



Best Practices for Conducting Systematic Reviews: Perspectives of Experienced Systematic Review Researchers in Educational Sciences

Yvonne M. Fromm¹ · Florence Martin² · Tuba Gezer³ · Dirk Ifenthaler^{1,4}

Accepted: 8 January 2025 / Published online: 4 February 2025
© The Author(s) 2025

Abstract

Systematic reviews have been gaining attention as a research methodology and are among the most frequently cited sources in educational sciences. However, best practices for conducting systematic reviews in educational sciences are still evolving. We conducted $N=12$ qualitative interviews to learn from experienced systematic review researchers across various educational disciplines and geographic locations. The interviews focused on the different steps of the systematic review process (i.e., designing, including/excluding, screening, coding, analyzing, reporting), benefits, challenges, team collaboration, ethical considerations, and technologies while conducting systematic reviews. Several themes were identified, providing best practices for conducting systematic reviews and highlighting the importance of a systematic, comprehensive, and transparent research process. The findings have implications for researchers who wish to conduct systematic reviews, instructors who teach students on conducting systematic reviews, as well as editors and reviewers of journals who publish systematic reviews.

Keywords Systematic review · Best practice · Guideline · Interview study

✉ Yvonne M. Fromm
yvonne.fromm@uni-mannheim.de

Florence Martin
fmartin3@ncsu.edu

Tuba Gezer
gezer.3@osu.edu

Dirk Ifenthaler
dirk@ifenthaler.info

¹ Learning, Design and Technology, University of Mannheim, Mannheim, Germany

² College of Education, North Carolina State University, Raleigh, USA

³ College of Education and Human Ecology, The Ohio State University, Columbus, USA

⁴ UNESCO Co-Chair on Data Science in Higher Education Learning and Teaching, Curtin University, Perth, Australia

1 Introduction

The scientific community in the context of education has seen a rise in systematic reviews since the beginning of the twenty-first century. For instance, a search in Scopus (<https://www.scopus.com>) for systematic reviews in educational sciences found 13 articles for the year 2002 and 749 articles for the year 2022. A systematic review is a type of secondary research that aims to identify, evaluate, and synthesize primary research in a transparent and analytical manner to draw conclusions based on the evidence (Martin et al., 2020). In previous research, the term systematic review has been used to designate either a research methodology that aims to synthesize primary studies in a qualitative manner (Grant & Booth, 2009) or a family of research methodologies (including meta-analyses) that aim to synthesize primary studies in a qualitative or quantitative manner (Page et al., 2021b; Pigott & Polanin, 2020). In this paper, we focus on the former, that is, systematic reviews as a research methodology that synthesizes (quantitative, qualitative, or mixed-method) primary studies in a qualitative manner.

Systematic reviews help summarize existing research in a standardized and reproducible way and can guide future research and practice in educational sciences (Alexander, 2020; Greyson et al., 2019). For example, systematic reviews may reveal issues in primary research that should be addressed in future studies or evaluate theories and educational interventions more comprehensively than individual primary studies (Alexander, 2020; Page et al., 2021a). Consequently, review journals and systematic review papers are among the most frequently cited sources in educational sciences (Alexander, 2020). Therefore, to guarantee that a systematic review is valuable for research and practice, authors, journal editors, and reviewers must ensure the quality and methodological rigor of systematic reviews (Alexander, 2020; Greyson et al., 2019; Page et al., 2021a).

Various guidelines for conducting and reporting systematic reviews exist (e.g., Higgins & Green, 2008b; Liberati et al., 2009). These guidelines aim to help authors conduct their systematic review in an analytical and comprehensive manner as well as write a transparent and reproducible systematic review report. However, several of these guidelines focus on health sciences (e.g., Higgins & Green, 2008b; Liberati et al., 2009) and may not meet the specific needs of systematic reviews in educational sciences, as research in educational sciences differs from health sciences (e.g., in research methodology and definition of constructs; Higgins & Green, 2008b; Liberati et al., 2009; Zawacki-Richter, 2020). Moreover, most existing guidelines mainly focus on the designing, including/excluding, screening, coding, analyzing and reporting phases (DISCAR) of systematic reviews, but do not consider aspects such as team collaboration, ethical considerations, or technology usage which have also been considered crucial when conducting systematic reviews (Borah et al., 2017; Tsafnat et al., 2014; West & Martin, 2023; Zawacki-Richter, 2020). For example, systematic reviews are usually conducted in a research team, but evidence-based best practices for effective team collaboration are still missing (Borah et al., 2017). Further, research on ethical considerations related to systematic reviews has mainly concentrated on biases in systematic review findings, but investigations on other ethical considerations as well as on how to deal with them when conducting systematic reviews are missing (Suri, 2020; Zawacki-Richter, 2020). Moreover, evidence-based best practices on technology usage when conducting systematic reviews may help authors conduct more comprehensive, efficient, and accurate systematic reviews (Tsafnat et al., 2014; Page et al., 2021a).

