

Anna Heinemann, Jan Koschorreck, Pia Sander, Annika Wilmers,
Iris Nieding & Jens Leber

Der Screening- und Kodierprozess im Critical Review

Eine vergleichende Auswertung und bildungssektorübergreifende
Lessons Learned

Abstract

Ein Critical Review beinhaltet neben der systematischen Recherche eine ebenso transparente und systematische Methodik für die Inkludierung und Exkludierung von Publikationen sowie der Kodierung der eingeschlossenen Literatur (Evans & Benefield, 2001). Die in diesem Kapitel vorgestellten Überblicke sowie Reflexionen über das methodische Verfahren im Critical Review entstammen aus vier Jahren, in denen im Metavorhaben „Digitalisierung im Bildungsbereich“ jeweils ein Review pro Jahr aus verschiedenen Bildungsbereichen zu einem übergeordneten Thema im Kontext des Metavorhabens entstanden sind. Die Autor:innen stellen die eigene Methodik dar und reflektieren sie anhand konkreter Erfahrungen.

Zunächst wird ein theoretischer Überblick über die methodischen Schritte des Screenings sowie Kodierens gegeben und das Verfahren des Critical Reviews methodisch eingeordnet. Im Anschluss daran wird ein Überblick über die in diesem Beitrag vorgestellten Reviews gegeben, indem z. B. Trefferzahlen sowie die ausgeschlossene Anzahl an Titeln vorgestellt werden. Im darauffolgenden Abschnitt werden die Erfahrungen der Autor:innen qualitativ in den aktuellen Stand der Methodik in den Bildungswissenschaften eingeordnet. Ziel ist es, einen Beitrag zu diesem methodischen Diskurs zu leisten und Lessons Learned aus dem vorliegenden Projekt vorzustellen.

Schlüsselwörter: Systematic Review, Critical Review, methodologische Einordnung, Screeningprozess, Kodieren

The screening and coding process in critical reviews. A comparative evaluation and lessons learned across educational sectors

In addition to systematic research, a critical review includes an equally transparent and systematic methodology for the inclusion and exclusion of publications and the coding of the included literature (Evans & Benefield, 2001). The overviews and reflections on the methodological procedure in the critical reviews presented in this chapter originate from four years in which the meta-project “Digitalisation in Education” has produced one review per year from different areas of education on an overarching topic in the context of the meta-project. The authors present and reflect their methodological approach based on concrete experiences.

First, a theoretical overview of the methodological steps of screening and coding is given and the critical review process is methodologically classified. This is followed by an overview of the reviews presented in this chapter, e.g. by presenting the number of hits

and the number of titles excluded. In the following section, the authors' experiences are qualitatively classified in the current state of methodology in the educational sciences. The article aims to contribute to this methodological discourse and to present lessons learned from the project.

Keywords: systematic review, critical review, methodological contribution, screening, coding

1. Einleitung

Ein Critical Review besteht aus mehreren Arbeitsschritten, die zunächst die Formulierung einer Forschungsfrage, die Anlegung von Suchstrings und die Durchführung der Suche in festgelegten Datenbanken und anschließend den Auswahlprozess der recherchierten Literatur und die Analyse der ausgewählten Titel umfassen. Der systematische Auswahlprozess nach definierten Kriterien ist ein wesentliches Charakteristikum in einem Critical Review, welches eine Form der Forschungssynthese ist (Grant & Booth, 2009). Ziel ist es, nur solche Studien einzuschließen, die auch der Forschungsfrage entsprechen, sowie die Qualität der Primärforschung einschätzen zu können, um valide sowie nützliche Ergebnisse in der Synthese zusammen zu tragen (Gough et al., 2017; Porritt et al., 2014). Darüber hinaus versucht die Methodik des Critical Reviews, Forschung und wissenschaftliche Beiträge zusammen zu tragen, um ggfs. modellerweiternd zu arbeiten (Wilmers et al., 2020).

Im Anschluss an die vorangegangenen Kapitel über den Rechercheprozess werden im folgenden Beitrag die Schritte des Screening- und Kodierprozesses (Gough et al., 2017; Siddaway et al., 2019) innerhalb der Critical Reviews dargestellt, die im Metavorhaben Digitalisierung im Bildungsbereich (Diekmann, 2022; Kerres et al., 2022) entstanden sind. Dabei beschreibt der folgende Beitrag zunächst die involvierten Bildungsbereiche und die Arbeit während der Erstellung der Reviews mit Bezug auf die Teamkonstellationen sowie einer Auswertung der Anzahlen der Titel, die im Auswahlprozess gescreent und letztendlich ausgewählt wurden als Hintergrund für darauf folgende primäre Anliegen dieses Kapitels: die Reflexion des Vorgehens in den Arbeitsschritten vom Titel- und Abstractscreening über das Volltextscreening bis zum Kodieren der eingeschlossenen Literatur in den Critical Reviews.

Dieses Kapitel bezieht sich auf 12 von insgesamt 20 Reviews¹, die während der Projektlaufzeit in den Bildungsbereichen BKJF (Bildung in Kindheit, Jugend und Familie), SB (Schulische Bildung) und EB/WB (Erwachsenen- bzw. Weiterbildung)² in insgesamt vier Bänden zwischen 2020–2023 entstanden sind. Den Screeningprozess reflektieren die Autor:innen anhand der Schritte Titel-, Abstract- und Volltextscreening sowie Kodierung basierend auf ausgewählten Beispielen aus diesen Cri-

1 Die hier behandelten 12 Reviews: Heinemann et al. (2022); Koschorreck und Gundermann (2020, 2021, 2022, 2023); Leber et al. (2023); Nieding (2022, 2023); Nieding und Klauudy (2020, 2021); Waffner (2020, 2021).

2 Auch im Bildungsbereich der Lehrkräftebildung und der beruflichen Bildung wurden acht Critical Reviews verfasst, die aber in den Inhalten dieses Kapitels nicht vertreten sind.

tical Reviews. Es wird jeweils der Frage nachgegangen, welche Herausforderungen und Gelingensbedingungen beim Screeningprozess sowohl auf methodischer als auch auf Arbeitsebene identifiziert werden konnten. Darüber hinaus wird der Versuch unternommen, die gesammelten Erfahrungen in den aktuellen Diskurs der Forschungsliteratur bezüglich der Methodik von Systematic Reviews einzuordnen. Hiermit möchten die Autor:innen einen Beitrag zur Verortung der Methode innerhalb der Bildungs- und Erziehungswissenschaften leisten. Der diesbezügliche Diskurs ist vergleichsweise jung, da das Format der Forschungssynthesen zunächst Eingang in die Disziplin der Medizin fand und daran anschließend in die Bildungs- und Erziehungswissenschaften übertragen wurde (Evans & Benefield, 2001; Gough et al., 2017; Heinemann et al., 2023; Petticrew & Roberts, 2006).

Der gesamte Reviewprozess teilte sich in zentrale und dezentrale Arbeitsschritte, wobei bei ersteren im gesamten Team über die Bildungsbereiche hinweg und v. a. unter der Leitung und Moderation des DIPF gearbeitet wurde und bei letzteren in den Teams der einzelnen Bildungsbereiche. Bevor die Teams in die Bearbeitung ihrer Bildungsbereiche gingen, wurde im Gesamtteam ein gemeinsames Kodierschema erstellt, welches als Grundlage für alle Kodierprozesse diente und je nach Bildungsbereich oder Forschungsfrage jeweils angepasst oder erweitert werden konnte. Der Auswahlprozess der Studien fand durch die Teams in einem dezentral angelegten Arbeitsprozess statt und wurde maßgeblich innerhalb der Bildungsbereiche durchgeführt. Jedoch wurden auch während dieser dezentralen Arbeitsschritte monatliche Austausch- und Reflexionstreffen im Gesamtteam abgehalten.

2. Methodische Einordnung des Screenings im Reviewprozess

Der Screeningprozess in Forschungssynthesen umfasst die Definition von Ein- sowie Ausschlusskriterien, das Screening der Titel und Abstracts (Screening 1) sowie der Volltexte (Screening 2). Hieran schließt sich die Kodierung und Überprüfung der Qualität der Primärliteratur an (Buntins, Mulders & Schroeder, 2023; Grant & Booth, 2009; Hallinger, 2013; Porritt et al., 2014; Wilmers et al., 2021).

