

What is Critical About a Critical Review? A Methodological Reflection on Conducting 20 Research Syntheses in Digital Education

Annika Wilmers^{1,2}, Pia Sander³, Jan Koschorreck⁴,
Carolyn Keller¹, Jens Leber³ & Anna Heinemann³

¹DIPF | Leibniz Institute for Research and Information in Education, Germany;

²Frankfurt University of Applied Sciences, Germany; ³University of Duisburg-Essen,

Germany; ⁴German Institute for Adult Education – Leibniz Centre for Lifelong Learning (DIE), Germany

ABSTRACT

This paper reflects on the use of critical reviews as a form of research synthesis in educational science. The analysis is based on 20 critical reviews conducted in the field of digital education including different areas of education, from early childhood education to adult education. By investigating the individual steps involved in a critical review, the paper explores what methodologically defines this type of review and to what extent standards can be applied to it. Additionally, it examines whether the approach aligns more closely with other review types, such as systematic reviews. The findings provide insights into what is critical about critical reviews, highlighting both the strengths and weaknesses associated with this review type. Critical reviews offer the potential to respond to heterogeneous research fields and dynamic subject areas in educational research. They also allow researchers to contextualise findings within a comprehensive literature search. On the other hand, challenges include a reduced emphasis on assessing the quality of individual studies and the difficulties in comparing literature. As a result of these findings, this paper concludes that critical reviews in the analysed examples proved to be particularly suitable for developing broader overviews in a field of educational research.

Keywords: *critical review, review type; educational science; digital education; methodological approach; research syntheses*

*Correspondence: Annika Wilmers, e-mail: a.wilmers@dipf.de

© 2025 Annika Wilmers, Pia Sander, Jan Koschorreck, Carolyn Keller, Jens Leber & Anna Heinemann. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: Wilmers, A., Sander, P., Koschorreck, J., Keller, C., Leber, J. & Heinemann, A. (2025). What is Critical About a Critical Review? A Methodological Reflection on Conducting 20 Research Syntheses in Digital Education. *Nordic Journal of Systematic Reviews in Education*, 3, 16–51. <http://doi.org/10.23865/njsre.v3.7059>

Highlights

- The review type *critical review* showed potential but also posed challenges. It proved useful to discuss specific topics within the wider research field of digitalisation in education.
- Potentials included the chance to manage heterogeneous educational research and integrate diverse findings into a cohesive synthesis using a narrative approach.
- Challenges lay in the imbalance between a comprehensive literature search and a limited quality assessment.
- Also, challenges shifted from one working step to another and influenced subsequent steps of the review process.
- The type of the review questions (broad or narrow) did not necessarily influence the size of the overall literature corpus or the set of included studies.
- The extent to which review questions referred to specific educational systems and country contexts or enabled relatively easy international transferability influenced both the size of the overall corpus and the number of included studies.

1. Introduction

The number of research syntheses in education has increased significantly over the past decade, reflecting their value in synthesising findings and providing systematic overviews on defined topics (Bedenlier et al., 2023; Haddon et al., 2023; Sutton et al., 2019; Zawacki-Richter et al., 2025). This trend highlights the need for methodological approaches that can accommodate the diversity and complexity of educational research (Zawacki-Richter et al., 2020). Among the various types of syntheses, *critical reviews* provide a flexible and reflective approach. They allow researchers to critically integrate and evaluate a broad range of research designs, literature types and topics while maintaining a narrative structure that identifies research gaps and future directions (Paré et al., 2015). Due to their reduced emphasis on formal quality assessment (Grant & Booth, 2009; Paré et al., 2015; Sutton et al., 2019), they allow for a quicker response to emerging topics and questions. However, critical reviews pose challenges and risks too, particularly in the processes of identifying, synthesising and critically discussing relevant literature (Heinemann et al., 2023). Despite the growing use of research syntheses, the methodology of critical reviews remains under-discussed and there is little discourse on critical reviews as a distinct methodological approach – especially for educational research, where rigid systematic review standards may not always align with the diversity of research questions and designs (Pigott et al., 2021). This is important in educational research as methodological standards are generally less established (Bedenlier et al., 2020) and clear procedures or hands-on experience are lacking for less common types of research syntheses such as critical reviews (cf. Haddon et al., 2023; Sutton et al., 2019).

This paper aims to address this gap by discussing critical reviews as a methodological approach in educational research. We examine the processes, challenges and opportunities, and thereby reflect on *what is critical about critical reviews*. Drawing on our experience in conducting 20 critical reviews, we investigate both typical characteristics as well as specificities of this review type. Furthermore, we provide strategies to guide researchers in decision-making during different phases of the review and that support a critical analysis of the literature.

While some of the findings presented here have been previously published in German,¹ they have been synthesised and translated specifically for this article. Other sections and investigations, especially those reflecting on the analysis and synthesis of included study results in critical reviews, have been uniquely designed and conducted for this publication, which is authored by six researchers from the review team that comprised up to 13 people throughout the project.

To ground our analysis, we reflect on insights from the research project Digi-EBF, which involved conducting critical reviews focused on digital education. Digital education serves as a particularly illustrative case for the methodological challenges and opportunities associated with synthesising research in a dynamically evolving and interdisciplinary domain. The surge in publications, combined with the often heterogeneous and fragmented nature of the evidence (Marín et al., 2023; Wilmers et al., 2023a), underscores the need for synthesis approaches that are both timely and conceptually appropriate. Moreover, the large number of strategy papers and frameworks related to digitalisation and education (e.g., European Commission, 2025; OECD, 2023; UNESCO, 2018) emphasises the political and academic urgency to conceptualise digitalisation and to advance the digital transformation of educational systems. At the same time, research in this field faces considerable pressure from a variety of stakeholders outside academia, who often demand guidance in dealing with the rapidly evolving technologies and associated instructional concepts (Terhart, 2021). Yet, educational research is still in the process of understanding how digitalisation unfolds in educational contexts. Critical reviews offer a promising means to synthesise and evaluate relevant individual studies and uncover research gaps: they have the potential to discuss controversies and contribute to conceptual innovations to support policy and practice (e.g., school development in the digital age (Waffner, 2021)). By critically analysing this methodological approach through the lens of digital education, this paper contributes to a deeper understanding of how critical reviews can be designed and conducted in education research, especially in fields marked by rapid change and conceptual uncertainty.

In the following sections, we describe the specific characteristics of critical reviews and introduce the conceptual and methodological foundation of our reflection. We outline the project context, the thematic scope and the rationale for selecting critical

¹ See results in Wilmers, 2024a; Wilmers & Keller, 2024 & Heinemann et al., 2024, in Wilmers (Ed.), 2024b.

reviews as an appropriate synthesis type within the project context. The section ends with the limitations related to our reflectional approach. The subsequent chapters detail the key steps involved in conducting a critical review. These include defining the search strategy, carrying out the screening and coding process, and synthesising and analysing the study results. Each section describes the theoretical foundations of the respective step, discusses important methodological considerations and highlights potential strategies. The paper concludes with a summary of the main insights derived from this process.

2. Specificities of critical reviews and research context

2.1 Characteristics of critical reviews

According to Grant and Booth (2009), a critical review follows a conceptual approach to explore a thematic field. In a defined thematic area, a systematic and exhaustive literature search is conducted resulting in a well-reasoned selection of literature. However, unlike a ‘classic’ systematic review, the assessment of literature is not as far-reaching. Rather, its objective is to gather important results, connect and place them in a larger scientific context to address an originally formulated research question. Owing to its conceptual design, critical reviews can contribute to building models, hypotheses and theories in a field (Grant & Booth, 2009). Therefore, broader or more openly formulated questions lend themselves to this format.

This definition positions critical reviews as a valuable tool for analysing and evaluating research findings at a specific point in time. In so doing, critical reviews support overall progress in a field. At the same time, the reduced emphasis on formal quality appraisal carries the risk of integrating study results that have not been subject to sufficient quality control. This can lead to subjective assessments and, consequently, the formulation of biased conclusions (Grant & Booth, 2009). Sutton et al. (2019) categorise critical reviews in the category of traditional reviews. These formats share a common feature in that they extensively explore the state of research in a field, usually with a narrative focus that also allows the integration of highly diverse research methods such as quantitative and qualitative research. Wright and Michailova (2023) further emphasise the relevance of critical discussions in critical reviews and remind authors of critical reviews that they should not lose sight of this aspect.

While research describes critical reviews as a separate type of review, as described above, there is still a lack of precise definition of its individual components and the associated work steps. Consequently, scholars may derive standards from systematic reviews, which are regarded as methodologically more complex. While it is helpful to use systematic review processes as points of orientation at each stage of a critical review, this approach requires extensive deliberation, which may lead to subjective and unsatisfactory decisions.

2.2 Foundation for methodological reflection: conducting 20 critical reviews

The reflections that constitute this article are based on the experience of working on four sets of critical reviews consisting of five critical reviews each (20 in total) over a period of five years (Wilmers et al., 2020, 2021, 2022a, 2023b). The project involved a research team of up to 13 persons (with some variations during the period of the project) and four institutional partners: one partner from a university and three partners from independent research institutes. Its overarching goals were to advance research on education and digitalisation while making this research visible, accessible and usable to the scientific community and practitioners. The critical reviews contributed to fulfilling these goals by synthesising and structuring current research on defined research questions related to digital education in five educational sectors. These sectors were ‘Education in Childhood, Youth and Family Contexts’, ‘School Education’, ‘Teacher Education’, ‘Vocational Education and Training’ and ‘Adult and Continuing Education’. In these sectors, the critical reviews summarise key findings and identify open research questions in four topic areas covering implications for educational staff, organisational development, learning and teaching, and participation in a digital world (see Appendix). Within the research team, one expert was responsible for conducting the literature search and the whole team met regularly to discuss challenges regarding both methodology and content. A senior coordinator facilitated these meetings to define common components in the methodology, review the methodological approach, reflect on individual processes and support each team member’s work.

The underlying database and all work steps in these processes are described in the following sections. The fact that the work process was repeated 20 times enables us to self-assess both the project’s overall contribution to advancing research on education and digitalisation and to identify methodological challenges and opportunities for critical reviews. This was possible due to the recurrent and transparent process, combined with the broad coverage of research topics and educational sectors (see Section 2.4 for a discussion of possible limitations). Such a reflectional approach may be relevant to similar research initiatives.

