orientieren, stellt ein solches Vorgehen eher den idealen Endpunkt auf einem Kontinuum oder gar einem mehrdimensionalen Modell *Evidenzorientierung* dar, der in der Praxis aus verschiedenen Gründen häufig nicht erreicht werden kann. In dem fiktiven Beispiel sollte deutlich geworden sein, dass die Orientierung an (bildungs-)wissenschaftlicher Evidenz selbst für eine motivierte Referendarin, die auf die Anwendung von Evidenzen durch ihr Studium vorbereitet wurde, erhebliche Probleme mit sich bringt. Eine Problemlösung bzw. die Beantwortung von Anwendungsfragen wird letztlich immer nur einen vorläufigen Charakter haben und immer auch mit Unsicherheiten behaftet sein.

Das Problem beginnt bereits mit der Auswahl geeigneter Studien. Dabei stellt das junge Alter der mit dem Bildungsbereich assoziierten Forschungsdomänen, wodurch Ansprüche an Quantität und Qualität der Evidenz sicherlich nicht zu hoch ausfallen dürften (Hartmann et al. 2016; Hedges 2018), nicht die einzige Hürde dar. Zu der Verfügbarkeitsproblematik tritt die oben beschriebene Passungs- und Interpretationsproblematik (Stark 2017). Selbst wenn es, wie in unserem fiktiven Beispiel, gelingt, anwendungsorientierte Studien zu finden, die zudem höheren methodischen Ansprüchen genügen, wird es immer multiple und schwer aufzulösende Passungsprobleme geben, da Faktoren wie Untersuchungskontext und -inhalte, Merkmale der Proband*innen sowie spezifische Implementationsanforderungen in der Regel nur teilweise mit der konkreten Anwendungssituation korrespondieren (Joyce und Cartwright 2020; Stark 2017). Inwieweit kann ein Befund über eine bestimmte Maßnahme (wie kooperatives Lernen mit Tablets), der (1) an einer bestimmten Stichprobe einer Altersgruppe oder Schulform, mit (2) bestimmter Stichprobenzusammensetzung, die sich zudem (3) zwischen den Experimentalgruppen bezüglich potenzieller, nicht kontrollierter oder kontrollierbarer Wirkfaktoren unterscheidet, sowie an (4) bestimmten Inhalten mit (5) bestimmten Zielen (unmittelbarer Lernerfolg, Zeugnisnote) gewonnen wurde, in eine andere Klassenstufe mit anderer Zusammensetzung (hinsichtlich bestimmter Heterogenitätsmerkmale) sowie in einen anderen Kontext mit anderen strukturähnlichen oder sogar strukturunähnlichen Inhalten und Zielen transferiert werden (Prenzel 2020)? Ebenso selten wie das Erfahrungswissen von Kolleg*innen auf identischen Situationen oder Kontexten (z. B. Schulform, Fachdisziplin, Jahrgangsstufe oder Klasse bzw. bestimmte Schüler*innen) beruhen wird, werden Lehrkräfte empirische Studien finden können, die in Bezug auf wesentliche Kontextmerkmale mit der eigenen Situation korrespondieren. Die Frage zu beantworten, was hierbei wesentliche Kontextmerkmale sind und was nicht, ist ebenfalls nicht trivial, zumal es hierauf keine einfach anzuwendende Standardantwort gibt. Hier ist demnach Transferkompetenz im engeren Sinne gefragt, die, selbst wenn es um einfache Problemstellungen geht, vor allem Personen mit niedrigem Vorwissen vor erhebliche Herausforderungen stellt (Detterman und Sternberg 1993; Galoyan und Betts 2021).

Metaanalysen und randomisierte kontrollierte Studien, der *Gold-Standard* evidenzbasierter oder -orientierter Argumentation, machen zwar prinzipiell, sofern sie methodisch einwandfrei sind, Kausalzuschreibungen im Sinne von „es funktioniert“ möglich, bieten aber letztlich keine verlässliche Grundlage, um Schlussfolgerungen für den individuellen Anwendungskontext im Sinne von „hier und für dieses Ziel wird es auch funktionieren“ zu ziehen (Joyce und Cartwright 2020). Ohnehin greifen einfache *wenn-dann*-Folgerungen bei der Planung von Unterricht zu kurz. So können neben der Stichprobenzusammensetzung auch bestimmte Spezifika der Intervention selbst (zum Beispiel Lerninhalt, konkrete didaktisch-methodische Umsetzung, verwendete Materialien) den Erfolg einer Intervention maßgeblich beeinflussen – Informationen, über welche Anwender*innen in vielen Studien nur wenig Präzises erfahren, unabhängig davon, ob sie sich an Metaanalysen oder Einzelstudien orientieren.

Hinzu kommt, dass die multiplen (*multikriterialen*) Ziele von Unterricht im Feld (Hartmann et al. 2016) in der Regel nicht von Metaanalysen und schon gar nicht von Einzelstudien konkret abgebildet werden. In dem systematischen Review von Connolly und Kolleginnen (2018), das 1017 randomisierte Kontrollstudien aus den Jahren 1980–2016 im Bildungsbereich in den Blick nahm, zeigte sich darüber hinaus, dass in fast zwei Drittel der einbezogenen randomisierten Kontrollstudien keine Prozessevaluationen (d. h. qualitative Analysen, welche die Erkenntnisse aus den quantifizierten Ergebnissen ergänzen) vorgenommen wurden, und die Analysen in etwa der Hälfte der Studien nicht über die Gesamteffekte der jeweiligen Interventionen auf die Stichprobe als Ganzes hinausgingen. Den Grad der Passung einer Studie für ihre Fragestellung und ihre Klasse müssen Anwender*innen letztlich selbst beurteilen und Schlussfolgerungen daraus ziehen, welche nicht selten anspruchsvolle Interpretationsleistungen notwendig machen und welche nie ganz frei von Spekulationen sind. Diese Art der Subjektivität ist auf der Basis empirischer Studien nicht zu eliminieren, sodass die Interpretation in der Regel mit einem hohen Grad an Unsicherheit behaftet ist. Neben wissenschaftlicher Evidenz müssen Lehrkräfte ihr Erfahrungswissen und ihr Wissen über diverse Kontextfaktoren bei ihrer Entscheidungsfindung professionell integrieren (Hartmann et al. 2021), was ebenfalls und in vielfacher Hinsicht anspruchsvoll ist – gerade, wenn es an klaren Richtlinien fehlt.

Neben die Passungsproblematik tritt die Problematik mangelnder Kohärenz, welche nicht nur widersprüchliche Befunde zwischen verschiedenen Quellen betrifft, sondern sich auch – wie das obige Beispiel (vgl. Kirkpatrick et al. 2018) zeigt – bereits zwischen einzelnen Befunden innerhalb einer Studie äußern kann. Angesichts der Vielzahl an Informationsquellen, die potenziell zur Lösung spezifischer Fragestellungen herangezogen werden können, ist es nicht unwahrscheinlich, dass aus verschiedenen Quellen beinahe ebenso viele verschiedene Handlungsempfehlungen abgeleitet werden können. Widersprüchliche und/oder uneindeutige Befunde, die im Bildungsbereich keine Ausnahme darstellen und vielleicht sogar ab einem bestimmten Komplexitätsgrad der Fragestellung die Regel sein dürften, sind nicht nur für Praktiker*innen schwer zu überwindende Barrieren – selbst Wissenschaftler*innen tun sich mit dieser Problematik oft schwer (Stark 2017).