Der systematische Suchprozess während der Recherche (siehe Wilmers & Keller in diesem Band) ergibt je nach Forschungskontext und Forschungsfrage ein individuell umfangreiches Literaturpaket, welches dahingehend überprüft werden muss, ob die identifizierte Literatur für das Review passend und relevant erscheint. Maßgeblich für diese Einschätzung ist die Relevanz der Studien für die Beantwortung der Forschungsfrage. Dieser Vorgang findet typischerweise in zwei aufeinanderfolgenden und teilweise auch ineinander übergehenden Arbeitsschritten statt, die als Screening bezeichnet werden (vgl. z.B. Higgins et al., 2023; Polanin et al., 2019). Zuerst werden die Titel und Abstracts der identifizierten Literatur durchsucht und diejenigen in die Volltextprüfung aufgenommen, die nach den Formalia sowie der in Titel und Abstract vorhandenen Information die entsprechenden Kriterien erfüllen. Ebenso

werden zur weiteren Beurteilung auch als uneindeutig kategorisierte Befunde in die Volltextprüfung aufgenommen. Danach folgt das Screening des Volltextes und, sofern der Text für das Review als relevant eingestuft wird, die Kodierung der Literatur anhand eines Kodierschemas (Harden & Gough, 2012; Petticrew & Roberts, 2006; Sutcliffe et al., 2017). Ein Vorteil des zweistufigen Screenings ist, dass ein großer Teil der Literatur durch die vorgeschaltete Prüfung im Screening 1 aussortiert werden kann und damit keiner deutlich zeitaufwendigeren Volltextprüfung bedarf. Gerade in der Bildungsforschung zeigt sich indes auch, dass ein solches Vorgehen nicht immer 1:1 umsetzbar ist, da Abstracts beispielsweise nicht immer die methodischen Kriterien listen, die zur Entscheidungsfindung notwendig sind. In jedem Fall ist das „critical appraisal“ (Newman & Gough, 2020, S. 13; Petticrew & Roberts, 2006) für das Screenen der Literatur sowie für das nachfolgende Kodieren der eingeschlossenen Literatur ein essenzieller methodischer Schritt um die Bewertung der Qualität systematisch durchzuführen. Dafür beschreiben Newman und Gough (2020, S. 12) drei wichtige Grundpfeiler:

“(i) details of the studies to enable mapping of what research has been undertaken; (ii) how the research was undertaken to allow assessment of the quality and relevance of the studies in addressing the review question; (iii) the results of each study so that these can be synthesised to answer the review question.”

Wie genau dabei das Kodierschema im individuellen Fall festgelegt wird, entscheidet sich über die Forschungsfrage und bestimmte Entscheidungen für das Review (Wetterich & Plänitz, 2021). Wilmers et al. (2022) führen diesbezüglich auf, dass sich das Vorgehen beim Kodieren durch „das gewählte Reviewformat (Methodik und Umfang des Reviews), der thematischen Schwerpunktsetzung (weit gefasste oder spezialisierte Fragestellung) und den Eckpunkten der Recherchestrategie (breite oder eingeschränkte Suche)“ (S. 20) definiert. Damit reflektiert die Kodierung auch die Methodenvielfalt in der Bildungsforschung und die Vielfalt möglicher Zugänge zum Feld, hält aber gleichzeitig an dem Anspruch fest, eine systematische (Meta-)Ebene zum Umgang mit der vorhandenen Forschungsliteratur bereitzustellen. Sutcliffe et al. (2017) legen für diesen Entscheidungsprozess ein Spektrum zum Kodierungsprozess vor, welches eine systematische Grundlage bietet, um das eigene Review methodisch einzuordnen. Je nach Forschungsfrage lassen sich für die Kodierung im Vorhinein bereits deduktiv Kategorien durch Rahmenmodelle und -theorien ableiten und/oder weitere Kategorien durch die Sichtung der Literatur im Reviewverfahren hinzufügen.

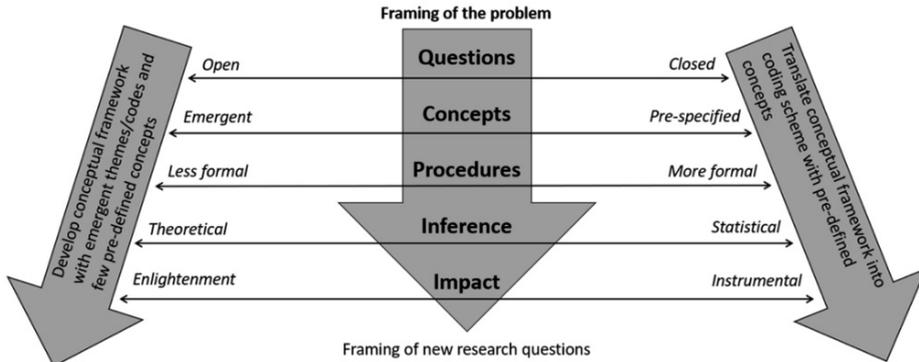


Abbildung 1: Spektrum zum Kodierprozess (Sutcliffe et al., 2017, S. 137)

Für den Kodierprozess stellen Sutcliffe et al. (2017, S. 135f.) zudem drei essentielle Prinzipien auf:

1. Kohärent und relevant: Im Sinne der Synthese und vergleichenden Analyse sollten sich die Codes kohärent zur Fragestellung des Reviews verhalten und die dafür relevanten Informationen erschließen.
2. Reliabel und explizit: Die Codes sollten klar und eindeutig formuliert sein, sodass alle Mitarbeitenden im Review ein gemeinsames Verständnis der Codes haben und gleich kodieren.
3. Bewusstsein für Subjektivität: Auch wenn das zweite Prinzip verfolgt wird, muss berücksichtigt werden, dass die Verdichtung von einer großen Menge an Informationen in ein Schema Interpretationen des Einzelnen mit sich bringt.

Vor allem das dritte Prinzip unterstreicht die essentielle Bedeutung der Zusammenarbeit und des iterativen kommunikativen Prozesses im Reviewverfahren.

In allen benannten Arbeitsschritten ist die Berechnung der Interrater-Reliabilität im Team, auch bereits durch eine einfache Unterstützung der prozentualen Übereinstimmung, und die iterative Diskussion möglicher Abweichungen u. a. im Sinne der Reliabilität des Critical Reviews wünschenswert (Buntins, Bedenlier et al., 2023; Viera & Garrett, 2005).

Die nachfolgenden Kapitel dienen der Darstellung und Reflexion der Zusammenarbeit über mehrere erstellte Critical Reviews hinweg. Beschrieben und reflektiert werden daher mehrere Jahre der Zusammenarbeit und des fachlichen Austauschs in den Arbeitsschritten des Titel- bzw. Abstract- und Volltextscreening sowie des Kodierens der eingeschlossenen Literatur.

3. Projekthintergrund

Die hier beschriebenen Reviews der einzelnen Bildungsbereiche waren eingebettet in die Arbeit eines Gesamtteams, das aus bis zu 13 Personen bestand. Alle Reviews eint die Systematik, dass eine übergeordnete Person verantwortlich war für die Durchführung der Recherche. Darüber hinaus fanden regelmäßige Treffen des Gesamtteams statt, in denen methodische und inhaltliche Herausforderungen diskutiert wurden. Dieser Diskurs wurde von einer Koordinatorin moderiert und zusammengeführt, u. a. um gemeinsame Komponenten in der Methodik festzulegen, das methodische Vorgehen zu prüfen, den eigenen Vorgang zu reflektieren und die Arbeitsprozesse gegenseitig zu unterstützen.

Die verschiedenen vertretenen Expertisen im Gesamtteam sowie in den einzelnen Bildungsbereichen waren für die Methodik im Review ein großer Mehrwert (Akers et al., 2009). Die Zusammensetzung der Mitarbeitenden in den Reviewteams unterschied sich zwischen den Bildungsbereichen und innerhalb der Bildungsbereiche über die Zeit (siehe Abb. 2): Die vier Critical Reviews im Bildungssektor Kindheit, Jugend und Familie (BKJF) wurden hauptverantwortlich von einer wissenschaftlichen Mitarbeiterin erstellt. Die Erarbeitung der Themen sowie Teile des Screenings erfolgten gemeinsam mit einer weiteren wissenschaftlichen Mitarbeiterin und punktuell wurde auf die Unterstützung einer studentischen Hilfskraft zurückgegriffen, beispielsweise bei der Literaturbeschaffung. Auch im Bereich der Schulischen Bildung (SB) sind vier Critical Reviews entstanden, die anfänglich von einer wissenschaftlichen Mitarbeiterin erstellt wurden und ab dem dritten Review von einem Autor:innenteam von drei wissenschaftlichen Mitarbeitenden. Darüber hinaus haben bis zu drei Hilfskräfte an den Reviews mitgearbeitet (siehe Ahmad et al. in diesem Band). Diese personelle Ergänzung war notwendig, da im schulischen Sektor grundsätzlich eine hohe Anzahl an potenziell relevanten Titeln recherchiert wurde, die in den Screeningprozess mit eingingen. Im Bereich der Erwachsenen- und Weiterbildung (EB/WB) arbeiteten hauptverantwortlich zwei wissenschaftliche Mitarbeitende an den vier Reviews. Die Zusammensetzung dieses Autor:innenteams änderte sich im Projektverlauf nicht. Im Laufe der Erarbeitung des ersten Reviews zeigte sich, dass die Literaturverwaltung und -digitalisierung einen erheblichen Mehraufwand bedeutete, deshalb wurden solche Arbeiten fortan von einer Assistentkraft unterstützt. Abstimmungsprozesse, z. B. hinsichtlich der Interrater-Reliabilität (Siddaway et al., 2019) wurden u. a. in den Autor:innenteams EB/WB und SB durch regelmäßige Besprechungen in allen Arbeitsschritten realisiert. Im Laufe der Zeit wurde ein größerer Bedarf erkannt, mehr Unterstützung – insbesondere in Form von Hilfs- und Assistentkräften – bei der Erstellung der Reviews in Anspruch zu nehmen.