2.3 Linking project objectives, thematic scope and review type

In the project described here, we have deliberately selected the critical review as the chosen review type, as it met the requirements of the project scope and the thematic field of digitalisation. One of the project requirements was to produce annual research syntheses starting from the second year of the project. Such an approach must inevitably involve the implementation of measures to ensure the timely preparation of reviews at the working level. On the one hand, these measures concerned coordinating and dividing work steps within the team. Over time, individual work processes became increasingly intertwined. For example, new literature searches began during the evaluation phase of the previous reviews and knowledge transfer

following the publication of the reviews occurred in parallel with the processing of subsequent reviews.

On the other hand, steps were taken that moved the review process close to rapid reviews. According to Sutton et al. (2019), rapid reviews form a distinct review family of their own. While the overall aim of research syntheses is to inform practice, policy and administration, rapid reviews often have a special connection to at least one of these areas, which suggests shorter processing times (Thomas et al., 2013; Wollscheid & Tripney, 2021). In the Digi-EBF project, abbreviated procedures were implemented both in the literature search (e.g., limitation of search tactics due to close time frames) and during the screening process (e.g., limited quality assessments and no continuous double-checking). Some of these steps are defined by the review type (especially shorter quality assessments). However, a critical review should not be equated with a rapid review because other processing steps might be clearly more exhaustive. In the case of Digi-EBF, the project frame defined processing times invested per work step rather than the selected review type itself.

Aside from internal project requirements, the rapid and dynamic developments of certain topics might justify a shorter processing time. This is particularly relevant to the field of digitalisation in education, which is characterised by societal transformation processes, recognisable by fundamental and ongoing dynamic changes (Kollmorgen et al., 2015). In this context, assessments of developments are highly future-oriented (see Adloff & Neckel, 2019) and the consequences of these processes are more difficult to predict. Research in the context of digitalisation is also characterised by the comparatively high volume of literature continually being produced and often quickly published. This trend became apparent following the COVID-19 pandemic, with the rather rapid emergence of publications using easily accessible online formats and Open Access options.

For research synthesis teams, this means deciding whether to prioritise publishing findings quickly or spending time exhaustively exploring a complex thematic field. Critical reviews offer the advantage that their narrative formats allow reaction to dynamic developments and the integration of highly heterogeneous research (Heinemann et al., 2023). Both factors are highly relevant in the case of digital education, which is driven by innovations in digital technology. Heterogeneity, which can be principally found in educational research, results in this case from research dynamics but also from the thematic field's international relevance, while also being subject to different cultural contexts. In this thematic field, research syntheses opting for shorter processing times do not necessarily aim to deliver an all-encompassing conclusive product; instead, they present research results at a certain point in time within the constant flow of new research, thus enabling future research to build on these results. This corresponds to Grant and Booth's (2009) observation that critical reviews, with their focus on a critical stance, tend to serve as a basis for future observations.

2.4 Limitations of our reflectional approach

While conducting 20 critical reviews within the same project provides a valuable and rich base for methodological reflection, it also imposes certain limitations. One limitation concerns the diversity of the reviews, which may restrict their comparability. Diversity arose from two sources: first, the reviews were carried out by different sub-teams of authors with varied disciplinary backgrounds, resulting in different methodological decisions (e.g., concerning the degree to which qualitative studies were included). Although a roughly standardised procedure was agreed on (e.g., a general coding scheme), the reviews were not all conducted in exactly the same way. Second, the reviews address a broader range of educational sectors (e.g., early education, school education, adult education) and topic areas (see Table 2 in Section 3.3), which led to different starting points in terms of availability and design of research in the field (e.g., see the necessary variations in different search strategies in Section 3). Consequently, sector- and topic-specific methodological decisions were made and shaped the screening and analysis procedures (see Sections 4 and 5). While this diversity limits comparability, it also enables us to identify a broad range of methodological potentials and challenges in critical reviews. In addition, all reviews were linked due to a focus on digital education, which allows us to identify particularities related to this topic and which are not necessarily given in other subject areas.

Another limitation relates to the generalisability of our reflection. All 20 reviews stem from the same overarching project. Other research syntheses from the field of digital education or other topical areas were not considered, which limits the extent to which our insights apply to critical reviews more broadly. Furthermore, as this reflection focuses on critical reviews, it is not possible to make a systematic comparison with other review types, such as rapid reviews. Consequently, we only draw to a limited extent conclusions about the relative strength or weakness of critical reviews compared to other approaches.

Finally, critical reviews pose inherent reproducibility challenges. The narrative nature of the synthesis, combined with the inclusion of qualitative data and less standardised procedures, makes it difficult to fully reproduce the total of 20 critical reviews (see Section 5). The methodological decisions and chosen narrative formats shaped the findings discussed here and may lead to variations in the interpretation of results.

3. Searching the field: strategies for critical reviews

3.1. Requirements for critical reviews

The literature search aims to achieve a high level of retrievability of relevant literature. At the same time, the precision of the findings should be maximised regarding the research question (Köstler, 2023; McFadden et al., 2012; Stock, 2007; Taylor et al., 2007). A broader search increases the recall and thereby increases the chance of identifying a larger amount of relevant literature. Regarding the systematic and

transparent approach, the steps involved in database research are in many ways independent of the characteristics of the review type and should ideally be applied to all types of research syntheses. In the specific case of our research syntheses, the choice of the critical review methodology influenced the literature search process in that the search was based on a broad database search and additional search tactics. However, the search was not designed to be completely exhaustive (e.g., grey literature was only included selectively). Furthermore, in contrast to research syntheses conducted in medicine, which often focus on highly specific research questions, our reviews were usually based on broader research questions, such as the identification of research literature on developments in digital participation.

The interdisciplinarity of educational research necessitates searching across different databases and sources, as no central database indexes all relevant journals or completely covers all fields (Kugley et al., 2017, p. 19). Moreover, findings from studies can be published in various formats so that a selection of document types for the literature searches is not useful and, in most databases, not possible either. These conditions result in a broad search that can only be adjusted to enhance sensitivity by focusing on its specific content. As a result, the literature search in research syntheses in educational research needs to be individually modified to the research field and the research question. This approach is distinct from the methodology used in systematic reviews in medical science.

In our critical reviews, the research questions in relation to specific educational sectors defined the scope of the search. By adding a method chapter in each volume and by storing and publishing the methodological procedure of the search strategy and the reference lists resulting from the literature searches at the Research Data Centre for Education (Forschungsdatenzentrum Bildung - FDZ) at the DIPF | Leibniz Institute for Research and Information in Education, we aimed for a transparent presentation of both the results and processes.²

3.2 Objectives and structure of the literature searches

Development of the literature searches

In light of the specific requirements, search strategies were developed for a total of 20 different research questions. The development process for these search strategies was similar in each case and was characterised by a systematic and structured approach (in accordance with the recommendations of the most relevant guidelines, such as Booth et al., 2016; Gough et al., 2017; Higgins et al., 2023; Kugley et al., 2017). Creating an extensive search strategy and searching for literature requires sufficient time resources, which are often underestimated in the course of a review project (Brunton et al., 2017, p. 104). In our case, the total duration of the search process for one set of reviews (including five reviews) varied between four and ten months.

² Vol. 1: [10.7477/414:1:0](#); Vol. 2: [10.7477/414:2:0](#); Vol. 3: [10.7477/414:3:0](#); Vol. 4: [10.7477/414:4:0](#)

However, as described above, these search phases also overlapped with other steps in the review process. Over time, this overlap extended not only within each individual review, but also across the different sets of reviews being produced. This circumstance points to the importance of having one person concentrating uniquely on the search process. For each individual review, the literature search required a minimum of one-third of the total working time allocated for one review. In general, the recommendations for research syntheses are similar and emphasise the importance of a sufficient time frame for the literature search and calculate up to 40% of the total duration of a review project (Booth et al., 2016, p. 68).

A precisely prepared search strategy aims to avoid potential problems such as biased content or unsuitable search terms. At the beginning of each literature search, a collection of search terms and keywords was prepared in German along with their respective English translations. One of the challenges was to define suitable synonyms and search terms for the various topics covered in the reviews. In addition, for some search terms it was challenging to find precise translations. An example is the term *digitalisation*, which is commonly used in German as *Digitalisierung* and reflects the main focus of our reviews. In English research literature, this term was rarely found. Instead, additional terms such as technology or media resulted in more relevant results in the literature search. The selection of keywords was guided by the keyword index of the German database Fachportal Pädagogik [Education Research Portal] and the thesaurus of the Education Resources Information Center (ERIC) database. Another method for identifying appropriate search terms involved retrieving relevant references from the databases Web of Science (Social Citation Index), ERIC and the Fachportal Pädagogik and analysing their keywords within the database systems. The process of creating collections of search terms followed an iterative approach, as additional search terms were discovered during the search process when suitable publications were found and checked for keywords. Another challenge was that the project repeatedly encountered cases in which there was a lack of precise definitions for terms in the field. One example of this was the search for ‘cooperation of learning locations’ in Volume Two and for forms of ‘orchestration of learning models’ in Volume Three.