In addition, most existing guidelines for conducting systematic reviews are based on the individual subjective experiences of the authors of these guidelines and have not

been obtained through empirical research (Alexander, 2020; Higgins & Green, 2008b). To develop evidence-based best practices for conducting systematic reviews in educational sciences, the systematic review process needs to be studied from the perspectives of multiple experienced systematic review researchers rather than from the limited perspectives of a single author or research team (Page et al., 2021a). Experienced systematic review researchers can reflect on the benefits of systematic reviews in educational sciences as well as on what worked well and what challenges they faced in the systematic review process. Little is known about how researchers in educational sciences perceive the systematic review process, how they use existing guidelines and what strategies (e.g., for searching and screening studies) they perceive as most useful. Therefore, in this study, we conducted semi-structured interviews to identify best practices for conducting systematic reviews in educational sciences based on the perspectives of experienced systematic review researchers across various educational disciplines. Such insights may provide guidelines for enhancing the quality and effectiveness of systematic reviews in educational sciences. In this paper, we first review existing research on and guidelines for systematic reviews. We then present our interview methodology and findings. The findings are discussed and implications for systematic reviews in educational sciences are derived.

2 Theoretical Background

2.1 Guidelines for Conducting Systematic Reviews

Table 1 provides an overview of existing guidelines for conducting and reporting systematic reviews. The systematic review approach stems from the concept of evidence-based medicine, in which systematic reviews have been primarily used to determine the effectiveness of healthcare interventions (Hammersley, 2020). Therefore, several existing guidelines for systematic reviews have been developed for the field of health sciences (e.g., Higgins & Green, 2008b; Page et al., 2021a, 2021b). For example, Higgins and Green (2008b) edited a handbook providing instructions for conducting systematic reviews in health sciences. Further, Needleman (2022) provided instructions on systematic review methods (e.g., search strategies, quality appraisals), pooling data, and drawing conclusions from systematic review findings. One of the most commonly used guidelines across disciplines is the PRISMA statement (Liberati et al., 2009; Page et al., 2021a, 2021b). The PRISMA statement provides a checklist for writing a transparent and complete systematic review of healthcare interventions (Page et al., 2021a). Although originating from health sciences, the PRISMA statement (Liberati et al., 2009; Page et al., 2021a, 2021b) has become popular in the field of educational sciences, and several educational review journals require authors to adhere to the PRISMA statement when writing their systematic review paper (e.g., *Review of Educational Research*, 2023). However, we suggest that there are at least three reasons why the PRISMA statement may have limitations when applied to educational sciences.

First, several items of the PRISMA statement focus on quantitative syntheses of primary studies in the form of meta-analyses and may not be applicable to summarize primary studies in a qualitative manner (Page et al., 2021a, 2021b). Therefore, more specific guidelines are needed for conducting qualitative research syntheses in educational sciences. Second, while in health sciences, disorders and diseases are defined by strict diagnosis criteria, most concepts in educational sciences (e.g., student engagement) are multi-faceted and