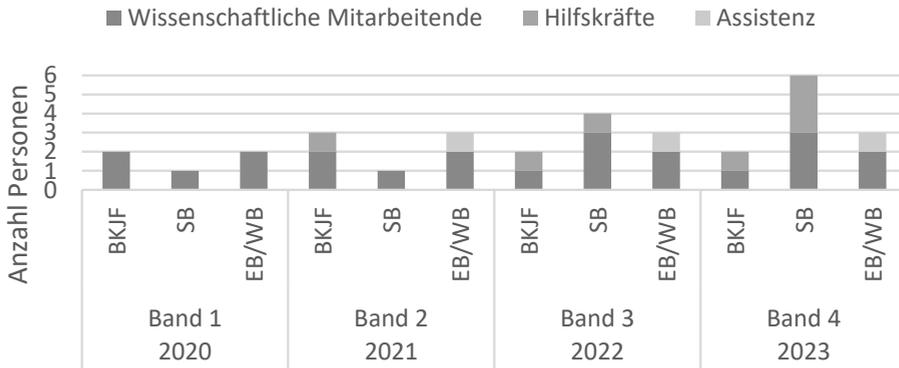


Abbildung 2: Teamkonstellationen in den Bildungssektoren und Bänden
 BKJF = Bildung in Kindheit, Jugend und Familie; SB = Schulische Bildung;
 EB/WB = Erwachsenen-/Weiterbildung

Der hohe Bedarf an personellen und zeitlichen Ressourcen wird auch deutlich, wenn die Anzahl der Titel betrachtet wird, die über den Projektzeitraum von fünf Jahren bearbeitet wurden. Insgesamt wurden in den 12 für diese Auswertung zu Grunde liegenden Reviews 31.098 Titel in der Recherche identifiziert und auf ihre grundsätzliche Passung für die Fragestellung gescreent. Davon wurden 2.826 Titel als potenziell relevant identifiziert und im Volltext gescreent. Schlussendlich wurden 421 Titel analysiert, synthetisiert und in die Reviews aufgenommen. Abbildung 3 zeigt die Anzahl der insgesamt bearbeiteten Titel pro Bildungssektor; Abbildung 4 schlüsselt diese Anzahlen zusätzlich für jeden Band auf. Solche hohen Treffermengen sind eine Herausforderung für den Screeningprozess, der man methodisch beispielsweise durch die Anwendung eines Referenzmodells begegnen kann (Heinemann et al., 2023). So können Publikationen qualitativ eingeordnet und eine Grundlage für die Synthese geschaffen werden.

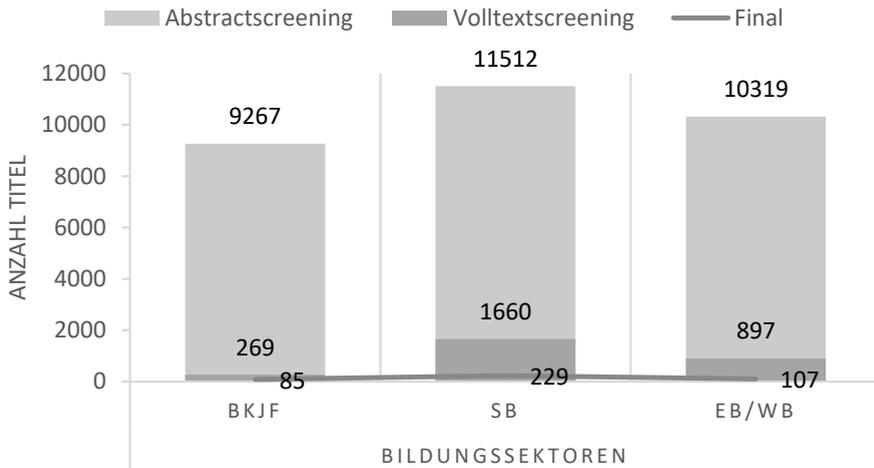


Abbildung 3: Anzahl der insgesamt bearbeiteten Titel pro Bildungssektor
 BKJF = Bildung in Kindheit, Jugend und Familie; SB = Schulische Bildung;
 EB/WB = Erwachsenen-/Weiterbildung

Betrachtet man die Anzahl der Titel, die im Abstractscreening bearbeitet wurden, gesondert für die einzelnen Reviews der Bildungsbereiche (Abb. 5), ist festzustellen, dass nicht allein der Sektor für eine hohe Anzahl an Treffern, die gescreent werden müssen, ausschlaggebend ist, sondern auch andere Erklärungen greifen. Zum Beispiel können die gewählte Fragestellung und die fachliche Expertise der Mitarbeitenden bei der Festlegung der Fragestellung und der Bestimmung der Suchstrings einen Einfluss auf die Zusammensetzung des Literaturbestandes haben.

Auffällig sind die Unterschiede in der Reduzierung der Gesamttreffermenge nach dem Abstractscreening: Während im Sektor SB im Durchschnitt 84% der Titel aussortiert wurden, wurden im Sektor BKJF 96% und im Sektor EB/WB 91% durchschnittlich aussortiert. Das Volltextscreening hat im Durchschnitt zu einer Reduzierung der relevanten Titel von 98% in der SB und 99% in BKJF und EB geführt. Abbildung 5 zeigt die prozentualen Anteile der ausgeschlossenen Titel nach den beiden Screeningschritten pro Sektor und Band. Insgesamt kann man feststellen, dass die prozentuale Anzahl der aussortierten Titel nach dem Abstractscreening deutlich zwischen den Sektoren variiert (zwischen 76,3% und 98,9%) und wenig innerhalb der Sektoren. Tendenziell wurden im Sektor SB am wenigstens Titel nach dem Abstractscreening ausgeschlossen. Nach dem Volltextscreening ist auch zwischen den Sektoren nur noch ein geringer Unterschied in der prozentualen Anzahl der ausgeschlossenen Titel zu erkennen.

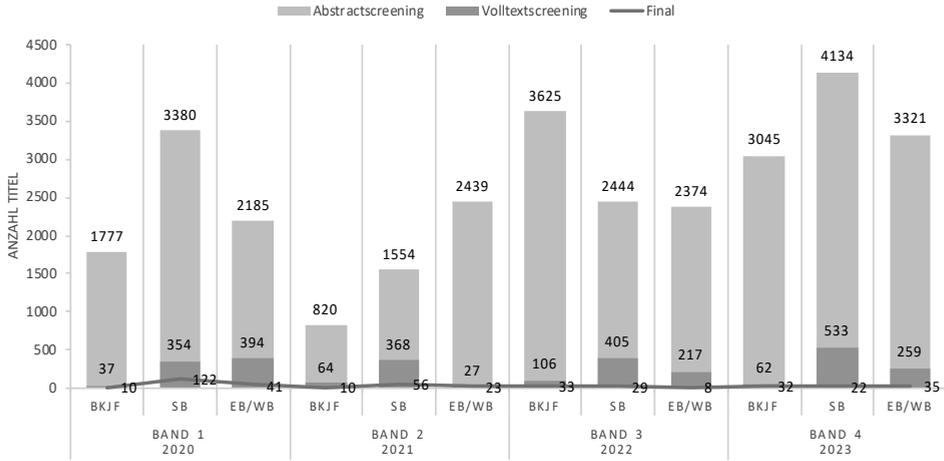


Abbildung 4: Anzahl der bearbeiteten Titel pro Bildungssektor und Band
 BKJF = Bildung in Kindheit, Jugend und Familie; SB = Schulische Bildung;
 EB/WB = Erwachsenen-/Weiterbildung

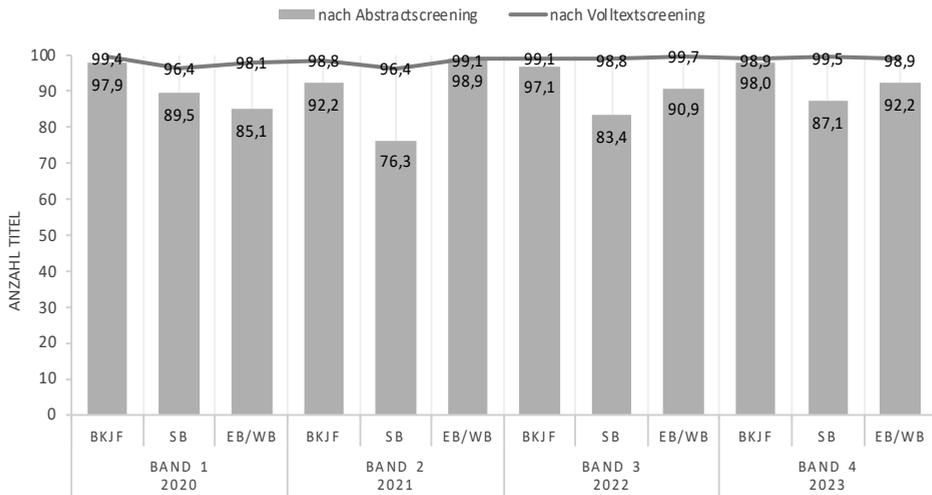


Abbildung 5: Ausgeschlossene Titel nach Abstract- und Volltextscreening in Prozent
 BKJF = Bildung in Kindheit, Jugend und Familie; SB = Schulische Bildung;
 EB/WB = Erwachsenen-/Weiterbildung