Selection of databases and sources

After collecting keywords, these were always tested in scoping searches at the beginning of each search process. On this basis, we could estimate the recall and precision of the search results as well as the appropriateness of the selected keywords. For the subsequent database search, each database was chosen based on its coverage of subject areas and its search options to support a systematic approach. The Fachportal Pädagogik and ERIC were used in every review. For the second set of reviews (Vol. 2) and all following reviews, Web of Science and Education Research Complete were added for all education sectors to retrieve literature not indexed in Fachportal

Table 1. Databases used in the research syntheses

Volumes with titles	Used databases sorted by educational sector	
Vol. 1: Meaning of digitalisation for educational staff and their training and further education	Childhood, Youth and Family	
	School Education	FPP, ERIC, DNB, GS = were used for all education sectors
	Teacher Education	
	VET	
	ACE	
Vol. 2: Organisational development regarding digitalisation	Childhood, Youth and Family	FPP, ERIC, LTL, ERC, WoS
	School Education	FPP, ERIC, LTL, ERC, WoS
	Teacher Education	FPP, ERIC, LTL, ERC, WoS
	VET	FPP, ERIC, LTL, ERC, WoS, BIBB/VET
	ACE	FPP, ERIC, LTL, ERC, WoS
Vol. 3: Teaching and learning with educational technologies	Childhood, Youth and Family	FPP, ERIC, ERC, WoS
	School Education	FPP, ERIC, LTL, ERC, WoS
	Teacher Education	FPP, ERIC, LTL, ERC, WoS
	VET	FPP, ERIC, LTL, ERC, WoS, SSOAR, SocioHub
	ACE	FPP, ERIC, LTL, ERC, WoS, BASE
Vol. 4: Impact of educational technologies for social learning and participation	Childhood, Youth and Family	FPP, ERIC, ERC, WoS
	School Education	FPP, ERIC, ERC, WoS, IEEE Xplore
	Teacher Education	FPP, ERIC, ERC, WoS, BASE
	VET	FPP, ERIC, ERC, WoS, BIBB/VET
	ACE	FPP, ERIC, ERC, WoS, IEEE Xplore

VET = Vocational Education and Training; ACE = Adult and Continuing Education; FPP = Fachportal Pädagogik [Education Research Portal]; ERIC = Education Resources Information Center; LTL = LearnTechLib, ERC = Education Research Complete; WoS = Web of Science Social Citation Index; BIBB/VET = VET Repository/Institute for Vocational Education and Training; SSOAR = Social Science Open Access Repository; SocioHub = Information Service for Sociology; IEEE Xplore = digital library by Institute of Electrical and Electronics Engineers; BASE = Bielefeld Academic Search Engine

Pädagogik or ERIC. The further selection of databases was guided iteratively by the thematic focus of the educational field and the specific research questions for each review. Table 1 provides an overview of all included databases.

Recognising that a significant amount of the indexed literature in bibliographic databases contains journal articles, a second search was conducted using additional search tactics to identify further references not available within these databases. These search tactics included, for instance, manual searching in relevant journals that were not indexed in the databases or the identification of relevant studies and publications from the websites of educational institutions or official organisations. In the educational sector of childhood, youth and family education, for example, the websites of welfare organisations were searched for publications that investigate informal education in a digital context. Due to the large amount of literature published in the field in recent years, the searched time frame ranged usually from the publication year of 2016 to the date when the literature search was performed. The search period in the first set of critical reviews included a longer time frame

starting with the year 2010. But it became apparent that mainly studies from more recent publication years were categorised as relevant. Consequently, in the second set of critical reviews in 2021, the last five years up to 2016 were searched. For the following critical reviews, the search period beginning in 2016 was continued. The search was conducted in both German and English and used subject headings and keywords.

3.3 Search results for a series of critical reviews: necessity of broad searches and advanced search tactics

Although the searches had clear thematic filters and covered a rather short search period, the results were relatively high for each search, even if some variations were observed. At the same time, the number of included studies was often relatively low, in many cases varying around one percent of all screened studies. Table 2 shows the total number of hits after removal of duplicates, as well as the final selection of included studies for each review after abstract and full text screening.

Table 2. Number of total and included studies for each synthesis

Educational sector	Childhood, Youth and Family	School Education	Teacher Education	Vocational Education & Training	Adult & Continuing Education
Volume I: Competences of educational staff					
Search results	1777	3380	2636	1620	2185
Included studies	10	122	15	14	41
Volume 2: Organisational development					
Search results	820	1554	1816	930	2439
Included studies	10	56	24	10	23
Volume 3: Teaching and learning with educational technologies					
Search results	3625	2444	2844	2442	2374
Included studies	33	29	38	6	8
Volume 4: Social learning and participation					
Search results	3045	4134	2929	2491	3321
Included studies	32	22	23	19	35

In principle, it is not surprising that the number of total records was relatively high, even with an approach that applied clear criteria. In contrast to a research synthesis on a precisely defined medical research topic and with defined research designs, a very broad search was initially required for the topics presented here to capture as many relevant fields as possible using this strategy (White, 2019, p. 57). This was often due to the lack of precise definitions of terms in the field or the use of different theoretical constructs in the literature. As a result, some of the search processes were difficult to plan at the beginning (see also Glanville, 2019, p. 76).

Educational research is not only highly interdisciplinary; the topics covered in the reviews also involve different approaches and contexts, which are reflected in the different terms and combinations of terms used in the searches. However, many findings, such as the appearance of unsuitable hits, only became apparent once all possible approaches to the subject area had been tested. The situation was similar with certain inappropriate synonyms, which had to be excluded from the total recall as far as possible by using further combinations of terms. A further challenge in the context of a broad search was that an open search, regarding study designs and publication formats, yielded a vast amount of literature that could fit in terms of content. However, these results did not necessarily reflect 1:1 in specific publication formats. For example, certain aspects often appeared as subordinate topics within a publication, but were deemed so relevant as to be included initially with the intention of reconsidering them later.

Furthermore, a persistent challenge was that the desired perspective on an educational sector or subject could not always be clearly covered in the search. Even if the education sector and the relevant group of individuals were theoretically well-defined (e.g., teacher, student, system level, etc.), there were overlaps in the literature between sectors (between youth and school or between further and higher education or between school and teacher education) and literature was retrieved that does not reflect the actual focus of the review (e.g., the perspective of teacher trainers). Incorrect allocations occurred despite the use of specific search terms such as school, day care centre or vocational school. As a result, such misallocations were shifted to the screening of the literature and had to be filtered in this step, which proved to be a time-consuming task.

In many cases, the keywording in the literature made a broad search necessary. The imprecise fit of keywords could stem from either inappropriate keywording for the research question of the review or the fact that a subject might already have been represented by different terms within a single language – a problem exacerbated by the translation of terms (see Bedenlier et al., 2020; Buntins et al., 2023a; Jäger-Dengler-Harles & Rittberger, 2022). Consequently, this meant that it was assumed as a precautionary measure that a search based only on key terms could not generate enough hits (see also O'Mara-Eves et al., 2013).

All in all, the comprehensive database searches, the combination of different databases in the field of education, the use of additional search tactics and the dynamic modifications made during the search process improved the recall and precision. Additional search tactics, such as checking the publications of individual authors or the manual search in (current) journal volumes, which were often highly time-intensive, helped to improve the precision. The results show indeed that a significant number of studies was identified via advanced search tactics, in total 99 studies out of the 570 included studies (see Table 3).

Based on the results and data of the 20 critical reviews, supplementary examinations analysed the efficiency of each of the used databases by searching the included

Table 3. Number of studies included via additional search tactics

Issues	Manual search	Review of publication lists of authors	Web search	Manually added publications	Added in total/ vol.
Volume 1:	4	-	-	28	32
Volume 2:	0	7	0	20	27
Volume 3:	0	0	0	5	5
Volume 4:	4	6	4	21	35

- = search tactic not applied; 0 = no further hits identified

570 studies of the respective critical reviews in these databases (Heck et al., 2024; Keller & Heck, 2024). As a result, 87 of the 99 studies that were originally found using advanced search tactics were also indexed in the used databases but were not identified via the database searches. In most cases, this was due to the insufficient indexing of titles. Therefore, additional search strategies proved to be useful. In general, the majority of the included studies were available in several of the databases, but none were indexed in all of the databases and only 12 studies were not available in any of the databases. However, since some studies were indexed in only one database and could not always be retrieved using the chosen search terms, searching across multiple sources also proved useful. Indeed, the analysis shows that a combination of national and international databases is highly recommended to maximise the coverage. For example, the highest coverage, at 98%, was identified for the third volume with an ideal combination of the databases ERIC / FIS / WoS / ERC / LTL / DNB. In contrast, the highest coverage, at only 89% for Volume 2, resulted from the new combination of ERIC / FIS / WoS / ERC / DNB. These examples demonstrate that the suitability of databases varies according to the specific content and research question. Therefore, each search strategy should always be individually tailored. Regarding the use of single databases, Google Scholar theoretically achieved the highest coverage in all analyses but nevertheless proved to be less suitable for systematic literature searches (Heck et al., 2024). As a result, it was excluded from the database searches after Volume 1.

To sum up, the iterative search procedure, which combined database searches and advanced search techniques, became more differentiated with each cycle of reviews to both cover the subject area and identify relevant literature. The nature of the search depended on the format of the critical review, such as the comparatively open question, but was also characterised by the subject area of digital education, which is marked by a high degree of dynamism and heterogeneity.

Some challenges of the review process already became apparent during the literature search but could not yet be addressed during this step. By aiming for a high recall, it was possible to ensure a broad coverage of the research question. However, this situation ultimately led to further challenges in the screening process. Conversely,

screening requirements also pointed to the need to cover the field as broadly as possible to enable the selection of appropriate literature. We therefore saw an interaction between searching and screening when dealing with the literature. From this experience, we can conclude: the more the search parameters were determined by content categories (as opposed to a selection based on criteria such as research design or publication format), the more filtering shifted to the following steps of the work process.

4. Finding the most relevant items: challenges and approaches in the screening and coding process

4.1 Theoretical insights for screening and coding the literature

The process of screening and coding the literature is a central element of each research synthesis. However, specific guidance in the literature for critical reviews is limited, with most activities typically adapted from procedures used in systematic reviews (e.g., Sutcliffe et al., 2017; Sutton et al., 2019). Screening and coding enable the identification and categorisation of relevant literature, considering its quality and usefulness in answering the research question (Gough et al., 2017; Porritt et al., 2014). In particular, in critical reviews ‘the emphasis is on the conceptual contribution of each item of included literature, not on formal quality assessment’ (Grant & Booth, 2009, p. 97). To identify relevant items, three main activities were conducted. These included: (a) the definition of selection criteria; (b) the screening of titles and abstracts as well as full texts; and (c) the coding and quality assessment of the included literature (Buntins et al., 2023b; Grant & Booth, 2009; Hallinger, 2013; Porritt et al., 2014). These activities did not necessarily follow a strict sequence but could occur in parallel or iteratively, based on new insights during the process (Heinemann et al., 2023).

Defining selection criteria

Decisive for defining the inclusion and exclusion criteria for selecting literature is the conceptual framework of a review, which includes its research question and scope (Newman & Gough, 2020). As the selection criteria derive from the search terms, this activity is closely related to the identification of search terms and keywords in the literature search (Haddon et al., 2023). However, to ensure a more reliable selection, it is usually necessary to adapt and specify the criteria. In general, some selection criteria will be straightforward and rather formal in nature (e.g., publication date) while others are strongly content-related, possibly fuzzy and rather complex – especially in educational science, where constructs often remain undefined (Zawacki-Richter et al., 2020). Examples of terms we analysed that are used inconsistently in the research literature include digital competences (e.g., Koschorreck & Gundermann, 2020, p. 167) and definitions of management (e.g., Koschorreck & Gundermann, 2021, p. 163). While there are no formal requirements for documenting and presenting selection criteria in critical reviews (cf. Grant & Booth, 2009), it is recommended

that the criteria are clearly defined before the selection process begins as this contributes to a transparent and more objective screening.