3 Mögliche Ansatzpunkte zur Auflösung der Problematiken

3.1 Erweiterung des Evidenzbegriffs auf *übergeordnete Wissensbestände*

Vor dem Hintergrund der oben beschriebenen komplexen und zum Teil kaum auflösbaren Problematiken ist es nicht nur angezeigt, im Sinne eines *moderaten Evidenzverständnisses*, Evidenz als Ressource zu verstehen, auf deren Basis erfahrungsbasierte Wissensbestände reflektiert, ergänzt oder korrigiert werden können (Bauer et al. 2015; Neuweg 2007; Stark 2017), sondern darüber hinaus den ohnehin vagen Evidenzbegriff auf sämtliche Komponenten wissenschaftlicher Wissensbestände zu erweitern, das bedeutet stärker auf *übergeordnete Wissensbestände* als auf einzelne empirische Studien und deren Ergebnisse zu fokussieren. Der Begriff *übergeordnete Wissensbestände* bezieht sich insbesondere auf wissenschaftliche Theorien und Konzepte, aber auch auf einschlägige wissenschaftliche Ansätze, Prinzipien und Modelle, deren Anwendung vor allem aufgrund des höheren Abstraktionsgrads deutlich weniger Passungs- und Interpretationsprobleme und damit auch weniger Barrieren mit sich bringen dürfte als eine Orientierung an empirischen Befunden einzelner Interventionsstudien oder gar Metaanalysen (Renkl 2022). Solche Wissensbestände können ebenfalls eine sinnvolle Ergänzung oder ein Korrektiv zu subjektiven Theorien von Praktiker*innen über die Wirksamkeit bestimmter Handlungsweisen oder Maßnahmen im Unterricht bieten (Groeben et al. 1988; Stark 2017).

Gewiss bilden Theorie und Empirie ebenso wie die Orientierung an wissenschaftlichen Theorien und die Orientierung an Metaanalysen bzw. randomisierten Kontrollstudien keine exklusiven Kategorien, zumal empirische Forschung stets nach theoretischer Einbettung verlangt und die Theorieentwicklung zumindest in empirisch arbeitenden Subdisziplinen der Bildungswissenschaften grundsätzlich auf empirische Forschung angewiesen ist (vgl. auch Connolly et al. 2018). Dennoch dürfte es in der Unterrichtspraxis im seltensten Fall darauf ankommen, *welche* Maßnahme implementiert wird, sondern vielmehr darauf, *wie* genau sie implementiert wird, und bei dieser Frage sind die angeführten übergeordneten Wissensbestände oft hilfreicher als bildungswissenschaftliche Evidenz im engeren Sinne. Organisationsformen, Methoden und Sozialformen geben zwar als *Oberflächenmerkmale* (auch *Sichtstrukturen*; zum Begriffsdiskurs vgl. Decristan et al. 2020) den Rahmen für Lernprozesse und sind damit bedeutsame Faktoren für die Gestaltung eines lernwirksamen Unterrichts, jedoch sollte neben diesen „Merkmale[n] der Inszenierung“ seitens der Lehrkraft (Pauli 2020, S. 133) der Qualität der Lernaktivitäten und -interaktionen seitens der Lernenden eine ebenso hohe, wenn nicht sogar höhere Relevanz beigegeben werden (Decristan et al. 2020; Pauli 2020). So können innerhalb des gleichen Oberflächenmerkmals sehr unterschiedliche Lernprozesse stattfinden. Anstelle der Frage, ob eine Maßnahme *per se* lernwirksam ist, oder lernwirksamer als eine andere, sollte somit vielmehr die Frage nach dem *Wie*, *Wann* und *Warum* gestellt werden (Decristan et al. 2020). Zwar können die auf Lern- und Verstehensprozesse bezogenen *Tiefenmerkmale* (auch *Tiefenstrukturen*; zum Begriffsdiskurs vgl. Decristan et al. 2020) – wenn überhaupt – nur indirekt beobachtet werden, sie erscheinen jedoch für die Effektivität von Unterricht umso bedeutsamer (Decristan et al. 2020; Kunter und

Trautwein 2013; Seidel und Shavelson 2007)². Insofern bestünde das Hauptproblem unserer Referendarin nicht darin, dass sie bei ihrer Recherche mit immer weiteren Fallstricken konfrontiert wird, sondern dass sie von einer zu engen bzw. nicht ziel-führenden Ausgangsfrage ausgegangen ist. Es sollte hier demnach weniger um ein evidenzorientiertes *Ob*, sondern vielmehr um ein evidenzorientiertes *Wie* gehen. Mit einer von der Evidenzhierarchie losgelösten Erweiterung des moderaten Evidenzbegriffs auf *übergeordnete Wissensbestände* trägt man der Tatsache Rechnung, dass Handeln in der pädagogischen Praxis in seiner ganzen Komplexität, selbst wenn man nur kognitive Dimensionen oder – noch reduktionistischer – Wissensbestände in den Blick nimmt, nur in Ausnahmefällen gänzlich auf empirisch abgesicherten Erkenntnissen basieren kann (Neuweg 2005).

Unter den Begriff *übergeordnete Wissensbestände* lassen sich einerseits solche Wissensbestände fassen, welche einen eher allgemeinen Orientierungsrahmen für den täglichen Unterricht bieten. So wurden in den letzten Jahren von der quantitativ-empirischen Unterrichtsforschung – auch im deutschsprachigen Raum – verschiedene Modelle formuliert und empirisch überprüft, welche Lehrkräften wichtige Orientierungspunkte für die Konzeption und Durchführung eines qualitativ hochwertigen und lernwirksamen Unterrichts bieten können. Als erste Orientierungshilfe dienen hier Angebot-Nutzungs-Modelle (zum Beispiel Helmke 2012), welche Unterricht als ein Angebot für Lernende verstehen, dessen Nutzung und Ertrag von verschiedenen Faktoren wie Vorwissen oder Motivation abhängt. Diese Modelle geben Lehrkräften erste Anhaltspunkte, um lernerfolgsrelevante Variablen und *Stellschrauben* von Unterricht zu reflektieren (aktuelle Übersicht bei Vieluf et al. 2020). Mit Blick auf die Frage, wie ein solches Angebot lernwirksam gestaltet werden kann, hat die Unterrichtsforschung außerdem verschiedene Merkmale qualitativ hochwertigen Unterrichts ausgearbeitet, aus denen Schlüsse für unterrichtliches Handeln gezogen werden können (z. B. Helmke 2012; Meyer 2004). Insbesondere hat sich im deutschsprachigen Raum das Modell der drei Basisdimensionen von Unterrichtsqualität etabliert, welches bereits 1995 im Rahmen der TIMSS-Videostudie entwickelt wurde (Klieme et al. 2001; vgl. weiterführend auch Praetorius et al. 2020a, b). In diesem Modell werden verschiedene Tiefenstrukturmerkmale den folgenden drei Hauptdimensionen zugeordnet: (1) Effektive Klassenführung, (2) kognitive Aktivierung und (3) konstruktive Unterstützung. Die theoretischen Bezüge des Modells und seine Vereinbarkeit mit Angebots-Nutzungs-Modellen bieten einen geeigneten Ansatz zur Konzeptualisierung von Unterrichtsqualität (Praetorius et al. 2020a, b). Da allerdings verschiedene empirische Untersuchungen, die sich in den vergangenen Jahren den Effekten der drei Basisdimensionen gewidmet haben – insbesondere in den MINT-Domänen und im deutschsprachigen Raum – darauf hindeuten, dass die drei Basisdimensionen keinen hinreichenden Ansatz zur Beschreibung von Unterrichtsqualität darstellen (Review bei Praetorius et al. 2018), beschäftigt sich die Unterrichtsforschung aktuell eingehend mit der Frage, inwieweit die drei Basisdimensionen um