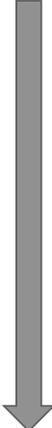
Abschließend lässt sich festhalten, dass unabhängig von der Fragestellung, der Gesamttreffermenge, der Anzahl der ausgeschlossenen Titel und des Vorgehens im Abstractscreening, die verschiedenen Autor:innenteams sehr ähnliche Anzahlen an Titeln nach dem Volltextscreening final in die Reviews aufgenommen haben. Die Anzahlen liegen dabei unter 4 % und meist sogar unter 2 %. Diese Zahlen sind für Reviews in der Bildungsforschung nicht ungewöhnlich, da das Forschungsfeld mitunter durch die Komplexität von Lehr- und Lernsituationen (Hammersley, 2020; Heinemann et al., 2023) sowie nicht deutlich abgegrenzte Begrifflichkeiten geprägt ist (Heinemann et al., 2022; Leber et al., 2023). Die Bildungsforschung weist darüber hinaus eine hohe Vielfalt an Methoden und Dokumententypen auf, deren Einbezug in das finale Review nicht immer den angelegten Qualitätskriterien entspricht (z. B. keine empirische Studie).

4. Reflexion des methodischen Vorgehens und Lessons Learned

Der folgende Abschnitt beschreibt und reflektiert das Screening- und Kodierverfahren. Generell erfolgte die Verarbeitung der Literatur über spezialisierte Anwendungen (in der Regel Citavi), ggf. wurden in einzelnen Bildungsbereichen weitere Programme verwendet, um den Prozess zu unterstützen oder einzelne Schritte ganz auszulagern (z. B. Excel). Dabei hat es sich auch als sinnvoll herausgestellt, die unterschiedlichen Schritte im Screening-Prozess in jeweils neuen Citavi-Dateien zu bearbeiten, um das systematische Vorgehen zu unterstützen und die Übersichtlichkeit und Handhabbarkeit der Literatur zu verbessern. Die Verwendung einer spezialisierten Software zur Erstellung von Forschungssynthesen wie dem EPPI-Reviewer war im Projektrahmen hingegen nicht vorgesehen.

Abbildung 6 veranschaulicht beispielhaft den Screening- und Kodierprozess anhand ausgewählter Ein- und Ausschlusskriterien des Reviews über die Förderung von Informationskompetenzen von Schüler:innen (Leber et al., 2023). Deutlich wird, dass die Prüfung formaler Kriterien, wie das Erscheinungsjahr oder die Sprache, in der die Publikation verfasst wurde, bereits im Titel- und Abstractscreening begonnen werden können, während sich viele inhaltliche Kriterien, wie der Bezug zur Fragestellung des Reviews teilweise erst bei der Analyse der Volltexte final bestimmen lassen. Nicht alle formalen Kriterien sind dabei aber immer eindeutig – ob die Publikation tatsächlich wissenschaftlichen Anforderungen, wie z. B. der Validität oder der Reliabilität entspricht oder ob sie nicht doch grauer Literatur zuzuordnen ist, kann daher manchmal auch erst im Volltextscreening geklärt werden.

Reviewtitel		
Informationskompetenz von Schüler:innen stärken: Wie kann man das Prüfen von Informationen durch den Einsatz digitaler Medien fördern?		
Kriterium	Einschluss	Ausschluss
Sprache	Englisch & Deutsch	-
Erscheinungsdatum	ab 2018	-
Dokumententyp	Publikationen in wissenschaftlichen Journals	Monografien, Sammelbände, Dissertationen, graue Literatur, Konferenzbeiträge
Publikationsart	empirische Publikationen	Konzeptuelle Publikationen oder keine transparente Methodik bei empirischen Publikationen
Bezug zur Fragestellung	Beschreibung einer Bildungsmaßnahme zur Förderung von Informationskompetenz oder eines ähnlichen Konstrukts unter Einbezug von digitalen Medien	<ul style="list-style-type: none"> • Keine Bildungsmaßnahme; • Kein Bezug zur schulischen Bildung • Schüler:innen sind nicht primäre Zielgruppe der Untersuchung; • Keine Förderung der Informationskompetenz oder eines ähnlichen Konstrukts; • Kein Einsatz digitaler Medien; mangelnde Transparenz in Bezug auf Medieneinsatz, forschungsmethodisches Vorgehen, didaktisch-methodisches Vorgehen oder die zu förmernde Kompetenz; • Digitalisierung oder Einbezug von digitalen Medien nicht im Fokus • Kein Bezug zur Fragestellung



Abstract-screening

Volltext-screening

Kodier-schema

Abbildung 6: Beispiel für Ein- und Ausschlusskriterien in einem Review (angelehnt an Leber et al., 2023)

4.1 Definition der Ein- und Ausschlusskriterien sowie Screening der Titel und Abstracts

In einem ersten Schritt haben alle Sektoren den Fokus auf Formalia gelegt. Dazu zählte eine Eingrenzung nach Veröffentlichungsjahren, die Sprache der Veröffentlichung in Deutsch oder Englisch und/oder der Ausschluss bestimmter Schlagworte oder Dokumentenarten wie graue Literatur, Konferenzbeiträge oder Dissertationen, wenn dies bereits anhand der Titel oder der ergänzenden Informationen im Literaturverwaltungsprogramm ersichtlich wurde. Je nach Thema bzw. Bildungsbereich kann auch das Bezugsland der wissenschaftlichen Arbeit bereits hier ein Ausschlusskriterium darstellen, wenn die internationale Übertragbarkeit nicht gegeben war.³ Dieser eher formale Teil des Prozesses kann mit vergleichsweise geringem Einarbeitungsaufwand durch wissenschaftliche Hilfskräfte unterstützt werden. So wurde beispielsweise im Sektor Bildung in Kindheit, Jugend und Familie eine wissenschaftliche Hilfskraft hinzugezogen, um das Literaturkorpus durch ein anhand klarer Kriterien festgelegten Screenings zu reduzieren. Aufgrund der Beschaffenheit des Sektors sowie seiner Inhalte gibt es zahlreiche Überschneidungen zum Sektor schulische Bildung. Diese Überschneidungen wurden bereinigt, indem die Hilfskraft aus dem deutschsprachigen Korpus alle Titel aussortierte, die sich dem Titel nach offensichtlich mit dem Aspekt „Schule“ beschäftigten. Stichpunkte, an denen sich die Hilfskraft orientieren konnte, waren bspw. die Begriffe Lehrkraft oder Unterricht.

³ Dies zeigt sich beispielsweise besonders deutlich im hier nicht weiter berücksichtigten Bildungsbereich der beruflichen Bildung (z. B. Hähn & Ratermann-Busse, 2020).

Dadurch konnte der Aufwand der darauffolgenden Schritte für die Autor:innen reduziert werden.

Nach dieser Prüfung von Ein- oder Ausschluss auf Basis der eben beschriebenen Formalia wurden die Ein- und Ausschlusskriterien stärker an inhaltlichen Aspekten orientiert. Dazu wurden insbesondere Titel, Abstracts sowie vorhandene Verschlagwortung der Einträge geprüft. Erfahrungsgemäß kam es selten vor, dass schon aus dem Titel eines Eintrags erkennbar war, ob es sich um eine empirische Arbeit handelte oder die Passung zum Bildungsbereich gegeben war. Dabei zeigte sich, dass in der Praxis die theoretische Prozessbeschreibung von Sutcliffe (vgl. Abschnitt 2) nicht ohne Anpassungen umsetzbar war. Daraus entwickelte sich ein Vorgehen, das Ähnlichkeiten mit einer qualitativ-inhaltsanalytischen Auswertungsstrategie aufwies; die Festlegung der inhaltlichen Kriterien erfolgte nach einer deduktiven und einer induktiven Herangehensweise. Die deduktiven Kriterien entsprachen dem ersten Teil eines gemeinsam in der Arbeitsgruppe entworfenen Kategorienschemas und wurden entsprechend des jeweiligen Bildungsbereichs und dem Thema bzw. der Forschungsfrage konkretisiert. Beispielsweise wurden im Bereich der EB/WB Arbeiten ausgeschlossen, bei denen es sich nicht um empirische Studien aus dem Bildungsbereich handelte oder Studien, die keine Erwachsenen untersuchten (Koschorreck & Gundermann, 2023). Dabei müssen die Kriterien ggf. induktiv weiter differenziert werden, wenn zum Beispiel im Titel, in Abstracts oder Schlagworten verwendete Synonyme eine Ergiebigkeit für die eigene Forschungsfrage bzw. das Thema des Reviews erwarten lassen (Heinemann et al., 2022). Dieses Vorgehen ist schon allein deshalb geboten, da nicht selten die in der Recherche verwendeten Begriffe weder im Titel noch im Abstract vorkommen, der Text aber dennoch eine Relevanz zur Beantwortung der Fragestellung haben kann (Bedenlier et al., 2020). Üblicherweise werden auf diese induktive Weise auch Schlagworte ermittelt, deren Vorkommen in Titel, Abstract oder Schlagworten zum Ausschluss des Titels aus der weiteren Auswertung führen. So entsteht im Laufe des Prozesses ein themenspezifisches Glossar an Ein- und Ausschlussbegriffen. Hier muss ggf. eine vorgeschaltete Recherche nach Synonymen und assoziierten Begriffen, in jedem Fall aber ausreichend Zeit für eine iterative Prüfung der vorgenommenen inhaltlichen bzw. begrifflichen Grenzziehungen des eigenen Auswertungsbereichs eingeplant werden. Letzteres kann auch eine Veränderung der Schwerpunktsetzung des Reviews notwendig machen, beispielsweise durch eine Verengung oder Erweiterung der ursprünglichen Fragestellung.