Screening

Next, the screening process was conducted in two distinct stages (Haddon et al., 2023; Newman & Gough, 2020). The first stage involved a primary assessment of titles and abstracts (abstract screening) with the aim of filtering out studies that did not meet the predefined criteria. Studies that could not be clearly categorised during abstract screening were passed on to the next stage for further investigation. The process effectively reduced the volume of literature, thereby limiting the more resource-intensive full-text screening to those studies that were considered potentially relevant. The full-text screening then allowed a thorough evaluation of the quality and relevance of the studies in relation to the review question. In fields such as educational research, where abstracts may not always adequately convey methodological details, this activity is essential (Hammersley, 2020). As such, the process of ‘critical appraisal’ becomes a crucial step in assessing the contribution of an item against selection criteria (Petticrew & Roberts, 2006; Sutton et al., 2019). Newman and Gough (2020) provide three fundamental pillars of critical appraisal in systematic reviews: detailed examination of study characteristics, assessment of study quality and relevance to the research question, and analysis of study results. While these pillars can also guide the full-text screening in critical reviews, the quality assessment is less formal and more emphasis lies on identifying the most significant items in the field regarding the research question (Grant & Booth, 2009).

Coding

In the final stage of the screening process, the coding scheme, which is often an intrinsic development of the previous activities, is finalised. This coding scheme helps to systematically organise the data in line with the research question and provides a framework for subsequent analysis (Wetterich & Plänitz, 2021). Sutcliffe et al. (2017) emphasise that an effective coding process in systematic reviews must adhere to three core principles: (1) determine coherent and relevant codes that align with the review question and extract relevant information; (2) clearly formulate reliable and explicit codes to ensure a common understanding within the team; and (3) be aware of subjectivity in coding. While coding is influenced by the review type (Wilmers et al., 2022b), these general principles may also apply to critical reviews. They emphasise the importance of iterative communication and collaboration within the research team to maintain rigour and minimise bias throughout the review process. By systematically categorising content – whether deductively from established theories (see Section 4.3) or inductively through the ongoing review of the literature – researchers create a robust basis for synthesising diverse evidence. This approach ultimately enhances the validity and reliability of the review’s

conclusions – despite the interpretative and subjective nature of critical reviews (Grant & Booth, 2009).

4.2 What is critical in screening and coding literature in educational science: facilitating transparency and reliability

The relevance of a systematic and transparent screening and coding process in critical reviews is evidenced by the low yield rate observed in our reviews, which rarely exceeds 1% (see Table 2). An illustrative example comes from the review on orchestration of blended learning in adult and continuing education (Koschorreck & Gundermann, 2022). In this review, more than 2,000 items were identified in the systematic search. However, after screening and coding, only 0.2% of these items met the selection criteria and were included in the final analysis. The specificity of the selection criteria that derived from the research questions is one reason for the enormous reduction in items. Consistent with other reviews in this project, the authors used general selection criteria derived from the search terms such as publication date (e.g., 2016–2021), educational sector (e.g., adult and continuing education) or concepts of interest (e.g., blended learning). Also, they specified some additional selection criteria such as the type of study design and the publication format. In this case, the criterion of identifying comparative experimental studies led to a large reduction in the number of items, as such studies were hard to find in adult and continuing education. An illustrative example of how search terms and selection criteria are defined, related, and yet differentiated according to their specific purpose in the review phase can be found in Capparozza and Kathmann's (2022) critical review. In general, we ensured the documentation of selection criteria, along with the specific rationale for each criterion to enhance the transparency and reproducibility of the review process (Booth et al., 2016). This approach served to mitigate bias and foster transparency across team members, enabling others to follow, verify or challenge the decisions made at any stage.

In addition to thorough documentation, we held training sessions and pilot tests among reviewers to calibrate the understanding and application of these criteria, ensuring the objectivity and consistency of the screening process (Buntins et al., 2023a; Viera & Garrett, 2005). During these sessions, any ambiguities or discrepancies in interpretation were addressed and consensus was reached on how to handle ambiguous cases. Furthermore, periodic checks of inter-rater reliability throughout the review process were helpful. In most cases, we achieved this by having multiple reviewers independently assess a subset of studies during abstract screening and full-text screening. The results of these assessments were then compared, and discrepancies were discussed and resolved collectively. When discrepancies were found, they served as a basis for further specifying the selection criteria or the coding scheme, ensuring that the process remained aligned with the research question and the review's objectives. Team collaboration and feedback loops proved particularly valuable as decision-making evolved with increased familiarity and expertise in

the literature (Buntins et al., 2023b; Polanin et al., 2019). Throughout the various research syntheses in the project, the value of collaboration between the information specialists and education scientists was particularly evident in the definition of relevant selection criteria (Bedenlier et al., 2020). While the information specialist was primarily responsible for determining the search string, her input on suitable search terms also proved invaluable in defining the inclusion and exclusion criteria. Likewise, the expertise of the researchers on a specific topic supported the information specialist in finding relevant search terms.

In the reviews considered here, we found it beneficial to start the screening process with rather simple formal and descriptive criteria and exclude studies immediately when these did not apply to these criteria – even without further examination of the content-based criteria. This made the process more efficient and prevented us from experiencing cognitive overload by balancing too many criteria at once. In the next step, we proceeded by evaluating items against content-based criteria and deciding whether a study's content was potentially relevant to the research question. The evaluation was usually much more complex and challenging. It was not guaranteed that publications addressing the topic of the review would be relevant to the specific aims of the review (Heinemann et al., 2023; Köstler, 2023). A review may concentrate on the underlying theory of a construct, the methodology employed to measure the construct and/or the specific findings related to the construct.

Deciding on relevance in all sectors was often a challenge due to the different research contexts they addressed, the disciplines involved and the associated differences in the conceptualisation and definition of terms. Similar to systematic reviews, this 'messiness' in educational research (Hammersley, 2020; Zawacki-Richter et al., 2020, p. xii) is also one of the major challenges in critical reviews. Therefore, it is not surprising that our experience indicates that the assessment of relevance did not adhere to a straightforward binary model. Instead, it entailed a degree of subjectivity, which introduced a challenge to the replicability of the process, even when clear selection criteria and a transparent coding scheme were employed. We experienced, however, that decision-making evolved with increased familiarity and expertise in the literature (Buntins et al., 2023b; Polanin et al., 2019), which also led to faster processing times. Nevertheless, as known from systematic reviews (Zawacki-Richter et al., 2020), screening and coding probably remain the most time-consuming activities in our reviews and working in a team helped to speed up the process. For example, it often became apparent that the abstracts and keywords retrieved during the search provided insufficient information to definitively include or exclude a study (see also Bedenlier et al., 2020; Haddon et al., 2023). In some cases, abstracts were incomplete or entirely missing – a well-documented issue in database research (Glanville, 2019). As a result, we allocated additional resources for rechecking abstracts and keywords to complete or correct the entries in the literature management system.

Working in a team necessitated agreement on a structured and transparent approach, which we tried to achieve through the implementation of established guidelines.

For example, pillars of critical appraisal, as described earlier (Newman & Gough, 2020), guided the systematic evaluation of studies during the full-text screening, ensuring that only methodologically sound and relevant studies contributed to the final syntheses. While these pillars were originally designed for quality assessment in systematic reviews, we prioritised the relevance of an item to answering the research question over rigorous quality checking (Grant & Booth, 2009). Additionally, we based the screening and coding process on Sutcliffe et al.'s (2017) theoretical process description. However, modifications were required to accommodate the open research questions in our reviews and the diversity of conceptualisations and methods in educational science (Haddon et al., 2023), which makes it particularly difficult to fully predefine coding categories. We found an approach resembling a qualitative content analysis strategy to be helpful, as it combined deductive and inductive categories for later coding, providing a balanced and flexible method of analysis that adapts to the evolving understanding of the literature (cf. Haddon et al., 2023). We developed a basic coding scheme for describing and categorising the items (Wilmers et al., 2022b). This scheme remained flexible and was refined across published review volumes and educational sectors as well as through inductive coding categories that emerged throughout the process (cf. Kitchenham & Charters, 2007; Shaffril et al., 2021).

As we dived deeper into the full texts, we encountered unexpected themes, methodological nuances and emerging patterns that were not anticipated during the initial stages. This demanded a dynamic interaction between deductive and inductive reasoning, allowing the review to evolve based on the data encountered. For instance, in their review on how digital media promotes students' information literacy, Leber et al. (2023) coded items addressing the initially determined concept of information literacy. Throughout the abstract screening they added 'critical thinking' and 'civic online reasoning' as separate coding categories as it became clear that these synonyms used in titles or abstracts suggested relevance to the research question and seemed to be worthy of differentiation. Through this iterative approach, the authors gradually refined the theoretical construct and the associated coding categories. Similarly, in the sector 'Education in Childhood, Youth and Family Contexts', recurring themes across the studies were noted and used to develop the thematic framework of the synthesis. By working with established guidelines while embracing an iterative process, the reviews remained responsive to the complexities of the literature, thereby enhancing its comprehensiveness and relevance (Bedenlier et al., 2020; Buntins et al., 2021). This blend of approaches, in our experience, is crucial for maintaining a balance between a manageable number of studies and comprehensive coverage of relevant literature. It also highlights that the activities during the screening and coding process often overlap – a phenomenon associated with the open research question in critical reviews (Sutcliffe et al., 2017). For example, we found that coding the full texts was already suitable for summarising studies for later synthesis. Accordingly, the authors

in the ‘Education in Childhood, Youth and Family’ sector carried out the coding, sorted the literature by topic and wrote initial summaries of the relevant content while reading the full texts. This approach allowed mental focus to remain solely on one text without being interrupted by other work steps, even if, in some cases, this work was then also carried out on studies that were later excluded.