² Bis dato mangelt es noch an einer empirischen Fundierung der theoretischen Annahmen bezüglich der Relevanz von Oberflächen- vs. Tiefenmerkmalen und deren Zusammenhängen, jedoch würde eine nähere Beleuchtung des Diskurses an dieser Stelle zu weit führen. Weiterführend soll daher auf die Beiträge von Decristan et al. (2020), Hess und Lipowsky (2020) sowie Pauli (2020) verwiesen werden.

weitere generische (d.h. allgemeine oder fächerübergreifende) und fachspezifische Aspekte erweitert werden sollten (vgl. Syntheseframework von Praetorius und Gräsel 2021). Zumindest könnte unsere fiktive Referendarin anhand der Qualitätskriterien qualitativ hochwertigen Unterrichts beispielsweise folgern, dass sie beim kooperativen Lernen mit Tablets mithilfe einer professionellen Klassenführung, einer kognitiv aktivierenden Gestaltung der (digitalen) Materialien und Aufgabenstellungen sowie einer Atmosphäre der gegenseitigen Wertschätzung für eine motivierte Beteiligung aller Sorge tragen muss.

Als zweifellose hilfreiche, allgemeine Leitlinien für die Unterrichtspraxis erscheinen darüber hinaus prominente Theorien und Konzepte, die aus dem pädagogisch-psychologischen Diskurs der letzten Jahrzehnte kaum wegzudenken sind (Renkl 2022), zum Beispiel zu Themen wie Cognitive Load (Sweller et al. 2011), Motivation (Deci und Ryan 1993; Schiefele et al. 2018) oder lehr-/lernbezogenen Emotionen (Pekrun 2000). Solche Theorien und Konzepte könnten der Referendarin beispielsweise dabei helfen, ihre Lehrkonzeption insgesamt zu optimieren, indem sie motivations- und emotionstheoretische Perspektiven bei der Konzeption von Materialien und Aufgabenstellungen einbezieht, die wahrgenommene Autonomie, Kompetenz und soziale Eingebundenheit der Lernenden unterstützt, deren Kontrolle erhöht, und günstige Wertzuschreibungen in Bezug auf die zu lernenden Inhalte oder zu erwerbenden Kompetenzen induziert.

Neben solchen Wissensbeständen, die eher einen allgemeinen Orientierungsrahmen oder Leitlinien geben, lassen sich dem Begriff der *übergeordneten Wissensbestände* auch für spezifischere Fragen relevante Wissensbestände zuordnen, aus denen sich konkrete Handlungsempfehlungen für den eigenen Unterricht ableiten lassen, so zum Beispiel die sogenannten, auf empirischen Befunden gründenden *fünf Basiselemente kooperativen Lernens* (Johnson und Johnson 2009) oder Kounins (2006) *Techniken der Klassenführung*. Diese können der Lehrerin in unserem fiktiven Beispiel entscheidende Hinweise bieten, um negativen Gruppeneffekten wie motivationalen Problemen bei einer kollektiven Bearbeitung von Aufgaben (Trittbrettfahreneffekt; Neber 2018) sowie Störungen im Unterricht vorzubeugen. Zwar kann eine Orientierung an übergeordneten Wissensbeständen der Referendarin letztlich nicht die Entscheidung darüber abnehmen, welcher konkrete Lehr-Lern-Ansatz für ihre Lehrkonzeption zu favorisieren wäre. Sie könnten die Referendarin jedoch – nach ihrer Entscheidungsfindung – bei der Konzeption ihres Lehrvorhabens, der Antizipation möglicher Probleme und vor allem der konkreten Implementation des Lehrvorhabens unterstützen.

Alles in allem stellt die bildungsbezogene Forschung, wenn auch nicht in Form von direkt anwendbaren empirischen *was funktioniert-* Befunden, durchaus belastbare Quellen bereit. Anzumerken ist jedoch, dass solche übergeordneten Wissensbestände ebenso wie empirische Befunde in der Regel weitergehender Reflexionen bedürfen, bevor sie erfolgreich in die Unterrichtspraxis implementiert werden können. Letztlich müssen Lehrkräfte unter Berücksichtigung der Interdependenz von Angebot und Nutzung sowie unter Integration ihres reichhaltigen Erfahrungs- und Kontextwissens eine Balance zwischen den Basisdimensionen der Unterrichtsqualität und zwischen den von ihnen angestrebten multiplen Zielen finden (Hartmann et al. 2016; 2021). So müsste unsere Referendarin bei ihrer Konzeption individu-

elle Stärken und Schwächen der Schüler*innen, lern- und kooperationsrelevante Eingangsvoraussetzungen sowie Gruppendynamiken in der Klasse berücksichtigen, welche sie wiederum durch eine systematische Reflexion der einzelnen Unterrichtseinheiten analysieren könnte. Aus einer evidenzorientierten Analyse dessen, inwieweit einzelne Lernende von den Lernarrangements profitiert haben, könnte sie wiederum am konkreten Fall gezielt Optimierungsmaßnahmen ableiten und evaluieren.

In Ergänzung zu der beschriebenen Orientierung an übergeordneten Wissensbeständen in Form von bildungswissenschaftlichen Theorien, Konzepten, Ansätzen, Prinzipien oder Modellen postulieren wir die heuristische Potenz von Prinzipien und Konzepten aus der Statistik, im Sinne einer weiteren Systematisierung pädagogischer Reflexion durch Integration einer zusätzlichen inter-individuellen Analyseperspektive. Statistische Prinzipien und Konzepte können jenseits aller inhaltlichen Bezüge sowie jenseits ihres Nutzens für den Nachvollzug und der Einordnung empirischer Befunde praktische Werkzeuge bereitstellen, die Praktiker*innen dabei helfen, eine zusätzliche inter-individuelle Analyseperspektive einzunehmen, die insbesondere bei der Entwicklung von umfassenderen Erklärungsmodellen in Bezug auf unterrichtsbezogene Sachverhalte und deren Genese, aber auch bei der Ableitung konkreter Konsequenzen für das Unterrichtshandeln in einer bestimmten Situation hilfreich sein kann.