Je nach Thema und Bildungsbereich wurde in dieser Phase bereits ein erstes grobes Clustering anhand inhaltlicher Aspekte vorgenommen (z.B. EB/WB im Review zu Bildungsteilhabe: Kategorisierung nach benachteiligter Gruppe). Auch in diesem Schritt können wissenschaftliche Hilfskräfte unterstützend eingebunden werden, jedoch ist damit ein deutlich höherer Einarbeitungsaufwand verbunden, da sie inhaltlich gebrieft werden müssen.

Titel, die mit Blick auf den Bildungsbereich und das Thema bzw. die Fragestellung passend erschienen, wurden entsprechend markiert und zur Prüfung im Voll-

text vorgesehen. In dieser Phase zeigte sich bildungsbereichsübergreifend, dass insbesondere die durch die Recherche ermittelten Abstracts und Schlagworte der Einträge nicht selten ungenügende Informationen lieferten, um eine Studie eindeutig ein- oder auszuschließen. Zum Teil waren die Abstracts auch unvollständig oder fehlten komplett, wenn sie datenbankseitig nicht oder nicht vollständig erfasst waren – ein bekanntes Problem in der Datenbankrecherche (Glanville et al., 2019). In Einzelfällen traten zudem Diskrepanzen zwischen den von den Datenbanken ausgegebenen Daten und den in der DOI kodierten Angaben auf. Daher müssen entsprechende Ressourcen zur Nachrecherche der Abstracts und/oder Schlagworte eingeplant werden, um die Einträge in der Literaturverwaltung zu vervollständigen bzw. zu korrigieren. Nicht eindeutig auszuschließende Arbeiten müssen ebenfalls zur Volltextprüfung vorgesehen werden, um Fehlauschlüsse zu vermeiden. Je nach Fragestellung und Bildungsbereich muss dadurch eine beträchtliche Zeitmenge für die Prüfung der Texte eingeplant werden. Dabei hat die Erfahrung gezeigt, dass der Ertrag an verwertbaren Studien aus den nicht eindeutig ausschließbaren Arbeiten in einem sehr schlechten Verhältnis zum Prüfungsaufwand der entsprechenden Volltexte steht. Auch in dieser Phase ist der regelmäßige Austausch im Team wichtig, um beispielsweise die eigene Einschätzung eines Titels im Zweifelsfall kommunikativ zu validieren und im Screening der Abstracts aufkommende Fragen zur Bewertung von inhaltlichen Aspekten zu klären, die möglicherweise eine grundlegende kommunikative Validierung im Team notwendig machen. Idealerweise entsteht so in einem gemeinsamen iterativen Prozess eine Instanz zur personenübergreifenden Absicherung des methodischen Vorgehens. Um die Abstimmungsnotwendigkeit im Verlauf des Abstract-Screenings zu reduzieren, hat es sich vor allem bei mehr als zwei Bearbeitenden als sinnvoll erwiesen, auf der Basis einer begrenzten Zahl von Titeln aus dem Korpus, die von allen gelesen und gemeinsam eingeordnet werden, einen Grundkonsens für die Beurteilung bestimmter Aspekte zu etablieren. Im Zusammenhang mit der kommunikativen Validierung, der Bewertung der Abstracts und der Sicherung der Nachvollziehbarkeit von Entscheidungen haben sich in diesem Projekt vor allem zwei Citavi-Funktionen als äußerst nützlich erwiesen: Zum einen das Verfassen von kurzen Kommentaren, insbesondere bei Titeln, deren Einordnung nicht unmittelbar augenscheinlich ist, zum anderen das Vergeben von Aufgaben mit entsprechender Betitelung (z. B. „kommunikative Validierung“) und dem Bearbeitungsstatus. Letztere erlauben das Wiederauffinden der Titel über die Citavi-Suchfunktion für die gemeinsame Abstimmung im Team.

4.2 Screening der Volltexte

Das Screening der Volltexte schließt sich an das Screening der Abstracts an. Die in diesem Kapitel vertretenen Sektoren hatten dabei jeweils verschieden große und heterogene Literaturpakete zu bearbeiten. Am Ende des Volltextscreenings wurde

deutlich, welche Studien in das Review mit aufgenommen und welche doch noch ausgeschlossen werden. Zunächst war es hilfreich, die für das Volltextscreening berücksichtigten Studien jeweils als erstes nach bestimmten Formalia zu prüfen (Porritt et al., 2014). Aufgrund der oben genannten Schwierigkeiten in der Abstractprüfung, wenn beispielsweise der Dokumententyp noch nicht ersichtlich war oder das Publikationsjahr nicht korrekt vom Literaturverwaltungsprogramm übernommen wurde, können sich auch formale Abfragen bis in den Folgeschritt des Volltextscreenings ziehen. Somit kann es passieren, dass Dokumente zwar durch das Abstractscreening kommen, aber gleich zu Beginn der Volltextprüfung aufgrund formaler Kriterien ausgeschlossen werden. An diesem Punkt sei aber auch auf die anhaltende kritische Diskussion über vermeintlich klare und objektive Kriterien und Entscheidungen im Reviewverfahren verwiesen (Wetterich & Plänitz, 2021; Wilmers et al., 2022), weshalb unter anderem auf das Kriterium vom Zitationsindex verzichtet wurde.

Der Arbeitsschritt des Volltextscreenings lässt sich nicht in jedem Fall scharf von anderen Arbeitsschritten trennen, da man sich v. a. in den hier vorliegenden Fällen des Critical Reviews und der damit verbundenen offenen Forschungsfrage (Sutcliffe et al., 2017) oftmals erst im Verlauf des Kodierens Sicherheit darüber verschaffen kann, ob ein Text dazu geeignet ist, in das Review aufgenommen zu werden oder nicht. So zeigte sich in der Erstellung der Reviews der Autor:innen, dass sich das Kodieren der Volltexte bereits dazu eignet, auch Textpassagen über die kodierten Studien zu verfassen. Dementsprechend verliefen die Volltextprüfung und die Analyse der Texte in vielen Fällen parallel und nicht unbedingt oder ausschließlich nacheinander. Dieses Vorgehen birgt jedoch den Nachteil, dass es dazu kommen kann, dass die Autor:innen stellenweise bereits weit im Analyse- und Leseprozess fortgeschritten waren und dann erst spät einen Aspekt identifizierten, der zum Ausschluss der Studie führte. Mit diesem Risiko hat bspw. der Sektor BKJF bereits während des Lesens der Volltexte gleichzeitig die Kodierung vorgenommen, die Literatur nach Themen sortiert sowie erste Zusammenfassungen der relevanten Inhalte verfasst. So konnte der gedankliche Fokus einzig auf einen Text gelenkt werden, ohne von anderen Arbeitsschritten unterbrochen zu werden, auch wenn dies in manchen Fällen zu einem ungenutzten Mehraufwand geführt hat.

Ein essentieller Entscheidungsschritt im Volltextscreening zeigte sich vermehrt bei der Grundsatzfrage, ab wann der Inhalt einer Studie als ausreichend relevant für das Critical Review eingestuft werden kann, denn auch wenn Publikationen den gewünschten Forschungsgegenstand eines Reviews aufgreifen, ist damit noch nicht garantiert, dass sie auch für das Ziel des Reviews relevant sind (Heinemann et al., 2023; Köstler, 2023; Liabo et al., 2017; s. Abschnitt 4.3 zum Kodierschema). In den Fokus des Reviews können potenziell sowohl die zugrundeliegende Theorie der Publikation als auch die Methodik und/oder die Ergebnisse rücken. Hierdurch werden das Kodierschema und die dort aufgeführten Kriterien zur Relevanzbestimmung bezüglich der Forschungsfrage des Reviews besonders wichtig. Auch das Kriterium der „Gene-

ralisierbarkeit“ erscheint demnach als ein wichtiger Grundsatz im Kodierschema, um die Frage nach der Relevanz für das Review immer wieder reflektieren zu können. Vor allem im Sektor Schule war dieses Kriterium nicht immer einfach zu entscheiden, da die Forschung typischerweise aus verschiedenen Fachdisziplinen kommt und damit mitunter auch verschiedene Lernkontexte abbildet. Die Frage der Vergleichbarkeit im Sinne einer sinnvollen Synthese war hier ein häufiger Fokus der gemeinsamen Kommunikation im Reviewteam (Heinemann et al., 2023).