To conclude, it became clear that the open research questions of critical reviews and the already heterogeneous context in educational research could pose a challenge in making clear decisions on the inclusion and exclusion of items (Hammersley, 2020). In addition, the selection of quantitative and qualitative studies for one’s own research question could sometimes be challenging, as the complexity of qualitative studies, for example, was not always easy to break down when merging the studies (Higgins et al., 2023; Petticrew, 2015). This can be countered with good documentation and a deductive coding scheme, which can be individually supplemented with inductive categories from the primary literature in each educational area and review. In general, the effort required for the screening and coding process is difficult to plan or estimate in advance. Influencing factors may be the diversity of research and the quantity of existing studies in the different educational sectors, or the selected research question and its limitation through existing expertise at the beginning of the review.

5. Extracting the gist: challenges of analysing and synthesising evidence from diverse research in education

5.1 Theoretical background on analysing data in critical reviews

The work steps of data analysis in research syntheses are described in different handbooks (e.g., Cooper et al., 2019; Higgins et al., 2023) and discussed in relation to specific preconditions and requirements stemming from the field of social science in general (e.g., Petticrew & Roberts, 2006) and educational science in particular (e.g., Gough et al., 2017). The main aim is to combine data from all included studies to draw conclusions from a body of evidence (McKenzie et al., 2023). Quality assessments evaluate both the relevance of the included studies in relation to the original review question and the quality of the applied methodology in the studies under consideration (Liabo et al., 2017). Depending on the research design of the included studies, this may, for example, concern questions of validity and generalisability or more content-related questions in the case of qualitative studies (see Petticrew & Roberts, 2006). Just as with the screening and coding process, the question arises as to which extent these standards, which are usually orientated towards systematic reviews, also apply to analysis steps in critical reviews. Specifically, the less formal quality assessment described earlier in critical reviews (Grant & Booth, 2009; Sutton et al., 2019) differs from other review types. However, regardless of the reduced quality control, a critical review must find a way to objectively assess literature findings so that their relevance to the review can be judged.

In this situation, what exactly characterises the analysis in critical reviews? In addition to using a narrative format, Sutton et al. (2019) emphasise the possibility of integrating different types of research (e.g., quantitative and qualitative research) and the ability to respond to current topics if this aligns with the goal of the review. In the reviews presented here, the focus was also less on the 1:1 comparability of study results and more on the critical discussion of these results and their categorisation in a larger research context. Wright and Michailova (2023) called for critical reviews to focus more on content, moving away from a pure presentation of results towards an explicit critical discussion of the literature examined, including a clear positioning of the respective authors of the reviews. This is also precisely where they see the added value of a critical review compared to a systematic review. In addition, Paré et al. (2015) emphasised that the primary advantage of critical reviews lies in their possibility of highlighting problems and contradictions in research and inadequate research contexts, thus advancing the development of the research field. In return, however, critical reviews must live with the weakness that they often contain subjective assessments (Paré et al., 2015, p. 189).

This intermediate position between more traditional and more systematic approaches of critical reviews (cf. Sutton et al., 2019) and consequently the opportunities and limitations of the analysis in our project are investigated in more detail below. To this end, we first retrospectively examine how the evidence from the 570 included studies across 20 research syntheses is presented and discussed in the critical review format. This includes, for example, the question of how often results could be verified multiple times and if and how often contradictions were discovered. Based on this approach, we aim to draw conclusions about the reliability of the analyses and their impact on the discussion of the research results.

5.2 A critical retrospective on 20 research analyses

Within the research syntheses of our project, findings are presented in a narrative text format that discusses the relevant results of the included studies and draws conclusions at a meta-level. These narrative passages are usually accompanied by tables indicating the nature and main characteristics of the included studies (e.g., the country context).

Procedure

To examine the analyses from all 20 reviews regarding characteristics and affordances, we applied a quantifying approach. While reading the sections of each review in which the selected studies and their respective results were presented, we noted all instances of multiple affirmations for a finding in a table document. Mentions of contradicting evidence were also noted. The statements included the number of referenced studies stating similar results (twofold verification, threefold, etc.), as well as the respective page number in the review. Additionally, we extracted the

research questions, the total number of selected studies and peculiarities or noteworthy features of the presentation or the presented studies.³ This process was conducted independently by three different authors of this publication, following the four-eyes-principle. Intercoder-correlation was computed for every pairing, which resulted in a kappa according to Brennan and Prediger of $k=0.72$ between coder a and b and $k=0.79$ between coder b and c (Brennan & Prediger, 1981). Still, agreements between coders varied considerably across some of the reviews. This may be due to the few multiple affirmations that could be identified in individual reviews and the narrated form of the findings, which resulted in challenges in the interpretation (see below). All individually recorded instances of multiple verifications and contradictions underwent subsequent communicative validation, through which all discrepancies between raters could be resolved. Counting references to the selected studies that accompany evidence-related statements seems, at first glance, to be a rather straightforward method of quantifying the robustness of the presented evidence on different aspects of digital education. However, we encountered several challenges in classifying the nature of the evidence in narrative text formats which are outlined in the subsequent paragraphs.

Challenges of quantifying weight of evidence via direct and indirect citations

Findings were predominantly presented in narrative form with varying degrees of interpretation by the authors rather than quoting numerical results from the statistical analysis, which were usually provided in tables. In two cases, the statistical values of the included studies were subjected to further computations to compare effects between studies in a basic statistical approach (Capparozza, 2021; Capparozza & Kathmann, 2023). We also regularly encountered instances where similar, but not identical, evidence was presented and, thus, the boundary between a critical and a mapping review became fluid. Sometimes, statements also allowed for a small margin of interpretation for the reader. We surmise that this is, on the one hand, a result of reviewing a heterogeneous body of research with studies usually following different design paradigms (quantitative, qualitative, mixed method) or exhibiting varying parameters (sample sizes, use of controls, etc.).

On the other hand, we assume that ambiguity in results is also due to the narrative form of presenting evidence. While the narrative format helped to follow the interpretation of results by providing a certain ‘reading flow’, the use of connecting vocabulary between statements and stylistic elements, such as the use of ‘similarly’ or ‘also’ opened a space for interpretation. Also, we encountered instances where two or more distinct yet strongly connected statements of evidence made within a single sentence (usually indicated by use of the appropriate conjunctions ‘and’, ‘or’) only

³ We focused solely on the result section of the review texts; discussions or summaries outside of the result section were not included in our examination, because these do not tend to relate directly to the selected studies but rather aggregate the results from a ‘bigger-picture’ perspective.

referred to one overarching reference. For our analysis, we recorded these passages in the reviews as instances of multiple affirmation along with the respective number of referenced studies, but also marked them as being unclear in their link to the presented evidence. These observations highlight both the need and the challenges of concise presentation of research evidence in narrative form. Consequently, the narrative format proved useful to describe the wider field and its developments but was not always precise when it came to judging single results.

All in all, the number of affirmations found in each review varied greatly, as presented in Figure 1, depending on the available literature for each research question and educational sector. Twofold affirmations occur most often, with the sectors ‘Teacher Education’ (grey columns) and ‘Childhood, Youth and Family’ (light blue columns) leading with 23 twofold affirmations of evidence each, albeit in different volumes. One would expect a link between the total number of selected studies, the total number of multiple affirmations and the ‘weight’ of affirmations: reviews with a larger corpus would show more affirmations as well as a larger number of referenced studies for each piece of narrated evidence. However, we could not extract this kind of interdependence from the data (see Fig. 1). Apart from the fact that there is a general tendency towards smaller numbers of affirmations, there is no clear, discernible pattern that is linked to the educational sector or volume.

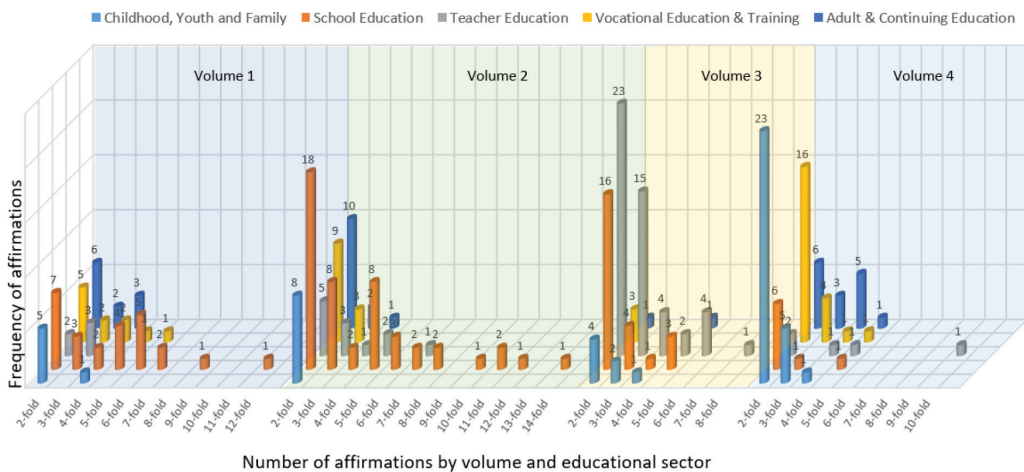


Figure 1. Number of multiple affirmations by volume and educational sector

Skimming for contradictory evidence

In retrospectively recording instances of contradictory evidence, it was sometimes challenging to decide whether a passage qualified as such, for example, when reported effects were similar between studies in size or alignment but the reported levels of significance differed. We solved this problem by discussing the passages between raters. While multiple affirmations of evidence were common in

all reviews, this was not the case with contradictory evidence. Overall, we found only 15 instances – distributed among nine of twenty reviews – where contradictory evidence was presented and analysed. This observation applies regardless of how open-endedly the respective research questions had been expressed by the author(s). Admittedly, the formulations of the research questions generally tended to invite a positivist approach to answering them. Still, reflecting upon our collected experiences in the screening process, we believe that this may be due to the heterogeneity in the field of digital education and difficulties of comparing single study results.