So könnte die Referendarin bei ihrer Analyse die Erfahrung gemacht haben, dass sich Lernprodukte von solchen Schüler*innen, deren Gruppen sich durch bestimmte Interaktionsprozesse auszeichnen, qualitativ von Produkten derjenigen unterschieden haben, deren Gruppen diese Interaktionsprozesse nicht oder weniger stark ausgeprägt erkennen ließen. Davon ausgehend könnte die Lehrerin einen korrelativen Zusammenhang zwischen Interaktionsprozessen in Gruppen und individuellem Lernerfolg formulieren, den sie weiter analysieren könnte, etwa in Bezug auf seine kausale Relevanz. Sie könnte beispielsweise die Arbeitshypothese formulieren, dass sich bestimmte Interaktionsprozesse positiv auf den individuellen Lernerfolg auswirken, die ihrerseits die Qualität der Interaktionsprozesse beeinflussen. Das heißt, sie würde eine Art Mediationshypothese formulieren. Ohne gezielte empirische Studien kann zwar keine fundierte Entscheidung zugunsten der einen oder der anderen Hypothese getroffen werden. Im vorliegenden Zusammenhang geht es jedoch primär um die *heuristische Potenz* des Mediationskonzepts – und dieses kann sich auch ohne empirisch-statistische Datengrundlage im engeren Sinn und vor allem ohne entsprechendes mathematisches bzw. statistisches Hintergrundwissen entfalten. Durch ein solches an Prinzipien und Konzepten der Statistik orientiertes Vorgehen würde – jenseits einer genuinen Evidenzorientierung – eine wissenschaftlich-analytische, systematische Vorgehensweise unterstützt, die versucht, diagnostizierte Merkmale im Sinne von Variablen und deren Ausprägungen miteinander in Beziehung zu setzen.

Eine evidenzorientierte Praxis also, die sich mindestens ebenso stark oder sogar stärker auf bildungswissenschaftliche Theorien, Konzepte, Ansätze, Prinzipien oder Modelle als auf Befunde einzelner empirischer Studien beruft, und die den Nutzen statistischer Kenntnisse für Lehrkräfte vor allem in ihrer heuristischen Potenz bei der Analyse von Lehr-Lern-Prozessen sieht, ist sicher nicht im engeren Sinne *evidenzbasiert*, sondern vielmehr *wissensbestandsorientiert*. Sie folgt jedoch eindeutig Prinzipien wissenschaftlichen Arbeitens vor dem Hintergrund einer falsifikationistischen

Grundhaltung. Die exemplarisch angeführten übergeordneten Wissensbestände sind allesamt als Bestandteil des aktuellen bildungswissenschaftlichen Diskurses empirisch mehr oder weniger stark *verankert*, die Reflexion und Integration der herangezogenen Erfahrungswissensbestände hat zudem bewusst kritisch und systematisch zu erfolgen, und das gesamte Vorgehen bei der Analyse, Planung bis zur späteren Handlung ist einer permanenten *Evaluation* zu unterziehen und dadurch prinzipiell immer der Falsifikation ausgesetzt. Auch wenn mit einer Öffnung der Evidenzorientierung im Sinne einer *Wissensbestandsorientierung* die externalen und internalen Barrieren, mit denen ein evidenzorientiertes Vorgehen stets konfrontiert ist, sicher nicht vollständig zu überwinden sind, sollten diese Barrieren deutlich niedriger ausfallen, als wenn Praktiker*innen angehalten werden, sich an empirischen Originalarbeiten, Kurzzusammenfassungen empirischer Studien oder gar Metaanalysen zu orientieren. Solche übergeordneten Wissensbestände können unterrichtliches Handeln zwar nicht direkt instruieren, besitzen allerdings „handlungsvorbereitenden, wahrnehmungs- und problemdefinitionsleitenden und handlungsrechtfertigenden Charakter“ (Neuweg 2014, S. 599), wodurch sie eine geeignete Basis für professionelles Handeln in der Praxis bieten (Hasselhorn et al. 2016).

3.2 Abbau von Barrieren durch einen Ausbau von Clearing Houses und weitere Forschung

Durch die obigen Ausführungen sollte deutlich geworden sein, dass die vielleicht relevanteste Barriere evidenzorientierter Praxis nicht auf Seiten der Praktiker*innen, sondern in der Konzeption von Evidenzbasierung und der Beschaffenheit von Evidenzen selbst liegt. Wichtige Voraussetzungen für die Anwendung wissenschaftlicher Wissensbestände in der Schul- und Unterrichtspraxis liegen in ihrer Passung zum situativen Kontext, in ihrem Aussage- und Erklärungsgehalt, und letztlich in der Ableitbarkeit von Handlungsoptionen (Hasselhorn et al. 2016). In Anlehnung an Hedges (2018) sind hiermit drei Herausforderungen verbunden: der Umgang mit der Replikationskrise, die Erhöhung der Verallgemeinerbarkeit der bildungswissenschaftlichen Forschung sowie die Anpassung von Forschungsdesigns an die zunehmende Komplexität von Interventionen und entsprechenden Wirkmechanismen. Joyce und Cartwright (2020) plädieren für eine Forschung, die sich um eine individuelle, d. h. lokale Anwendbarkeit von Evidenz unter den jeweiligen Kontextbedingungen bemüht. Eine flächendeckende Etablierung von Unterrichts- (und auch Schul-)entwicklungsprozessen, im Rahmen derer auf Basis intern erhobener Daten innovative Wege und Möglichkeiten von Unterricht entwickelt werden (Altrichter et al. 2016), würde dieser Forderung Rechnung tragen.

Auch wenn diese zweifellos relevante Aspekte für zukünftige Forschung darstellen, stellt sich immer noch die Frage, wie bisherige Befunde gewinnbringend in der Praxis angewendet werden können. Gräsel (2019) betont hier die Relevanz eines adäquaten *Transfers* (des aktuell vorhandenen) wissenschaftlichen Wissens in die Praxis durch eine konkrete Anpassung von Forschungserkenntnissen an die Rahmenbedingungen des für eine Implementation vorgesehenen Kontextes. Das *Clearing House Unterricht* der TU München bildet sicherlich einen ersten Meilenstein hinsichtlich der anwendungsorientierten Aufbereitung von Evidenzen. Ihm und ähn-

lichen Portalen kommt ein erhebliches Potenzial zu, das in dieser Form einzigartig ist. Da sich das *Clearing House Unterricht* zunächst an lehrkräftebildende Personen als Multiplikator*innen richtet (Diery et al. 2020a), ist es jedoch vor allem für die systematische Weiterbildung von Personen gewinnbringend, die bereits über umfassende forschungsmethodische Kenntnisse und möglichst über eigene Erfahrung mit empirischer Forschung verfügen. Wie am Beispiel der fiktiven Referendarin dargestellt wurde, ist der Nutzen für Anwendungssituationen in der Unterrichtspraxis bisher eher eingeschränkt.

Damit jedoch aufbereitete Evidenzen nicht nur von lehrkräftebildenden Personen, sondern auch von Praktiker*innen als für sie und ihre Schulpraxis bedeutsam wahrgenommen und mit Blick auf unterrichtspraktische Entscheidungen adäquat interpretiert, verarbeitet und angewandt werden können, erscheint ein gezielter Ausbau von breit angelegten Clearing Houses notwendig. Dies betrifft eine Ausweitung der berücksichtigten Domänen (über die MINT-Fächer und den Sekundarbereich hinaus) sowie eine Erweiterung der aufbereiteten Wissensbestände selbst. Das *Clearing House Unterricht* dürfte mit in 29 Kurzreviews verarbeiteten Metaanalysen (Stand Mai 2022) bislang nur einen kleinen Ausschnitt an relevanten Fragestellungen abdecken, die im Unterrichtsalltag auftreten können. Hier ist es angezeigt, auch Wissensbestände, die nicht dem *Gold-Standard* evidenzbasierter oder -orientierter Argumentation entsprechen, strukturiert und themenspezifisch zusammenzufassen. Des Weiteren könnten wie oben dargelegt übergeordnete Wissensbestände, also Theorien, Konzepte, Ansätze, Prinzipien oder Modelle zu einem bestimmten Gebiet (wie zum Beispiel kooperativem Lernen) zusammengetragen und mit konkreten Anwendungsbeispielen für die Unterrichtspraxis unterfüttert werden. Besonders vielversprechend erscheinen hier Fallszenarien und Unterrichtsvideos (Piwowar et al. 2018; Seidel et al. 2021). In diesem Zusammenhang wäre eine Verknüpfung des Clearing House mit dem jüngst entstandenen *Meta-Videoportal für die Lehrkräftebildung* (<https://unterrichtsvideos.net/metaportal/>) sicherlich vielversprechend. Solche konkreten Anwendungsbeispiele würden auch eine hilfreiche Ergänzung des *Fazits für die Unterrichtspraxis* darstellen, welches vom *Clearing House Unterricht* bislang in jedem Kurzreview gezogen wird.