Eine solche Abwägung wird erfahrungsgemäß auch bei derselben Person unterschiedlich ausfallen, je mehr sie sich mit der Literatur befasst hat und je mehr Expertenwissen sie zu dem besagten Zeitpunkt bereits aufbauen konnte (Buntins, Mulders & Schroeder, 2023; Polanin et al., 2019). Dementsprechend ist es ebenfalls möglich, dass Entscheidungsprozesse, die zum Ein- oder Ausschluss von Literatur führen, im Laufe des Arbeitsprozesses mit anderen Gewichtungen oder Einschätzungen vollzogen werden. Aufgrund dieser Veränderungen des eigenen Wissens erschienen die Zusammenarbeit im Team und damit verbundene Feedbackschleifen (ebenso) sehr hilfreich. Als besonders nützlich wurde die methodisch vorgesehene Dokumentation empfunden, da hierdurch Transparenz zwischen verschiedenen Arbeitsschritten und Mitgliedern im Team auf einfachem Wege hergestellt werden konnte. Entscheidungen im Bearbeitungsprozess konnten so zu jeder Zeit von allen Teammitgliedern nachvollzogen werden.

4.3 Kodierschema

Das Projektteam entwickelte ein einheitliches, übergeordnetes Kodierschema für die Beschreibung und Einordnung der Volltexte (siehe Tab. 1). Das Kodierschema wurde dabei sowohl zwischen den veröffentlichten Reviewbänden als auch innerhalb der Bildungsbereiche weiter ausdifferenziert sowie punktuell und thematisch passend um einzelne Kategorien erweitert. Die Kodierung der Studien erfolgte im grundlegenden Schema nach den drei Bereichen der Beschreibung der Publikationen, der Ziele der Publikationen sowie der Einordnung des methodischen Vorgehens der Publikationen. Eine genaue Beschreibung des Kodierschemas ist bei Wilmers et al. (2022) vorzufinden.

Der generelle Ansatz des Kodierens erfolgte im Sinne des Kodierspektrums nach Sutcliffe et al. (2017) in einer Kombination aus diesen deduktiv festgelegten Codes und weiteren induktiv in jedem Bildungsbereich und jeweiligem Review neu erschlossenen Codes, die aus der gesichteten Literatur heraus entstanden. So wurde ausgehend von der Methode des Critical Reviews auf theorie- und modellerweiternde bzw. -analysierende offene Forschungsfragen hingearbeitet.

Tabelle 1: Übergeordnetes Kodierschema für alle Reviews aus dem Projekt (2020–2023)

Beschreibung der Studien	Inhalte der Studien	Einordnung des methodischen Vorgehens
<ul style="list-style-type: none"> • Publikationstyp (z. B. Journalartikel, Sammelbandbeitrag) • Literaturbeschaffung (z. B. Online Volltext, Autor:innenkontakt) • Art der Studie (z. B. quantitativ, qualitativ, Mixed Methods) • Beschreibung des Forschungsdesigns (z. B. Stichprobengröße; Zeitpunkt der Datenerhebung) • Geografischer Fokus (z. B. Land) • Kontext (z. B. Passung zum Bildungssektor) 	<ul style="list-style-type: none"> • Fragestellung / Hypothese / Forschungsziel • Ergebnisse der Studie • Untersuchungsgegenstände 	<ul style="list-style-type: none"> • Orientierung an wissenschaftlichen Kriterien (z. B. Validität, Reliabilität) • Relevanz für die Beantwortung der Forschungsfrage des Reviews

In den Bereich „Beschreibung der Studien“ (Tab. 1) fallen der Publikationstyp, die Art der Literaturbeschaffung, die Art der Studie, die Beschreibung des Forschungsdesigns, der geografische Fokus sowie der Kontext der vorliegenden Studie. Diese einfachen Deskriptoren von Publikationen dienen v. a. der Übersicht über das Forschungsfeld, z. B. ob eine Forschungsfrage in bestimmten Regionen besonders stark beforscht wird oder ob bestimmte Forschungsdesigns über andere dominieren. Im zweiten Block, „Inhalte der Studien“ (Tab. 1), des Kodierschemas wurden die Fragestellung und die Hypothesen der kodierten Studien einschließlich der Ergebnisse in Bezug auf die Fragestellung und die für das Critical Review relevanten Themen bzw. Untersuchungsgegenstände erfasst. Im letzten Block zur „Einordnung des methodischen Vorgehens“ (Tab. 1) wurde nach quantitativen, qualitativen und Mixed-Method-Studien sowie der Relevanz der Studien für die Beantwortung der Forschungsfrage des Critical Reviews kodiert.

Festzuhalten ist bei diesem grundlegenden Kodierschema, dass der erste Punkt recht klare Formalia umfasst, die ohne weitere Herausforderungen transparent kodiert werden können. Bei den Zielen der Studien und bei der Einordnung des methodischen Vorgehens wurde der Kodierprozess jedoch anspruchsvoller (Siddaway et al., 2019). Zunächst ist auch hier auf die Beurteilung von Relevanz zu verweisen, die nicht oder nur selten einem dualistischen Ja-Nein-Prinzip unterliegt und sich nicht nur aus einem Kriterium ergibt. Eine Einschätzung der Autor:innen ist immer nötig und unter Berücksichtigung des individuellen Bias macht dieser Punkt deutlich, dass auch bei einem transparenten Kodierschema die Replizierbarkeit eines Reviews herausfordernd sein kann (s. Prinzip des Bewusstseins für Subjektivität bei Sutcliffe et al., 2017). Zumal viele der eingeschlossenen Studien auch qualitative und offene Forschungsfragen bearbeiteten und demnach immer einzeln entschieden werden musste, inwiefern der Inhalt zu der Fragestellung des Reviews passt. Des Weiteren passierte es häufig, dass Studien das Thema des Reviews nicht als Hauptfokus be-

handeln, jedoch relevante Teilergebnisse für die Forschungsfrage des Reviews bieten (Heinemann et al., 2023). Für den Umgang mit dieser Gemengelage war v. a. der zweite Block des Kodierschemas hilfreich, in dem die Autor:innen die verschiedenen Forschungsgegenstände der Studie kategorisierten und beleuchteten.

Einzelne Subkategorien, z. B. bezüglich der Relevanz und thematischen Einordnung der Studien, blieben in der Beantwortung und Kodierung jedoch relativ offen, damit die Sektoren wie zuvor beschrieben in der Lage waren, diese je nach Fragestellung zu nutzen (Kitchenham & Charters, 2007; Shaffril et al., 2021) und mit weiteren Codes zu erweitern. So konnte abermals gewährleistet werden, dass die unterschiedlichen Perspektiven, die die verschiedenen Bildungsbereiche sowie die thematischen Ausrichtungen der einzelnen Synthesen mit sich bringen, Berücksichtigung finden. Beispielsweise wurde in einem Review des Bildungsbereichs Schule (Leber et al., 2023) eine weitere Subkategorie für verwendete digitale Tools in den Studien genutzt, da man in der Synthese sowohl die Methodik als auch die genutzten Tools im Lernsetting ausführlich darstellen wollte. Im Bildungsbereich Kindheit, Jugend und Familie wurde auf eine andere Art mit weiteren Subkategorien gearbeitet: Hier wurden wesentliche und wiederkehrende Themen (Oliver & Sutcliffe, 2012, S. 139; s.a. Wilmers et al., 2020) innerhalb der Studien festgehalten und am Ende schließlich als Grundlage für den thematischen Leitfaden der Synthese verwendet. In einem Review des Sektors Schule (Heinemann et al., 2022) stießen die Autor:innen bei einigen Studien im Volltextscreening auf ein Bezugsmodell, welches die Fragestellung des Reviews untermauerte. Die Kategorien des Modells wurden demnach induktiv zum Kodierschema hinzugefügt und als eine Grundlage für die Synthese unter einer offen angelegten Fragestellung (s. Kodierschema nach Sutcliffe et al., 2017, Abb. 1) herangezogen.

5. Erkenntnisse zum Screeningprozess

Die Methode des Critical Reviews im Kontext der Bildungswissenschaften benötigt weiteren Fokus, um eine klarere Vorgehensweise für Forschende aufzeigen zu können (Heinemann et al., 2023). Hierzu möchte das vorliegende Kapitel einen Beitrag leisten, indem es Möglichkeiten und Herausforderungen im Screening- und Kodierprozess von Literatur innerhalb eines Critical Reviews aufgezeigt hat. Dazu wurden Prozesse und Entscheidungsfindungen dargestellt sowie Lessons Learned beschrieben und reflektiert.