Handling thematic multiplicity

As indicated above, the thematic spread was substantial on different levels. Naturally, the thematic foci varied greatly between editions in the book series; this was a conceptual decision to cover different facets of the overarching field of digitalisation in education. Also, looking at the contributions in the different editions, the range of topics and research questions was considerable. As the reviews tended towards open-ended research questions, they also tended to cover larger topics that were split into different aspects by the authors, showing less overlap than one would assume. This naturally reduced the chances of multiple affirmations since studies are less likely to simultaneously present evidence on different thematic aspects (see Fig. 1). So more openly formulated research questions do not necessarily lead to more comparable or closely linked evidence, but they always lead to a substantial delineation in the research field. We suggest that this is not solely owing to the nature of the questions posed, but also to the fact that many terms and theoretical concepts related to the field of digital education are not clearly defined or have undergone substantial changes through discourse in the scientific community (e.g., blended learning, Koschorreck & Gundermann, 2022; media and information literacy, Leber et al., 2023). Another reason for the low number of multiple affirmations could be that a variety of different technologies were considered in the individual studies. Even if these can be categorised similarly, a comparison between two media systems and their effects could be difficult due to their complexity and the number of confounding variables, so that only limited comparable evidence could be identified. In our view, critical reviews are well-suited to these circumstances: they can highlight heterogeneity and ‘fuzziness’ in the field and potentially serve as a basis for furthering theoretical discourse.

On the other hand, the usefulness of our critical reviews as a means of supporting research results or revealing contradictory evidence in a research field is clearly limited. To this end, other forms of reviews should be favoured, while selecting a more homogenous body of research and aiming for comparability of evidence, for example, by selecting only studies that follow a quantitative paradigm. However, we argue that, given its nature, comparability of evidence is a lingering concern in educational research in comparison to other fields of science.

Instead, critical reviews can excel by revealing broader directions in empirical evidence related to clearly delineated topics while also outlining blank areas in the respective research. They are inherently eclectic regarding different research paradigms, allowing for the integration of diverse types of evidence and interpreting them meaningfully on a higher level. Ideally, this approach can generate new insights from a holistic perspective. This is especially true when authors deliberately try to extract meaning from the reviewed evidence while adopting perspectives from educational practice or policy (see Section 4.4).

The challenge of imposing order on evidence from diverse origins remains critical and requires either systematically using existing categories, adapting them where necessary, or even generating new sets of categories. We found that incorporating existing models related to the respective topic or research question could be an asset for adapting or generating categories grounded in related theory.

5.3 Referencing models in critical reviews

One way of dealing with the set of included studies can be the explicit reference to models, which – if available in the subject area – can provide an important frame of reference for categorising study results. The contribution on model building (Grant & Booth, 2009) in research syntheses can consist of integrating existing models into the analyses, adding to existing models, or setting incentives for new models. Reference to a model is not essential for a critical review. However, the review format often proves suitable for referencing models due to its reflective character. If such a decision is made during the review process, for example, because a model is frequently cited and therefore serves as an appropriate framework (Heinemann et al., 2022), an added perspective is offered. Thus, the synthesis process can be structured and serve as a matrix for category development or at least support the task of category building. This is particularly helpful when dealing with a heterogeneous corpus of literature.

One example for model referencing in our project, which provided both a conceptual framework and facilitated educational policy categorisation, was the integration of framework models developed by the European Union (EU) in two of the review volumes. The first volume, focusing on the competences of educational staff, incorporated the DigCompEdu model of the European Commission in several places in the analysis (Redecker & Punie, 2017; Wilmers et al., 2020). In the second volume, which addresses organisational development and digital education, the EU DigCompOrg model provided an important frame of reference for the development of the reviews (Kampylis et al., 2015; Wilmers et al., 2021). Both models made it possible to include an additional and overarching perspective in the research syntheses, particularly due to their educational policy and practical as well as international orientation.

In line with the topic of competences of educational staff, the use of existing competence models from educational research serves as a further example. The individual research syntheses explicitly mentioned existing competence models to map the

state of research (Wilmers et al., 2020). Besides the important role these models play within the research syntheses, contrasting models enabled an overarching comparison of existing competence models from diverse education sectors pointing to research gaps and avenues for future research (Wilmers et al., 2023c).

A third example is the inclusion of explicit frameworks for the analysis process in research syntheses from the school sector and the sector for adult and continuing education in the third review volume. For the school sector, the model for the comparative analysis of political measures by Kozma (2008) served as a frame of reference for the classification of literature, while for adult education, consideration of the models for technology-assisted learning by Weinberger (2018) supported the process of including and excluding references (Heinemann et al., 2022; Koschorreck & Gundermann, 2022). In both cases, the categorisation of literature was facilitated by reference to the models (Heinemann et al., 2023).

As these examples highlight, critical reviews, with their reflexive and narrative approach, are particularly well-suited to incorporating a framework for analysis. This approach provides a suitable guideline for structuring and synthesising heterogeneous literature in response to the open research question of the review. A framework provides the authors of the review with a common thread that can be further differentiated or can serve as a starting point for further thoughts. At the same time, it can be helpful to the reader, offering a logical structure and presentation of the results.

Nevertheless, and especially if a framework is used to structure the coding process, bias might arise in the interpretation and presentation of the results. In general, the integration of such a guiding framework in the review process can already be applied in the conception of a critical review or in the later synthesis of results and thus structure the development (Newman & Gough, 2020). Referring to a suitable model for the critical review from the beginning of the approach can facilitate the conceptualisation of the review, for example, by aiding in the formulation of the guiding research questions or the definition of suitable study types. Nevertheless, it can also lead to a massive exclusion of studies due to other keywords or other variables that do not fit into the original reference framework. The decision as to whether a framework should be used and the choice of the respective strategy of model integration should be carefully weighed up in terms of effort and benefits with the topic and the research field in mind.

5.4 Formulating implications: considering audiences and aspirations

A crucial part of a critical review lies in deriving implications in terms of perspectives for future research. Across all educational sectors and reviews in the project, the authors were consistently able to clearly identify either gaps or reveal insufficient evidence based on their findings, from which they derived implications for possible future research activities in digital education. In the context of the reviews reflected in this paper, another focus was on implications for practice in the different fields of education. These were addressed to varying degrees in the reviews – sometimes in

single paragraphs (e.g., Leber et al., 2023) – other times in a whole section of the discussion and conclusion (e.g., Koschorreck & Gundermann, 2023). Regardless of their extent, the implications distilled from the evidence cater to two very different audiences: researchers and practitioners. Both endeavours are united by a common challenge when addressing digitalisation in education: they relate to research published in a highly dynamic field. Consequently, when formulating implications and desiderata, the underlying evidence must be considered from two perspectives: (1) against the background of the current status quo in the respective field of education concerning digitalisation; and (2) in the light of technological developments that have occurred since the evidence was generated. In this context, the authors applied an array of strategies, ranging from more commonly providing contextual evidence on digitalisation to discussing implications (e.g., Capparozza & Kathmann, 2022) and going as far as considering results generated through a survey on the challenges of digitalisation in adult and continuing education involving educational practitioners (Koschorreck & Gundermann, 2024). Nevertheless, formulating implications for practice and research perspectives is inevitably a matter of balancing concreteness and generality. To provide orientation, implications for practice and topics for future research need to be derived from the reviewed evidence and avoid vagueness. At the same time, implications and desiderata need to be formulated openly enough to account for both the contingency and the dynamics of technological development to go beyond snapshots.

Although reviews often tend to be regarded as more useful than individual studies by practitioners, reviews alone cannot guarantee the successful transfer of evidence into educational practice (Spoden et al., 2022). The transfer of synthesised knowledge needs to be made more salient and visible through a cascading approach that uses a variety of complementary activities, such as publications in different outlets and formats for educational professionals (e.g., blogs, podcasts, and the organisation of dialogue events between professionals and researchers). In our case, an approach that combined sustainable online platforms and various forms of brokerage activities helped to facilitate the research syntheses' findings (for an overview, see Al-Baghdadi et al., 2024).

6. Conclusion

Taking all the insights into account, we can draw several conclusions about the limitations of critical reviews, the opportunities that this type of review offers and, more importantly, the aspects that authors need to be aware of to maximise the potential of critical reviews. The review type is characterised by a comprehensive literature search on the one hand and a limited quality assessment on the other hand, leading to a partial imbalance in this regard. In between these two requirements lies the difficulty of implementing a systematic and transparent work process, as well as the chance of handling a very heterogenous set of studies and integrating diverse findings into the

overall synthesis via broader discussions and narrative analyses. We found that leveraging the project setting, which included several educational sectors and interdisciplinary subgroups, proved very helpful in discussing, defining and reiterating core templates and/or procedures to tackle these issues.

The literature searches, very similar to those conducted in systematic reviews, aimed to identify all relevant literature, at least within certain search parameters. However, the screening processes revealed the difficulty of handling large datasets without explicit inherent quality controls beyond checking for inclusion and exclusion criteria, and assessing the overall relevance. This is where the focus of critical reviews becomes clear: to categorise a topic first and foremost in the field and to consider it with an overarching focus. This approach has both strengths and weaknesses. Its strength lies in providing an overview of what is often a complex and even confusing field of research. Its weakness, however, is the potential for inaccurate statements and bias. Challenges that first became apparent in the literature search and arose particularly due to the combination of the review type with rather open review questions resulted in a very broad literature corpus. These challenges then shifted to the screening phase, where the time-consuming task of study selection had to take place. They also influenced the analysis, which consequently focused on the depiction of thematic diversity and could not always guarantee exact comparability. As a result, statements were ultimately emphasised, which did not intensively validate individual results, but characterised the field relatively reliably.

In conclusion, when choosing a specific review type, we stress the importance of being very clear on the ultimate purpose of each research synthesis. In our case, the format of the critical review has proven useful to discuss specific topics within the wider and heterogeneous research field of digitalisation in education while aiming to keep pace with dynamic changes in this area. Nevertheless, more specific methodological standards for critical reviews are needed for a clearer distinction between this and other review types. In cases where the primary aim of a review is to inform practice and policy on such dynamic developments within a short time frame, a rapid review might be a more suitable choice. Moreover, our results, in many cases, could not offer in-depth and detailed recommendations enabling clear statements on, for example, what works and what does not. For such purposes, a meta-analysis based on a specific set of comparable empirical studies offers an effective approach to synthesise findings and produce more reliable estimates of an effect. In the presented work, the focus was on critical analysis, corresponding to the characteristics of research in digital education, which is investigated from very different perspectives, dealing with very different contextual factors and constantly adapting to new requirements and developments – factors that limit comparability. The wide variety in the field also became clear when comparing research on digital education from different educational sectors, as some sectors, such as the school sector, could rely on a larger research base whereas other educational sectors had both fewer studies in general and less empirical data on which assumptions could be based.