Weiterhin wäre es notwendig, nicht nur Anwendungsmöglichkeiten, sondern auch konkrete Grenzen der Anwendbarkeit aufzuzeigen. Hier könnten die als Ergänzung zu der in den Kurzreviews vorgenommenen Bewertung bereitgestellten Rating-Sheets dahingehend erweitert werden, dass offengelegt wird, worin konkret die Stärken und Schwächen der aufbereiteten Studie liegen. Hieraus müsste deutlich werden, was Praktiker*innen aus der Bewertung schlussfolgern können – und was nicht. Dies betrifft vor allem das Gesamturteil, das am Ende eines jeden Rating-Sheets präsentiert wird. Für lehrkräftebildende Personen können die Prozentangaben für eine grobe Einschätzung der Qualität einer Metaanalyse sicherlich nützlich sein, als Entscheidungshilfe für Praktiker*innen erscheint der Mehrwert der aktuellen Darstellungsweise ohne konkrete Erläuterungen und Implikationen für die Praxis jedoch eher gering, könnte mitunter sogar zu falschen Schlüssen führen.

Auch wenn die vielleicht größte Barriere evidenzorientierter Schul- und Unterrichtspraxis in den Eigenschaften von Evidenz selbst liegt, so bedarf es – um das Angebot von Clearing Houses auch für Lehrkräfte anwendungsorientierter zu ge-

stalten – weiterer Forschung, die sich nicht nur mit möglichen Barrieren auf Seiten lehrkräftebildender Personen (Diery et al. 2020b), sondern auch mit möglichen Barrieren auf Seiten von Praktiker*innen auseinandersetzt. Ein Forschungsdesiderat liegt dabei insbesondere in der Frage, wie Praktiker*innen mit (1) einer geringen Passung der Evidenz zu ihrer Fragestellung bzw. Problematik und dem vorliegenden Kontext sowie (2) einer geringen Kohärenz einzelner Befunde zwischen verschiedenen Studien oder sogar einer geringen Kohärenz einzelner Befunde zu den eigenen Erfahrungsbeständen umgehen. Ein tieferes Verständnis anspruchsvoller Integrationsprozesse wäre nicht nur mit Blick auf die adressatengerechte Aufbereitung von Evidenzen in Kurzreviews in hohem Maße gewinnbringend, sondern eröffnete darüber hinaus die Ableitung konkreter Förderansätze für die Lehrkräfteausbildung und -weiterbildung.

Weitere relevante Voraussetzungen für eine evidenzorientierte Praxis im Sinne einer *Wissensbestandsorientierung* sind in günstigen Einstellungen und Überzeugungen von Praktiker*innen gegenüber wissenschaftlichen Wissensbeständen zu sehen (Merk et al. 2017; Rochnia und Trempler 2019; Schellenbach-Zell und Rochnia 2020). Sicherlich würde bereits der oben beschriebene Ausbau von Clearing Houses der wahrgenommenen Nützlichkeit wissenschaftlicher Wissensbestände zugutekommen, jedoch sollte zudem der Frage nachgegangen werden, durch welche konkreten Maßnahmen ein positiver Nützlichkeitswert (z. B. durch bestimmte Wertinduktionsverfahren; Johnson und Sinatra 2013; Stark et al. 2022, in Druck) sowie günstige epistemologische Überzeugungen (z. B. durch bestimmte Verfahren zur Induktion epistemischen Wandels; Kerwer und Rosman 2020) bei angehenden und berufstätigen Lehrkräften weiter gefördert werden können.

Alles in allem darf der Transfer von Forschung in die Praxis nicht nur als ein individueller Prozess einzelner Lehrkräfte verstanden werden, welche von Clearing Houses aufbereitete Quellen in ihrer Praxis anwenden (Gräsel 2019). Es bedarf immer noch eines gemeinsamen Diskurses, in dem Potenziale und Barrieren evidenzorientierter Praxis reflektiert, diskutiert und erarbeitet werden, und damit einer flächendeckenden Etablierung von Wissenschafts-Praxis-Kooperationen, wie sie beispielsweise Hartmann und Kolleg*innen (2016) skizzieren. Ein gemeinsamer Diskurs würde darüber hinaus, im Sinne von Joyce und Cartwright (2020), zu einer Forschung beitragen, die nicht nur zeigt, *was funktioniert*, sondern *was in einem bestimmten Kontext mit bestimmten Zielen wie, wann, und warum funktioniert*.

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>.

Interessenkonflikt T. Wilkes und R. Stark geben an, dass kein Interessenkonflikt besteht.

Literatur

- van Ackeren, I., Heinrich, M., & Thiel, F. (2013). *Evidenzbasierte Steuerung im Bildungssystem? Befunde aus dem BMBF-SteBis-Verbund*. Die deutsche Schule: Beiheft. Münster: Waxmann.
- Altrichter, H., Moosbrugger, R., & Zuber, J. (2016). Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 235–277). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-18942-0_9.
- Asberger, J., Thomm, E., & Bauer, J. (2020). Empirische Arbeit: Zur Erfassung fragwürdiger Überzeugungen zu Bildungsthemen: Entwicklung und erste Überprüfung des Questionable Beliefs in Education-Inventars (QUEBEC). *Psychologie in Erziehung Und Unterricht*, 67, 178–193. <https://doi.org/10.2378/peu2019.art25d>.
- Bauer, J., Prenzel, M., & Renkl, A. (2015). Evidenzbasierte Praxis – im Lehrerberuf?! Einführung in den Thementeil. *Unterrichtswissenschaft*, 43(3), 188–192. <https://doi.org/10.3262/UW1503188>.
- Bauer, J., Berthold, K., Hefter, M.H., Prenzel, M., & Renkl, A. (2017). Wie können Lehrkräfte und ihre Schülerinnen und Schüler lernen, fragile Evidenz zu verstehen und zu nutzen? *Psychologische Rundschau*, 68(3), 188–192. <https://doi.org/10.1026/00333042/a000363>.
- Baumann, S. (2013). Tablet-PCs im Unterricht: Erste Erkenntnisse einer Fallstudie. *Beiträge zum Mathematikunterricht, 2013*, 104–107. <https://doi.org/10.17877/DE290R-1267>.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Beelmann, A. (2014). Möglichkeiten und Grenzen systematischer Evidenzakkumulation durch Forschungssynthesen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaften*, 17(4), 55–78. <https://doi.org/10.1007/s11618-014-0509-2>.
- Biesta, G. (2011). Warum „What works“ nicht funktioniert: Evidenzbasierte pädagogische Praxis und das Demokratiedefizit der Bildungsforschung. In J. Bellmann & T. Müller (Hrsg.), *Wissen, was wirkt: Kritik evidenzbasierter Pädagogik* (S. 95–121). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-93296-5_4.
- Bleck, V., & Lipowsky, F. (2020). Dröge, nutzlos, praxisfern? Wie verändert sich die Bewertung wissenschaftlicher Studieninhalte in Praxisphasen? In I. Ulrich & A. Gröschner (Hrsg.), *Praxissemester im Lehramtsstudium in Deutschland: Wirkungen auf Studierende* (S. 97–127). Wiesbaden: Springer VS.
- Bromme, R., Prenzel, M., & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. *Zeitschrift Für Erziehungswissenschaft*, 17(S4), 3–54. <https://doi.org/10.1007/s11618-014-0514-5>.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2. Aufl.). New York: Routledge.
- Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, 60(3), 276–291. <https://doi.org/10.1080/00131881.2018.1493353>.
- Csanadi, A., Kollar, I., & Fischer, F. (2021). Pre-service teachers' evidence-based reasoning during pedagogical problem-solving: better together? *European Journal of Psychology of Education*, 36(1), 147–168. <https://doi.org/10.1007/s10212-020-00467-4>.
- Deci, E.L., & Ryan, R.M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238. <https://doi.org/10.25656/01:11173>.
- Decristan, J., Hess, M., Holzberger, D., & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale. Eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. In A.-K. Praetorius, J. Grünkorn & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 102–116). Weinheim, Basel: Beltz Juventa.
- Detterman, D.K., & Sternberg, R.J. (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Norwood: Ablex.
- Diery, A., Knogler, M., Mazziotti, C., Schneeweiss, A., Hetmanek, A., Holzberger, D., & Seidel, T. (2020a). Das Clearing House Unterricht. Ein Service für die Lehrer*innenbildung?! *Journal für LehrerInnenbildung*, 2, 42–49. https://doi.org/10.35468/jlb-02-2020_03.