Generell lässt sich konstatieren, dass in der Methode des Critical Reviews das Ziel verfolgt wird, einen Überblick über aktuelle Forschungsbefunde zu einer bestimmten Forschungsfrage zu geben und dabei möglichst transparent zu arbeiten sowie (gegebenenfalls auftretenden) Bias zu reflektieren und methodisch zu minimieren (Buntins, Bedenlier et al., 2023; Higgins et al., 2023; Torgerson, 2006). In diesem Sinne wird beispielsweise die zuvor recherchierte Literatur zuerst nach bestimm-

ten, festgelegten Kriterien geprüft, dann (für die Analyse) ein- bzw. ausgeschlossen und letztendlich die eingeschlossene Literatur nach einem (möglichst) transparenten Schema kodiert. Critical Reviews besitzen aufgrund des Reviewformats, das zwar auf einer systematischen Erfassung des Forschungsfeldes beruht, aber keine strikte Qualitätsprüfung vorschreibt und gleichzeitig der kritischen Auswertung von Literatur Priorität einräumt, per Definition eine gewisse konzeptuelle Offenheit, die die Einordnung der Forschungserkenntnisse beeinflusst (Grant & Booth, 2009; Wright & Michailova, 2023).

Deutlich wurde, dass in allen hier fokussierten Arbeitsschritten im Reviewprozess teilweise ähnliche und damit wiederkehrende Herausforderungen bestehen. Deshalb erscheint es sehr sinnvoll, antizipierte Herausforderungen und Probleme so früh wie möglich im Autor:innenteam zu besprechen und mögliche Lösungsstrategien zu entwickeln. Idealerweise können so bereits in den vorgelagerten Erarbeitungsphasen wie etwa der Konzeption und Recherche Problemlagen vermieden werden, die sich ggf. auf das Screening auswirken (Buntins, Bedenlier et al., 2023). So zieht sich eine hohe Treffermenge in der Literaturrecherche, wie in allen hier zugrundeliegenden Reviews des Bildungssektors Schule, bis in den Kodierprozess. Begegnet wurde der Gemengelage z. B. mit formalen Kriterien, wie dem ausschließlichen Einbezug bestimmter Publikationstypen. Im weiteren Verlauf des Screenings und Kodierens unterstützten Bezugsmodelle und ein nachvollziehbares sowie transparentes Kodierschema.

Eine eher arbeitspraktische Limitierung beim Screening ist die Konzentrationsfähigkeit der involvierten Personen. So kann es bspw. aufgrund der Kürze der Texte beim Screening der Abstracts leicht passieren, dass die kognitive Belastung durch das Lesen und Beurteilen der Abstracts unterschätzt wird und zu viele Einträge geprüft und/oder zu wenige Pausen durch die prüfenden Personen eingelegt werden. Beides birgt Potenzial für Fehlentscheidungen im Screeningprozess, weshalb z. B. empfohlen wird, die Ein- und Ausschlusskriterien nach Komplexität zu ordnen und einfache Fragen voranzustellen (Polanin et al., 2019). Zudem sollte der Workflow unbedingt ausreichend Pausen sowie ein Tagesmaximum geprüfter Titel beinhalten. Darüber hinaus sollte unter Berücksichtigung der individuellen Ressourcen eine sinnvolle Verteilung der zu prüfenden Titel auf die bearbeitenden Personen gewährleistet werden. Ebenso konnte es sich als hilfreich herausstellen, mit mehreren Mitarbeitenden zu screenen und zu kodieren. Eine andere Möglichkeit für den Umgang mit einer hohen Trefferzahl wäre beispielsweise eine randomisierte Auswahl an Publikationen in das Screening und Kodieren aufzunehmen (Bedenlier et al., 2020).

Es wird deutlich, dass vor allem durch die hier vorliegenden offenen Forschungsfragen der Critical Reviews und dem sowieso schon heterogenen Kontext in der Bildungsforschung eine Herausforderung in der klaren Entscheidungsfindung zum Einchluss und Ausschluss der Literatur bestehen kann (Hammersley, 2020). Zudem stellt sich die Auswahl quantitativer und qualitativer Studien zu einer eigenen Forschungsfrage mitunter als herausfordernd dar, da die Komplexität z. B. durch qualitativ angelegte Studien in der Zusammenführung der Studien nicht immer einfach

für die eigene Forschungsfrage herunterzubrechen ist (Higgins et al., 2023; Petticrew, 2015). Dem lässt sich mit einer guten Dokumentation sowie einem deduktiv angelegten Kodierschema, welches in jedem Bildungsbereich und Review individuell mit induktiven Kategorien aus der Primärliteratur ergänzt werden kann, begegnen. Generell lässt sich der Aufwand, der im Screening und Kodierprozess geleistet werden muss, nur schwer vorab planen bzw. einschätzen. Hierzu können verschiedene Faktoren, wie die unterschiedliche Beforschung und die Quantität vorhandener Studien in den verschiedenen Bildungssektoren, oder die gewählte Fragestellung und deren Eingrenzung durch bereits vorhandene Expertise zu Beginn des Reviews beitragen.

Fasst man die Erfahrungen der Autor:innen zusammen, so haben Critical Reviews als Methode der Synthese von Forschungsergebnissen sowohl Vor- als auch Nachteile: Durch die im Vergleich zu Systematic Reviews (Grant & Booth, 2009) tendenziell offenere Fragestellung und größere Offenheit hinsichtlich der qualitativen Bewertung der Studien in Verbindung mit dem hochdynamischen Thema Digitalisierung ist die Replizierbarkeit der Ergebnisse eingeschränkt. Andererseits eröffnet sich dadurch die Möglichkeit ein Thema explorativer zu erschließen, was sonst eher Scoping Reviews vorbehalten ist. Rückblickend erscheint die Methode des Critical Review als geeignet um mit einem systematischen und transparenten Vorgehen Fragestellungen zu bearbeiten, bei der die Empirie erwartbar heterogen oder wenig gesättigt ist.

Literatur

Die durchgeführten Reviews, auf die sich diese methodische Reflexion bezieht, werden mit einem * markiert.

- Akers, J., Aguiar-Ibáñez, R. & Baba-Akbari, A. (2009). *Systematic reviews: CRD's guidance for undertaking reviews in healthcare* (3. Aufl.). York Publ. Services.
- Bedenlier, S., Bond, M., Buntins, K., Zawacki-Richter, O. & Kerres, M. (2020). Learning by doing? Reflections on conducting a systematic review in the field of educational technology. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond & K. Buntins (Hrsg.), *Systematic reviews in educational research* (S. 111–127). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27602-7_7
- Buntins, K., Bedenlier, S., Marín, V., Händel, M. & Bond, M. (2023). Methodological approaches to evidence synthesis in educational technology: A tertiary systematic mapping review. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 167–191. <https://doi.org/10.21240/mpaed/54/2023.12.20.X>
- Buntins, K., Mulders, M. & Schroeder, N. (2023). Kognitive Prozesse in Screening-Prozessen – Suchstrategien in der Bildungstechnologieforschung. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 103–124. <https://doi.org/10.21240/mpaed/54/2023.11.19.X>
- Diekmann, D. (2022). *Praxisbericht: Mit Metavorhaben die Kommunikation und Kooperation zwischen Bildungsforschung und Bildungspraxis unterstützen*. Vorab-Onlinepublikation. <https://doi.org/10.25539/bildungsforschung.v0i2.899>.