Furthermore, a specific advantage within our project was that review processes were undertaken 20 times, which allowed for some generalisations at least within our project frame. This enabled the review team to differentiate methodological observations between individual cases and common characteristics and to examine possible dependencies and biases. For instance, regarding the relation between review questions and the size of the literature corpus and the number of included studies, the type of review questions (broad or narrow) did not necessarily influence the size of the overall literature corpus or the set of included studies. Conversely, the extent to which a question referred to specific educational systems and country contexts or enabled relatively easy international transferability influenced both the size of the overall corpus and the amount of included studies.

Finally, we can only point to the importance of extracting implications and ensuring transfer processes that enable review findings to have some influence outside of academia. In our experience, through experimenting with very different formats of knowledge transfer, critical reviews can provide a solid foundation to discuss research results in broader settings with educational practice and policy, provided sufficient time is allocated to this additional task.

Funding

This work was funded by the German Ministry of Education and Research under the numbers 01JD1800A to D.

Author biographies

Annika Wilmers is a professor at Frankfurt University of Applied Sciences and an affiliated researcher at the DIPF | Leibniz-Institute for Research and Information in Education. As the coordinator of a working group on research syntheses within the project ‘Digitalisation in Education’ (Digi-EBF), she has co-edited and co-authored several books, book chapters and articles on research synthesis methodology in education science.

Pia Sander is a post-doctoral researcher in the Learning Lab at the University of Duisburg-Essen, where she contributes to research on digital education and knowledge transfer. In the Digi-EBF project she served as a coordinator and contributed to research syntheses and methodological reflections in the field of digital education.

Jan Koschorreck works as a research associate at the German Institute for Adult Education. He published several research syntheses during his time on the project Digi-EBF. In addition, he has a strong focus on knowledge transfer.

Carolyn Keller works as a research associate at the DIPF | Leibniz-Institute for Research and Information in Education. As an information specialist, she was responsible for the literature searches within the project ‘Digitalisation in Education’ (Digi-EBF). Her research focuses on literature searches for research syntheses in education science and collaboration with academic librarians.

Jens Leber works as a research associate in the Learning Lab at the University of Duisburg-Essen. His focus is on educational technology and instructional design. Within the project Digi-EBF, he contributed to several articles and research syntheses.

Anna Heinemann works as a research associate in the Learning Lab at the University of Duisburg-Essen. Her work and publications focus on the implications of educational technology in K-12 schools and knowledge mobilisation between educational research and practice. In the Digi-EBF project, she published syntheses on digital education as well as methodological reflections on conducting syntheses in educational research.

References

- Adloff, F., & Neckel, S. (2019). Modernisierung, Transformation oder Kontrolle? In K. Dörre, H. Rosa, K. Becker, S. Bose & B. Seyd (Eds.), *Große Transformation? Zur Zukunft moderner Gesellschaften* (Sonderband des Berliner Journals für Soziologie, pp. 167–180). Springer VS. https://doi.org/10.1007/978-3-658-25947-1_8
- Al-Baghdadi, S., Gundermann, A., Wilmers, A., Kathmann, J., Anda, C., & Achenbach, M. (2024). Transferformate im Metavorhaben Digitalisierung im Bildungsbereich. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF - Methode, Auswertung, Perspektiven* (pp. 211–231). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.10>
- 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 (Eds.), *Systematic reviews in educational research: Methodology, perspectives and application* (pp. 111–127). Springer VS. https://doi.org/10.1007/978-3-658-27602-7_7
- Bedenlier, S., Buntins, K., Wilmers, A., & Kerres, M. (2023). Editorial: Forschungssynthesen in der Mediendidaktik: Ansätze und Herausforderungen. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 54, i–viii. <https://doi.org/10.21240/mpaed/54/2023.12.21.X>
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Sage.
- Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687–699. <https://doi.org/10/d22q4b>
- Brunton, G., Stansfield, C., Caird, J., & Thomas, J. (2017). Finding relevant studies. In D. Gough, S. Oliver & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 93–122). Sage.
- Buntins, K., Kerres, M., & Heinemann, A. (2021). A scoping review of research instruments for measuring student engagement: In need for convergence. *International Journal of Educational Research Open*, 2. <https://doi.org/10.1016/j.ijedro.2021.100099>
- Buntins, K., Bedenlier, S., Marín, V., Händel, M., & Bond, M. (2023a). 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. (2023b). 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>

- Capparoza, M. (2021). Maßnahmen für die curriculare Verankerung mediendidaktischer Kompetenzen im Lehramtsstudium: Ein Critical Review. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen* (pp. 105–129). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830994558.04>
- Capparoza, M., & Kathmann, J. (2022). Gelingensbedingungen für die Nutzung digitaler Medien von Lehramtsstudierenden in der Schulpraxis: Ein Critical Review. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (pp. 103–128). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996224.04>
- Capparoza, M., & Kathmann, J. (2023). Innovative Technologien für effektives Classroom Management? Ein Critical Review über Virtual Reality, Mixed Reality und 360°-Videos in der Aus- und Fortbildung von Lehrkräften. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (pp. 107–133). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830998464.04>
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation. <https://doi.org/10.7758/9781610448864>
- European Commission (European Education and Culture Executive Agency). (2025). *Driving digital transformation in education: Innovative practices from the EU*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2797/9398200>
- Glanville, J. (2019). Searching bibliographic databases. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 73–100). Russell Sage Foundation. <https://doi.org/10.7758/9781610448864.8>
- Gough, D., Oliver, S., & Thomas, J. (Eds.). (2017). *An introduction to systematic reviews* (2nd ed.). 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>
- Haddon, L., Cino, D., & Doyle, M. A. (2023). The challenges of conducting systematic evidence reviews: A case study of factors shaping children’s digital skills. *Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique*, 158(1), 71–90. <https://doi.org/10.1177/07591063221141730>
- 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 (Eds.), *Systematic reviews in educational research: Methodology, perspectives and application* (pp. 23–39). Springer VS. https://doi.org/10.1007/978-3-658-27602-7_2
- Heck, T., Keller, C., & Rittberger, M. (2024). Coverage and similarity of bibliographic databases to find most relevant literature for systematic reviews in education. *International Journal on Digital Libraries*, 25(2), 365–376. <https://doi.org/10.1007/s00799-023-00364-3>
- Heinemann, A., Leber, J., Sander, P., & Ahmad, C. (2022). Strategien und Maßnahmen für die schulische Curriculumsentwicklung im Kontext der Digitalisierung. Ein Critical Review zur vergleichenden Analyse politischer Initiativen. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (Digitalisierung in der Bildung. Forschungsstand und -perspektiven, 3, pp. 69–102). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996224.03>
- Heinemann, A., Koschorreck, J., Wilmers, A., Sander, P., & Leber, J. (2023). Potenziale und Herausforderungen von Critical Reviews im Kontext der Digitalisierung im Bildungsbereich. *Forschungssynthesen in Der Mediendidaktik. Ansätze Und Herausforderungen. MedienPädagogik, Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 79–102. <https://doi.org/10.21240/mpaed/54/2023.08.13.X>
- Heinemann, A., Koschorreck, J., Sander, P., Wilmers, A., Nieding, I., & Leber, J. (2024). Der Screening- und Kodierprozess im Critical Review: Eine vergleichende Auswertung und bildungssektorübergreifende Lessons Learned. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF - Methode, Auswertung, Perspektiven* (pp. 109–131). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.06>
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2023). *Cochrane handbook for systematic reviews of interventions version 6.4*. www.training.cochrane.org/handbook.

- Jäger-Dengler-Harles, I., & Rittberger, M. (2022). Datenerhebung zum Wissensstand. Systematische Literaturrecherche zur Forschungssynthese über die Erscheinungsformen sozialer Ungleichheiten des Bildungserwerbs. In A. Bachsleitner, R. Lämmchen & K. Maaz (Eds.), *Soziale Ungleichheit des Bildungserwerbs von der Vorschule bis zur Hochschule. Eine Forschungssynthese zwei Jahrzehnte nach PISA* (pp. 17–42). Waxmann Verlag GmbH.
- Kampylis, P., Punie, Y., & Devine, J. (2015). *Promoting effective digital-age learning. A European framework for digitally-competent educational organisations*. Publication Office of the European Union.
- Keller, C., & Heck, T. (2024). Die Effizienz von Datenbankrecherchen: Eine systematische Reflexion am Beispiel von Digi-EBF. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF - Methode, Auswertung, Perspektiven* (pp. 67–90). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.04>
- Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (Technical Report EBSE-2007-01). Keele University.
- Kollmorgen, R., Merkel, W., & Wagener, H. J. (2015). Transformation und Transformationsforschung: Zur Einführung. In R. Kollmorgen, W. Merkel & H. J. Wagener (Eds.), *Handbuch Transformationsforschung* (pp. 11–27). Springer VS. https://doi.org/10.1007/978-3-658-05348-2_1
- 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 (Eds.), *Bildung im digitalen Wandel. Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung* (pp. 159–193). Waxmann Verlag GmbH. <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 (Eds.), *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen* (pp. 161–192). Waxmann Verlag GmbH. <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 (Eds.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (pp. 161–186). Waxmann Verlag GmbH. <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 (Eds.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (pp. 171–201). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830998464.06>
- Koschorreck, J., & Gundermann, A. (2024). Implikationen und Desiderate aus vier Reviews zum Bereich der Weiterbildung. Eine systematische Verknüpfung mit Wissensbedarfen aus der Praxis. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF – Methode, Auswertung, Perspektiven* (pp. 191–209). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.09>
- Köstler, V. (2023). Zwischen Präzision und Sensitivität. Generierung eines Studienkorpus am Beispiel einer Fragestellung zu Künstlicher Intelligenz (KI) in Bildungsprozessen. *MedienPädagogik. Zeitschrift für Theorie und Praxis der Medienbildung*, 54, 1–27. <https://doi.org/10.21240/mpaed/54/2023.08.10.X>
- Kozma, R. B. (2008). Comparative analysis of policies for ICT in education. In J. Voogt & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education* (Vol. 20, pp. 1083–1096). Scholars Portal. https://doi.org/10.1007/978-0-387-73315-9_68
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13 (1), 1–73. <https://doi.org/10.4073/cmg.2016.1>
- 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 (Eds.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (pp. 75–106). Waxmann Verlag GmbH. <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 (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 261–277). Sage.