- Diery, A., Vogel, F., Knogler, M., & Seidel, T. (2020b). Evidence-based practice in higher education: teacher educators' attitudes, challenges, and uses. *Frontiers in Education*, 5, 62. <https://doi.org/10.3389/educ.2020.00062>.
- Diery, A., Knogler, M., & Seidel, T. (2021). Supporting evidence-based practice through teacher education: a profile analysis of teacher educators' perceived challenges and possible solutions. *International Journal of Educational Research Open*, 2, 100056. <https://doi.org/10.1016/j.ijedro.2021.100056>.
- Egger, C., & Ophoff, G. J. G. (2020). Die Einschätzung des Nutzens von Forschung als Voraussetzung für die Entwicklung einer forschenden Haltung von Lehramtsstudierenden. In I. Gogolin, B. Hannover & A. Scheunpflug (Hrsg.), *Evidenzbasierung in der Lehrkräftebildung* (S. 73–93). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-22460-8_4.
- Fischer, F. (2021). Some reasons why evidence from educational research is not particularly popular among (pre-service) teachers: a discussion. *Zeitschrift Für Pädagogische Psychologie*, 35(2), 209–214. <https://doi.org/10.1024/1010-0652/a000311>.
- Galoyan, T., & Betts, K. (2021). Integrative transfer of learning model and implications for higher education. *The Journal of Continuing Higher Education*, 69(3), 169–191. <https://doi.org/10.1080/07377363.2020.1847970>.
- Gogolin, I., Hannover, B., & Scheunpflug, A. (2020). Evidenzbasierung als leitendes Prinzip in der Ausbildung von Lehrerinnen und Lehrern – Editorial. In I. Gogolin, B. Hannover & A. Scheunpflug (Hrsg.), *Evidenzbasierung in der Lehrkräftebildung* (S. 1–9). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-22460-8_1.
- Gräsel, C. (2019). Transfer von Forschungsergebnissen in die Praxis. In C. Donie, F. Foerster, M. Obermayr, A. Deckwerth, G. Kammermeyer & G. Lenske (Hrsg.), *Jahrbuch Grundschulforschung. Grundschulpädagogik zwischen Wissenschaft und Transfer* (S. 2–11). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-26231-0_1.
- Groeben, N., Wahl, D., Schlee, J., & Scheele, B. (1988). *Das Forschungsprogramm Subjektive Theorien: eine Einführung in die Psychologie des reflexiven Subjekts*. Tübingen: Francke.
- Gröschner, A., & Seidel, T. (2015). Evidenzbasierte Lehrerbildung. Der Beitrag der Bildungsforschung zur Gestaltung von Bildungspraxis. *Journal Für LehrerInnenbildung*, 4, 9–15.
- 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, 179–199. <https://doi.org/10.1007/s11618-016-0712-4>.
- Hartmann, U., Kindlinger, M., & Trempler, K. (2021). Integrating information from multiple texts relates to pre-service teachers' epistemic products for reflective teaching practice. *Teaching and Teacher Education*, 97(Article 103205), 1–12. <https://doi.org/10.1016/j.tate.2020.103205>.
- Hasselhorn, M., Drewes, S., & Seifried, K. (2016). Wissenschaftliches Selbstverständnis schulpsychologischen Handelns. In K. Seifried, S. Drewes & M. Hasselhorn (Hrsg.), *Handbuch Schulpsychologie: Psychologie für die Schule* (2. Aufl. S. 23–29). Stuttgart: Kohlhammer.
- Hastall, M. R., & Lang, B. (2019). Grundlagen einer evidenzbasierten Gesundheitskommunikation. In C. Rossmann & M. R. Hastall (Hrsg.), *Handbuch der Gesundheitskommunikation: Kommunikationswissenschaftliche Perspektiven* (S. 15–27). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-10727-7_2.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hedges, L. V. (2018). Challenges in building usable knowledge in education. *Journal of Research on Educational Effectiveness*, 11(1), 1–21. <https://doi.org/10.1080/19345747.2017.1375583>.
- Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (4. Aufl.). Seelze: Klett.
- Hess, M., & Lipowsky, F. (2020). Zur (Un-)Abhängigkeit von Oberflächen- und Tiefenmerkmalen im Grundschulunterricht. Fragen von Lehrpersonen im öffentlichen Unterricht und in Schülerarbeitsphasen im Vergleich. In A.-K. Praetorius, J. Grünkorn & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 117–131). Weinheim, Basel: Beltz Juventa.
- Hinzke, J.-H., Gesang, J., & Besa, K.-S. (2020). Zur Erschließung der Nutzung von Forschungsergebnissen durch Lehrpersonen. Forschungsrelevanz zwischen Theorie und Praxis. *Zeitschrift für Erziehungswissenschaft*, 23(6), 1303–1323. <https://doi.org/10.1007/s11618-020-00982-6>.
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: social interdependence theory and cooperative learning. *Educational Researcher*, 38, 365–379.