- Evans, J. & Benefield, P. (2001). Systematic Reviews of Educational Research: Does the medical model fit? *British Educational Research Journal*, 27(5), 527–541. <https://doi.org/10.1080/01411920120095717>
- Glanville, J., Cooper, H. M., Hedges, L. V. & Valentine, J. C. (2019). Searching bibliographic databases. *The handbook of research synthesis and meta-analysis*, 73–100. <https://doi.org/10.7758/9781610448864.8>
- Gough, D., Oliver, S. & Thomas, J. (Hrsg.). (2017). *An introduction to systematic reviews* (2. Aufl.). Sage.
- Grant, M. J. & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health information and libraries journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Hähn, K. & Ratermann-Busse, M. (2020). Digitale Medien in der Berufsbildung – Eine Herausforderung für Lehrkräfte und Ausbildungspersonal? In A. Wilmers, C. Anda, C. Keller & M. Rittberger (Hrsg.), *Bildung im digitalen Wandel: Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (S. 129–158). Waxmann. <https://doi.org/10.31244/9783830991991.05>
- Hallinger, P. (2013). A conceptual framework for systematic reviews of research in educational leadership and management. *Journal of Educational Administration*, 51(2), 126–149. <https://doi.org/10.1108/09578231311304670>
- Hammersley, M. (2020). Reflections on the methodological approach of systematic reviews. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond & K. Buntins (Hrsg.), *Systematic reviews in educational research* (S. 23–39). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27602-7_2
- Harden, A. & Gough, D. (2012). Quality and relevance appraisal. In D. Gough, S. Oliver & J. Thomas (Hrsg.), *An Introduction to systematic reviews* (S. 153–178). Sage.
- Heinemann, A., Koschorreck, J., Wilmers, A., Sander, P. & Leber, J. (2023). Potenziale und Herausforderungen von Forschungssynthesen im Kontext der Digitalisierung im Bildungsbereich am Beispiel des Critical Reviews. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 79–102. <https://doi.org/10.21240/mpaed/54/2023.08.13.X>
- *Heinemann, A., Leber, J., Sander, P. & Ahmad, C. (2022). Strategien und Maßnahmen für die schulische Curriculumentwicklung im Kontext der Digitalisierung. Ein Critical Review zur vergleichenden Analyse politischer Initiativen. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (S. 69–102). Waxmann. <https://doi.org/10.31244/9783830996224.03>
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J. & Welch, V. A. (2023). *Cochrane Handbook for Systematic Reviews of Interventions*. www.training.cochrane.org/handbook
- Kerres, M., Sander, P. & Waffner, B. (2022). *Zum Zusammenwirken von Bildungsforschung und Bildungspraxis: Gestaltungsorientierte Bildungsforschung als Ko-Konstruktion*. Vorab-Onlinepublikation. <https://doi.org/10.25539/BILDUNGSFORSCHUN.V0I2.935>
- Kitchenham, B. A. & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Technical Report EBSE-2007-01). Keele University.
- *Koschorreck, J. & Gundermann, A. (2020). Die Implikationen der Digitalisierung für das Lehrpersonal in der Erwachsenen- und Weiterbildung: Ein Review ausgewählter empirischer Ergebnisse und weiterer theoriebildender Literatur. In A. Wilmers, C. Anda, C. Keller & M. Rittberger (Hrsg.), *Bildung im digitalen Wandel: Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (S. 159–193). Waxmann. <https://doi.org/10.31244/9783830991991.06>

- *Koschorreck, J. & Gundermann, A. (2021). Die Bedeutung der Digitalisierung für das Management von Weiterbildungsorganisationen. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen* (S. 161–192). Waxmann. <https://doi.org/10.31244/9783830994558.06>
- *Koschorreck, J. & Gundermann, A. (2022). Orchestrierung von Blended Learning in der Erwachsenen- und Weiterbildung: Ein Critical Review. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (S. 161–186). Waxmann. <https://doi.org/10.31244/9783830996224.06>
- *Koschorreck, J. & Gundermann, A. (2023). Bildung mit digitalen Medien für benachteiligte Erwachsene – Chancen und Herausforderungen. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (S. 171–201). Waxmann. <https://doi.org/10.31244/9783830998464.06>
- Köstler, V. (2023). Zwischen Präzision und Sensitivität. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 1–27. <https://doi.org/10.21240/mpaed/54/2023.08.10.X>
- *Leber, J., Heinemann, A., Sander, P., Ahmad, C., Meneghinello, L. & Peters, L. (2023). Informationskompetenz von Schüler:innen stärken: Wie kann man das Prüfen von Informationen durch den Einsatz digitaler Medien fördern? In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (S. 75–106). Waxmann. <https://doi.org/10.31244/9783830998464.03>
- Liabo, K., Gough, D. & Harden, A. (2017). Developing justifiable evidence claims. In D. Gough, S. Oliver & J. Thomas (Hrsg.), *An introduction to systematic reviews* (2. Aufl., S. 252–277). Sage.
- Newman, M. & Gough, D. (2020). Systematic reviews in educational research: Methodology, perspectives and application. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond & K. Buntins (Hrsg.), *Systematic reviews in educational research* (S. 3–22). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27602-7_1
- *Nieding, I. (2022). Von digitalen Lernspielen bis Social Media. Wie Kinder und Jugendliche in non-formalen Bildungskontexten mit, von und über digitale Medien lernen. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (S. 31–67). Waxmann. <https://doi.org/10.31244/9783830996224.02>
- *Nieding, I. (2023). Teilhabe durch digitale Medien im häuslichen Umfeld. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (S. 33–74). Waxmann. <https://doi.org/10.31244/9783830998464.02>
- *Nieding, I. & Klaudy, E. K. (2020). Digitalisierung in der frühen Bildung: Der Umgang mit digitalen Medien im Spannungsfeld zwischen Schutzraum und Schlüsselkompetenz. In A. Wilmers, C. Anda, C. Keller & M. Rittberger (Hrsg.), *Bildung im digitalen Wandel: Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (S. 31–56). Waxmann. <https://doi.org/10.31244/9783830991991.02>
- *Nieding, I. & Klaudy, E. K. (2021). Die Umsetzung von Digitalisierung in Organisationen der non-formalen Bildung. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen* (S. 33–66). Waxmann. <https://doi.org/10.31244/9783830994558.02>
- Oliver, S. & Sutcliffe, K. (2012). Describing and analysing studies. In D. Gough, S. Oliver & J. Thomas (Hrsg.), *An introduction to systematic reviews* (S. 135–152). Sage.

- Petticrew, M. (2015). Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. *Systematic reviews*, 4, 36. <https://doi.org/10.1186/s13643-015-0027-1>
- Petticrew, M. & Roberts, H. (2006). *Systematic Reviews in the Social Sciences*. Wiley. <https://doi.org/10.1002/9780470754887>
- Polanin, J. R., Pigott, T. D., Espelage, D. L. & Grotzinger, J. K. (2019). Best practice guidelines for abstract screening large-evidence systematic reviews and meta-analyses. *Research Synthesis Methods*, 10(3), 330–342. <https://doi.org/10.1002/jrsm.1354>
- Porritt, K., Gomersall, J. & Lockwood, C. (2014). JBI's systematic reviews: Study selection and critical appraisal. *The American journal of nursing*, 114(6), 47–52. <https://doi.org/10.1097/01.NAJ.0000450430.97383.64>
- Shaffril, H. A. M., Samsuddin, S. F. & Samah, A. A. (2021). The ABC of systematic literature review: the basic methodological guidance for beginners. *Quality & Quantity*, 55(4), 1319–1346. <https://doi.org/10.1007/s11135-020-01059-6>
- Siddaway, A. P., Wood, A. M. & Hedges, L. V [Larry V.] (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology*, 70, 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Sutcliffe, K., Oliver, S. & Richardson, M. (2017). Describing and analysing studies. In D. Gough, S. Oliver & J. Thomas (Hrsg.), *An introduction to systematic reviews* (2. Aufl., S. 123–144). Sage.
- Torgerson, C. J. (2006). Publication bias: the achilles' heel of systematic reviews? *British Journal of Educational Studies*, 54(1), 89–102. <https://doi.org/10.1111/j.1467-8527.2006.00332.x>
- Viera, A. J. & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 5(37), 360–363.
- *Waffner, B. (2020). Unterrichtspraktiken, Erfahrungen und Einstellungen von Lehrpersonen zu digitalen Medien in der Schule. In A. Wilmers, C. Anda, C. Keller & M. Rittberger (Hrsg.), *Bildung im digitalen Wandel: Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (S. 57–102). Waxmann. <https://doi.org/10.31244/9783830991991.03>
- *Waffner, B. (2021). Schulentwicklung in der digital geprägten Welt: Strategien, Rahmenbedingungen und Implikationen für Schulleitungshandeln. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen* (S. 67–103). Waxmann. <https://doi.org/10.31244/9783830994558.03>
- Wetterich, C. & Plänitz, E. (2021). *Systematische Literaturanalysen in den Sozialwissenschaften: Eine praxisorientierte Einführung*. Verlag Barbara Budrich. <https://doi.org/10.3224/84742430>
- Wilmers, A., Anda, C., Keller, C., Kerres, M. & Getto, B. (2020). Reviews zur Bildung im digitalen Wandel: Eine Einführung in Kontext und Methodik. In A. Wilmers, C. Anda, C. Keller & M. Rittberger (Hrsg.), *Bildung im digitalen Wandel: Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (S. 7–29). Waxmann. <https://doi.org/10.31244/9783830991991.01>
- Wilmers, A., Keller, C., Achenbach, M. & Rittberger, M. (2022). Reviews zur Bildung im digitalen Wandel: Methodisches Vorgehen im Reviewprozess. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (S. 7–30). Waxmann. <https://doi.org/10.31244/9783830996224.01>
- Wilmers, A., Keller, C. & Anda, C. (2021). Reviews zur Organisationsentwicklung in Bildungseinrichtungen. Zur Methodik und Diskussion von Bias in Reviewverfahren der Bildungsforschung. In A. Wilmers, M. Achenbach & C. Keller (Hrsg.), *Bildung im di-*

gitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen. Waxmann. <https://doi.org/10.31244/9783830994558>

Wright, A. & Michailova, S. (2023). Critical literature reviews: A critique and actionable advice. *Management Learning*, 54(2), 177–197. <https://doi.org/10.1177/13505076211073961>