- Marin, V. I., Buntins, K., Bedenlier, S., & Bond, M. (2023) Invisible borders in educational technology research? A comparative analysis. *Education Technology Research and Development*, 71(3), 1349–1370. <https://doi.org/10.1007/s11423-023-10195-3>
- McFadden, P., Taylor, B. J., Campbell, A., & McQuilkin, J. (2012). Systematically identifying relevant research: Case study on child protection social workers' resilience. *Research on Social Work Practice*, 22(6), 626–636. <https://doi.org/10.1177/1049731512453209>
- McKenzie, J. E., Brennan, S. E., Ryan, R. E., Thomson, H. J., & Johnston, R. V. (2023). Chapter 9: Summarizing study characteristics and preparing for synthesis. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions version 6.4*.
- 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 (Eds.), *Systematic reviews in educational research: Methodology, perspectives and application* (pp. 3–22). Springer VS. https://doi.org/10.1007/978-3-658-27602-7_1
- OECD. (2023). *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- O'Mara-Eves, A., Brunton, G., McDaid, D., Kavanagh, J., Oliver, S., & Thomas, J. (2013). Techniques for identifying cross-disciplinary and 'hard-to-detect' evidence for systematic review. *Research Synthesis Methods*, 5(1), 50–59. <https://doi.org/10.1002/jrsm.1094>
- Paré, G., Trudel, M. C., & Jaana, M. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information and Management*, 52(2), 183–199. <https://doi.org/10.1016/j.im.2014.08.008>
- Petticrew, M. (2015). Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. *Systematic Reviews*, 4. <https://doi.org/10.1186/s13643-015-0027-1>
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences. A practical guide*. Blackwell Publ. <https://doi.org/10.1002/9780470754887>
- Pigott, T. D., Tocci, C., Ryan, A. M., & Galliher, A. (2021). Quality of research evidence in education: How do we know? *Review of Research in Education*, 45(1), vii–xii.
- Polanin, J. R., Pigott, T. D., Espelage, D. L., & Grotpeter, 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>
- Redecker, C., & Punie, Y. (2017). *European framework for the digital competences of educators*. DigCompEdu. Publication Office of the European Union.
- 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>
- Spoden, C., Bernhard-Skala, C., Koschorreck, J., & Rohwerder, J. (2022). *Transfer von Forschungsergebnissen zur Digitalisierung in die Bildungspraxis*. <https://doi.org/10.57776/qy87-qx94>
- Stock, W. G. (2007). *Information retrieval. Informationen suchen und finden*. Oldenbourg.
- Sutcliffe, K., Oliver, S., & Richardson, M. (2017). Describing and analysing studies. In D. Gough, S. Oliver & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 123–144). Sage.
- Sutton, A., Clowes, M., Preston, L., & Booth, A. (2019). Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Information & Libraries Journal*, 36(3), 202–222. <https://doi.org/10.1111/hir.12276>
- Taylor, B., Wylie, E., Dempster, M., & Donnelly, M. (2007). Systematically retrieving research: A case study evaluating seven databases. *Research on Social Work Practice*, 17(6), 697–706. <https://doi.org/10.1177/1049731507304402>
- Terhart, E. (2021). Zukunftsthemen der Schulforschung. In D. Fickermann, B. Edelstein, J. Gerick & K. Racherbäumer (Eds.), *Schule und Schulpolitik während der Corona-Pandemie: Nichts gelernt?* (pp. 147–175). Waxmann Verlag GmbH. <https://doi.org/10.25656/01:23876>
- Thomas, J., Newman, M., & Oliver, S. (2013). Rapid evidence assessments of research to inform social policy: Taking stock and moving forward. *Evidence & Policy. A Journal of Research, Debate and Practice*, 9(1), 5–27. <https://doi.org/10.1332/174426413X662572>

- UNESCO. (2018). *Global framework of reference on digital literacy skills (Information paper No. 51)*. UNESCO Institute for Statistics. <https://uis.unesco.org/sites/default/files/documents/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf>
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5), 360–363.
- Voogt, J., & Knezek, G. (Eds.). (2008). *International handbook of information technology in primary and secondary education*. Scholars Portal.
- Waffner, B. (2021). Schulentwicklung in der digital geprägten Welt: Strategien, Rahmenbedingungen und Implikationen für Schulleitungshandeln. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel: Organisationsentwicklung in Bildungseinrichtungen* (pp. 67–103). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830994558.03>
- Weinberger, A. (2018). Orchestrierungsmodelle und -szenarien technologie-unterstützten Lernens. In S. Ladel (Ed.), *Digitalisierung und Bildung* (pp. 117–139). Springer VS. https://doi.org/10.1007/978-3-658-18333-2_7
- Wetterich, C., & Plänitz, E. (2021). *Systematische Literaturanalysen in den Sozialwissenschaften: Eine praxisorientierte Einführung*. Verlag Barbara Budrich. <https://doi.org/10.3224/84742430>
- White, H. D. (2019). Scientific communication and literature retrieval. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 51–72). Russell Sage Foundation. <https://doi.org/10.7758/9781610448864.7>
- Wilmers, A., Anda, C., Keller, C., & Rittberger, M. (Eds.). (2020). *Bildung im digitalen Wandel. Die Bedeutung für das pädagogische Personal und für die Aus- und Fortbildung*. Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830991991>
- Wilmers, A., Achenbach, M., & Keller, C. (Eds.). (2021) *Bildung im digitalen Wandel. Organisationsentwicklung in Bildungseinrichtungen*. Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830994558>
- Wilmers, A., Achenbach, M., & Keller, C. (Eds.) (2022a), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (pp. 7–30). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996224>
- Wilmers, A., Keller, C., Achenbach, M., & Rittberger, M. (2022b). Reviews zur Bildung im digitalen Wandel: Methodisches Vorgehen im Reviewprozess. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Lehren und Lernen mit digitalen Medien* (pp. 7–30). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996224.01>
- Wilmers, A., Keller, C., Stöbe-Blossey, S., Achenbach, M., & Nieding, I. (2023a). Digitalisierung und Teilhabe in der Bildung. Eine Abdeckung des Themenfeldes und Erläuterungen des methodischen Vorgehens in den Forschungssynthesen. In A. Wilmers, M. Achenbach & C. Keller (Eds.), *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe* (pp. 7–32). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830998464.01>
- Wilmers, A., Achenbach, M., & Keller, C. (Eds.). (2023b). *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe*. Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996224>
- Wilmers, A., Gundermann, A., Hähn, K., Irle, G., Koschorreck, J., Nieding, I., Ratermann-Busse, M., Waffner, B., Anda, C., & Keller, C. (2023c). Kompetenzen des pädagogischen Personals in der digitalen Welt. Eine vergleichende Übersicht zum Forschungsstand in verschiedenen Sektoren des Bildungssystems. In K. Scheiter & I. Gogolin (Eds.), *Bildung für eine digitale Zukunft* (pp. 293–315). Springer VS. https://doi.org/10.1007/978-3-658-37895-0_12
- Wilmers, A. (2024a). Reviewformat und Projektkontext: Eine Auseinandersetzung mit der Zuordnung Critical Review. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF – Methode, Auswertung, Perspektiven* (pp. 16–38). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.02>
- Wilmers, A. (Ed.). (2024b). *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF – Methode, Auswertung, Perspektiven*. Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126>
- Wilmers, A., & Keller, C. (2024). Forschungssynthesen in der digitalen Bildung: Eine vergleichende Auswertung von 20 Literaturrecherchen. In A. Wilmers (Ed.), *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF – Methode, Auswertung, Perspektiven* (pp. 39–66). Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830999126.03>

What is Critical About a Critical Review?

- Wollscheid, S., & Tripney, J. (2021). Rapid reviews as an emerging approach to evidence synthesis in education. *London Review of Education*, 19(1). <https://doi.org/10.14324/LRE.19.1.32>
- 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>
- Zawacki-Richter, O., Cefa, B., & Bai, J. Y. H. (2025). Towards reproducible systematic reviews in open, distance, and digital education—An umbrella mapping review. *Review of Education*, 13(1), Article e70031. <https://doi.org/10.1002/rev3.70031>
- Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (Eds.). (2020). *Systematic reviews in educational research: Methodology, perspectives and application*. Springer VS. <https://doi.org/10.1007/978-3-658-27602-7>

Appendix

Table 4. Listing of all 20 critical reviews (original titles are in German)

Volume 1	Meaning of digitalisation for educational staff and their training and further education; https://doi.org/10.31244/9783830991991
Childhood, Youth & Family	Dealing with digital media in early education in the conflict between protected space and key competence
School	Teaching practices, experiences and attitudes of teachers towards educational technologies at school
Teacher Education	Teacher educators as facilitators of technology integration in teacher education
Vocational Training	Educational technologies in vocational education and training as a challenge for teachers and staff?
Adult & Continuing Education	Implications of digitalisation on teachers and trainers in adult education and training
Volume 2	Organisational development in educational institutions with regard to digitalisation; https://doi.org/10.31244/9783830994558
Childhood, Youth & Family	The implementation of digitalisation in non-formal educational organisations
School	School development in a digital world: strategies, frameworks and implications for school leadership
Teacher Education	Implementing technology infusion in teacher education programmes
Vocational Training	Digitally supported cooperation of organisations in vocational education and training
Adult & Continuing Education	The relevance of digital transformation for organisational management in adult and continuing education
Volume 3	Teaching and learning with educational technologies; https://doi.org/10.31244/9783830996224
Childhood, Youth & Family	How children and teenagers can learn with, from and about educational technologies in non-formal educational contexts
School	Strategic and operational policies in the context of curriculum development
Teacher Education	Conditions for pre-service teachers' technology integration in field experiences
Vocational Training	Designing digital concepts for disadvantaged young people in vocational orientation and preparation
Adult & Continuing Education	Orchestration of blended learning in adult and continuing education

Volume 4	Impact of educational technologies for social learning and participation; https://doi.org/10.31244/9783830998464
Childhood, Youth & Family	Participation through educational technologies at home
School	Strengthening students' information literacy through the use of educational technologies
Teacher Education	Innovative technologies for effective classroom management
Vocational Training	Effects and side effects of digitalisation for people with disabilities in the primary labour market
Adult & Continuing Education	Education with educational technologies for disadvantaged adults – opportunities and challenges