- Johnson, M. L., & Sinatra, G. M. (2013). Use of task-value instructional inductions for facilitating engagement and conceptual change. *Contemporary Educational Psychology*, 38(1), 51–63. <https://doi.org/10.1016/j.cedpsych.2012.09.003>.
- Joyce, K. E., & Cartwright, N. (2020). Bridging the gap between research and practice: predicting what will work locally. *American Educational Research Journal*, 57(3), 1045–1082. <https://doi.org/10.3102/0002831219866687>.
- Kerwer, M., & Rosman, T. (2020). Vom Großen ins Kleine – Übertragbarkeit eines Modells der epistemischen Lebensspannenentwicklung auf den epistemischen Wandel in Kurzzeitinterventionen. *Psychologische Rundschau*, 71(2), 127–133. <https://doi.org/10.1026/0033-3042/a000485>.
- Kierner, K., & Kollar, I. (2018). Evidence-based reasoning of pre-service teachers: A script perspective. In J. Kay & R. Luckin (Hrsg.), *Rethinking learning in the digital age*. Making the Learning Sciences count, 13th International Conference of the Learning Sciences (ICLS) 2018. (Bd. 2, S. 1037–1040). London: International Society of the Learning Sciences.
- Kierner, K., & Kollar, I. (2021). Source selection and source use as a basis for evidence-informed teaching. *Zeitschrift Für Pädagogische Psychologie*, 35(2–3), 127–141.
- Kirkpatrick, L., Brown, H. M., Searle, M., Smyth, R. E., Ready, E. A., & Kennedy, K. (2018). Impact of a one-to-one ipad initiative on grade 7 students' achievement in language arts, mathematics, and learning skills. *Computers in the Schools*, 35(3), 171–185. <https://doi.org/10.1080/07380569.2018.1491771>.
- Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Videodokumente* (S. 42–57). Bonn: Bundesministerium für Bildung und Forschung.
- Knogler, M., Wiesbeck, A. B., & CHU Research Group (2018a). Kollaboratives Lernen und mobile digitale Geräte: Eine wirksame Kombination? www.clearinghouse-unterricht.de Kurzreview 15. Zugegriffen: 18.05.2022.
- Knogler, M., Wiesbeck, A. B., & CHU Research Group (2018b). Kollaboratives Lernen und mobile digitale Geräte: Eine wirksame Kombination? www.clearinghouse-unterricht.de Rating-Sheet zu Kurzreview 15. Zugegriffen: 18.05.2022.
- König, J. (2016). Lehrerexpertise und Lehrerkompetenz. In M. Rothland (Hrsg.), *Beruf Lehrer/Lehrerin. Ein Studienbuch* (S. 127–148). Münster: Waxmann.
- Kounin, J. S. (2006). *Techniken der Klassenführung*. Münster: Waxmann. Original der deutschen Ausgabe, 1976
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>.
- Kultusministerkonferenz (KMK) (2014). Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004 i.d. F. vom 12.06.2014. https://www.fu-berlin.de/sites/bologna/dokumente_zur_bologna-reform/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf. Zugegriffen: 12. Juni 2021.
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Schöningh.
- Kunter, M., Baumert, J., & Blum, W. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lysenko, L. V., Abrami, P. C., Bernard, R. M., Dagenais, C., & Janosz, M. (2014). Educational research in educational practice: predictors of use. *Canadian Journal of Education*, 37(2), 1–26.
- Malouf, D. B., & Taymans, J. M. (2016). Anatomy of an evidence base. *Educational Researcher*, 45(8), 454–459. <https://doi.org/10.3102/0013189X16678417>.
- Merk, S., Rosman, T., Rueß, J., Syring, M., & Schneider, J. (2017). Pre-service teachers' perceived value of general pedagogical knowledge for practice: Relations with epistemic beliefs and source beliefs. *PLOS ONE*, 13(2), e193632. <https://doi.org/10.1371/journal.pone.0184971>.
- Meyer, H. (2004). *Was ist guter Unterricht?* Berlin: Cornelsen – Scriptor.
- Neber, H. (2018). Kooperatives Lernen. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 354–362). Weinheim: Beltz.
- Neuweg, G. H. (2005). Implizites Wissen als Forschungsgegenstand. In F. Rauner (Hrsg.), *Handbuch Berufsbildungsforschung* (S. 581–588). Gütersloh: Bertelsmann.
- Neuweg, G. H. (2007). Wie grau ist alle Theorie, wie grün des Lebens goldener Baum? LehrerInnenbildung im Spannungsfeld von Theorie und Praxis. In *bwp@ Berufs- und Wirtschaftspädagogik – online*, Ausgabe 12. www.bwpat.de/ausgabe12/neuweg_bwpat12.pdf. Zugegriffen: 12. Juni 2021.

- Neuweg, G.H. (2014). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl. S. 583–614). Münster; New York: Waxmann.
- Pauli, C. (2020). Kommentar zum Themenblock „Oberflächen- und Tiefenstruktur des Unterrichts“. Nutzen und Grenzen eines prominenten Begriffspaars für die Unterrichtsforschung – und das Unterrichten. In A.-K. Praetorius, J. Grünkorn & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 132–137). Weinheim, Basel: Beltz Juventa.
- Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Hrsg.), *Advances in psychology, 131. Motivational psychology of human development: Developing motivation and motivating development* (S. 143–163). Amsterdam: Elsevier. [https://doi.org/10.1016/S0166-4115\(00\)80010-2](https://doi.org/10.1016/S0166-4115(00)80010-2).
- Piowar, V., Barth, V.L., Ophardt, D., & Thiel, F. (2018). Evidence-based scripted videos on handling student misbehavior: the development and evaluation of video cases for teacher education. *Professional Development in Education, 44*(3), 369–384.
- Praetorius, A.-K., & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft, 49*, 167–188. <https://doi.org/10.1007/s42010-021-00119-6>.
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of three basic dimensions. *ZDM, 50*(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>.
- Praetorius, A.-K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Sandy, T., & Charalambous, C. (2020a). Towards developing a theory of generic teaching quality. Origin, current status, and necessary next steps regarding the Three Basic Dimensions Model. In A.-K. Praetorius, J. Grünkorn & K. Eckhard (Hrsg.), *Empirische Forschung zur Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 15–36). Weinheim, Basel: Beltz Juventa.
- Praetorius, A.-K., Rogh, W., & Kleickmann, T. (2020b). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft, 48*(3), 303–318. <https://doi.org/10.1007/s42010-020-00072-w>.
- Prenzel, M. (2020). „Nützlich, praktisch, gut“: Erwartungen an die Forschung in der Lehrerinnen- und Lehrerbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung, 38*(1), 8–20. <https://doi.org/10.25656/01:21771>.
- Reiss, J. (2015). A pragmatist theory of evidence. *Philosophy of Science, 82*(3), 341–362. <https://doi.org/10.1086/681643>.
- Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? A plea for theories as primus inter pares. *Zeitschrift für Pädagogische Psychologie. https://doi.org/10.1024/1010-0652/a000345*.
- Rochnia, M., & Trempler, K. (2019). Welche externen Wissensquellen bevorzugen Lehrkräfte für ihr professionelles Handeln? *Lehrerbildung auf dem Prüfstand, 12*(2), 125–142.
- Rost, D.H. (2007). *Interpretation und Bewertung pädagogisch-psychologischer Studien. Eine Einführung* (2. Aufl.). Weinheim: Beltz.
- Rothland, M. (2016). Der Lehrerberuf als Gegenstand der Lehrerbildung. In M. Rothland (Hrsg.), *Beruf Lehrer/Lehrerin. Ein Studienbuch* (S. 7–15). Münster: Waxmann.
- Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ, 312*(7023), 71–72. <https://doi.org/10.1136/bmj.312.7023.71>.
- Scharlau, I. (2019). Sich verständigen. Überlegungen zur Frage der Evidenzbasierung. In T. Jenert, G. Reinmann & T. Schmohl (Hrsg.), *Hochschulbildungsforschung. Theoretische, methodologische und methodische Denkanstöße für die Hochschuldidaktik* (S. 105–123). Wiesbaden: Springer VS.
- Schellenbach-Zell, J., & Rochnia, M. (2020). Die Rolle der epistemologischen Überzeugungen von Studierenden im Praxissemester bei der Herstellung von Theorie-Praxis-Bezügen. In K. Rheinländer & D. Scholl (Hrsg.), *Verlängerte Praxisphasen in der Lehrer*innenbildung* (S. 198–212). Bad Heilbrunn: Klinkhardt.
- Schiefele, U., Köller, O., & Schaffner, E. (2018). Intrinsische und extrinsische Motivation. In D.H. Rost, J.R. Sparfeldt & S.R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 309–319). Weinheim: Beltz.

- Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the last decade: Role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>.
- Seidel, T., Knogler, M., Schneeweiss, A., Diery, A., Mazziotti, C., & Hetmanek, A. (2020). Auf dem Weg zu einer evidenzbasierten Lehrerbildung: Meilensteine und aktuelle Entwicklungen im Clearing House Unterricht. In Bundesministerium für Bildung und Forschung (Hrsg.), *Profilbildung im Lehramtsstudium. Beiträge der „Qualitätsoffensive Lehrerbildung“ zur individuellen Orientierung, curricularen Entwicklung und institutionellen Verankerung* (S. 85–94). Berlin: BMBF.
- Seidel, T., Renkl, A., & Rieß, W. (2021). Basisdimensionen für Unterrichtsqualität im Fachkontext konkreter Situationen: Die Rolle von Unterrichtsartefakten und Bestimmung von Standardsituationen. *Unterrichtswissenschaft*, 49, 293–301. <https://doi.org/10.1007/s42010-021-00108-9>.
- Shulman, L. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>.
- Slavin, R.E. (2020). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, 55(1), 21–31. <https://doi.org/10.1080/00461520.2019.1611432>.
- Slavin, R.E., Cheung, A.C.K., & Zhuang, T. (2021). How could evidence-based reform advance education? *ECNU Review of Education*, 4(1), 7–24. <https://doi.org/10.1177/2096531120976060>.
- Stark, R. (2017). Probleme evidenzbasierter bzw. -orientierter pädagogischer Praxis. *Zeitschrift für Pädagogische Psychologie*, 31(2), 99–110. <https://doi.org/10.1024/1010-0652/a000201>.
- Stark, R., & Mandl, H. (2007). Bridging the gap between basic and applied research by an integrative research approach. *Educational research and evaluation*, 13(3), 249–261. <https://doi.org/10.1080/13803610701632075>.
- Stark, L., Wilkes, T., Uhlenbrock, J., Klein, M., & Stark, R. (2022, in Druck). Förderung evidenzorientierten Problemlösens in der Lehramtsausbildung: Effekte zweier Verfahren zur Induktion eines neutralen vs. positiven Nützlichkeitswertes. In K.S. Besa, D. Demski, J. Gesang & J.-H. Hinzke (Hrsg.), *Evidenzorientierung in Lehrer*innenbildung, Schule und Bildungspolitik und -administration – neue Befunde zu alten Problemen*. Wiesbaden: Springer VS.
- Sung, Y.T., Yang, J.M., & Lee, H.Y. (2017). The effects of mobile-computer-supported collaborative learning: meta-analysis and critical synthesis. *Review of educational research*, 87(4), 768–805. <https://doi.org/10.3102/0034654317704307>.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, Dordrecht, Heidelberg, London: Springer.
- Taylor, J.A., Davis, E., & Michaelson, L.E. (2021). Considerations for evidence frameworks in education research. *Review of Research in Education*, 45(1), 101–128. <https://doi.org/10.3102/0091732X20985077>.
- Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. *Reading and Writing*, 29(8), 1629–1652. <https://doi.org/10.1007/s11145-016-9638-8>.
- Thomm, E., Gold, B., Betsch, T., & Bauer, J. (2021a). When preservice teachers' prior beliefs contradict evidence from educational research. *British Journal of Educational Psychology*, 91(3), 1055–1072. <https://doi.org/10.1111/bjep.12407>.
- Thomm, E., Sälzer, C., Prenzel, M., & Bauer, J. (2021b). Predictors of teachers' appreciation of evidence-based practice and educational research findings. *Zeitschrift für Pädagogische Psychologie*, 35(2–3), 173–184. <https://doi.org/10.1024/1010-0652/a000301>.
- Thomm, E., Seifried, E., & Bauer, J. (2021c). Informing professional practice: (Future) teachers' choice, use, and evaluation of (non-)scientific sources of educational topics. *Zeitschrift für Pädagogische Psychologie*, 35(2–3), 121–126. <https://doi.org/10.1024/1010-0652/a000309>.
- Trempler, K., Hetmanek, A., Wecker, C., Kiesewetter, J., Wermelt, M., Fischer, F., Fischer, M., & Gräsel, C. (2015). Nutzung von Evidenz im Bildungsbereich. Validierung eines Instruments zur Erfassung von Kompetenzen der Informationsauswahl und Bewertung von Studien. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden*. Zeitschrift für Pädagogik, Beiheft 61. (S. 144–166). Weinheim: Beltz Juventa.
- Vieluf, S., Praetorius, A.-K., Rakoczy, K., Kleinknecht, M., & Pietsch, M. (2020). Angebots-Nutzungsmodelle der Wirkweise des Unterrichts: eine kritische Auseinandersetzung mit ihrer theoretischen Konzeption. *Zeitschrift für Pädagogik*, 66(1), 63–80.
- Wagner, K., Bergner, M., Krause, U.-M., & Stark, R. (2018). Förderung wissenschaftlichen Denkens im Lehramtsstudium: Lernen aus eigenen und fremden Fehlern in multiplen und uniformen Kontexten. *Zeitschrift Für Pädagogische Psychologie*, 32(1–2), 5–22. <https://doi.org/10.1024/1010-0652/a000219>.

- Wecker, C., Vogel, F., & Hetmanek, A. (2017). Visionär und imposant – aber auch belastbar? Eine Kritik der Methodik von Hatties „Visible Learning“. *Zeitschrift für Erziehungswissenschaft*, *20*(1), 21–40.
- Wekerle, C., & Kollar, I. (2021). Fostering pre-service teachers' situation-specific technological pedagogical knowledge—does learning by mapping and learning from worked examples help? *Computers in Human Behavior*, *115*(106617), 1–14. <https://doi.org/10.1016/j.chb.2020.106617>.
- Wenglein, S., Bauer, J., Heininger, S., & Prenzel, M. (2015). Kompetenz angehender Lehrkräfte zum Argumentieren mit Evidenz: Erhöht ein Training von Heuristiken die Argumentationsqualität? *Unterrichtswissenschaft*, *43*(3), 209–224.
- Zlatkin-Troitschanskaia, O., Zimmer, L., Mater, O., Laier, B., Koch, A., Binnewies, C., Dormann, C., van Ackeren, I., Clausen, M., Preisendörfer, P., Schmidt, U., Demski, D., Preuß, D., & Stump, M. (2016). Schulische und individuelle Einflussfaktoren auf das evidenzbasierte Handeln von Lehrkräften und Schulleitungen – eine mehrbenenanalytische Studie. In Bundesministerium für Bildung und Forschung (Hrsg.), *Implementation und Wirkung neuer Steuerungsinstrumente im Schulwesen* (Bd. 43, S. 8–38). Berlin: BMBF.