Table 1 Guidelines for conducting and reporting systematic reviews

References	Discipline	Main focus	Empirical	Guidelines for		
				DISCAR process	Team collaboration	Eth. considerations
Alexander (2020)	Education	Conducting	No	Yes	No	No
Gough et al. (2017)	Multiple	Conducting	No	Yes	No	Yes
Hallinger (2013)	Education	Conducting	No	Yes	No	No
Higgins and Green (2008)	Health sciences	Conducting	No	Yes	No	No
Kitchenham (2004)	Software engineering	Conducting	No	Yes	No	No
Liberati et al. (2009)	Health sciences	Reporting	Yes	Yes	No	No
Needleman (2022)	Health sciences	Conducting	No	Yes	No	No
Okoli and Schabram (2010)	Information systems	Conducting	No	Yes	No	No
Page, McKenzie, et al. (2021a, 2021b); Page and Moher et al., (2021a, 2021b)	Health care	Reporting	Yes	Yes	No	No
Petticrew and Roberts (2005)	Social sciences	Conducting	No	Yes	No	No
Reed et al. (2005)	Medical education	Conducting	Yes	Yes	No	No
Shaffril et al. (2021)	Multiple	Conducting	Yes	Yes	No	No
Siddaway et al. (2019)	Psychology	Conducting and reporting	No	Yes	No	No
West and Martin (2023)	Education	Conducting	No	Yes	No	No
Xiao and Watson (2019)	Multiple	Conducting	Yes	Yes	No	No
Zwacki-Richter et al. (2020)	Education	Conducting	No	Yes	No	Yes

have various definitions in the literature. Therefore, searching for relevant studies and the definition of inclusion and exclusion criteria may be especially challenging in educational sciences. On the one hand, all synonyms of relevant search terms need to be considered to cover all relevant studies (Alexander, 2020; West & Martin, 2023). On the other hand, since some authors may apply the same term to different constructs, such comprehensive searches may also lead to a large number of irrelevant search results. Accordingly, finding the ideal balance between relevance and completeness of the search strategy has been considered a key issue in conducting systematic reviews in educational sciences (Zawacki-Richter, 2020). Third, the PRISMA statement mainly focuses on *reporting* systematic reviews and meta-analyses rather than best practices for *conducting* systematic reviews (Page et al., 2021a, 2021b). When conducting systematic reviews, researchers may face various challenges, such as defining appropriate search parameters, and evidence-based best practices to deal with these challenges in educational sciences are still missing (Alexander, 2020; Dowd & Johnsons, 2020).

To meet the specific needs of systematic reviews in educational sciences, researchers have begun to develop guidelines for conducting systematic reviews of educational research. For example, Zawacki-Richter et al. (2020) edited a handbook including chapters on methodological approaches to conducting systematic reviews (Hammersley, 2020; Newman & Gough, 2020). Moreover, Alexander (2020) and Reed et al. (2005) summarized various challenges of conducting systematic reviews in educational sciences as well as guidelines for dealing with these challenges. Hallinger (2013) developed a conceptual framework for systematic reviews in educational leadership and management by summarizing previously conducted systematic reviews. Further, West and Martin (2023) provided several strategies for effectively writing a systematic review paper in educational sciences as well as a taxonomy of review papers. However, most of these guidelines are limited to the individual subjective experiences of the authors of these guidelines and have not been obtained through empirical research. Further, existing guidelines for conducting systematic reviews (in educational sciences and other research areas) mainly focus on the DISCAR process of systematic reviews (West & Martin, 2023) and do not consider further aspects, such as team collaboration, ethical considerations, and technology usage when conducting systematic reviews (Borah et al., 2017; Suri, 2020; Tsafnat et al., 2014).

2.2 DISCAR Process

The DISCAR process (West & Martin, 2023) summarizes the systematic review process into the phases designing (D), including/excluding (I), screening (S), coding (C), as well as analyzing (A) and reporting (R). In the designing phase, researchers identify research questions as well as a theoretical and methodological framework for their systematic review. For the designing phase, existing guidelines for educational sciences recommend authors to check whether the identified research questions are meaningful and can be answered through a systematic review as well as to determine whether there is an appropriate number of primary studies to conduct a systematic review (Alexander, 2020; West & Martin, 2023). The including/excluding phase includes strategies for defining inclusion and exclusion criteria as well as appropriate search parameters. Existing guidelines for educational sciences highlight the importance of relating the inclusion and exclusion criteria to the research questions, identifying all synonyms for search terms and an appropriate time frame for the search, as well as considering unpublished literature (Alexander, 2020; Newman & Gough, 2020; West & Martin, 2023). In the screening phase, researchers apply their inclusion and

exclusion criteria to their search results. Existing guidelines recommend authors to remove duplicates before screening, to involve multiple researchers in the screening phase, as well as to check titles and abstracts before screening full texts (Newman & Gough, 2020; West & Martin, 2023). In the coding phase, the selected studies are classified according to a predefined coding system. For this phase, existing guidelines for educational sciences recommend authors to establish a codebook that includes all relevant variables as well as to involve multiple researchers in the coding phase (Newman & Gough, 2020; West & Martin, 2023). Finally, in the analyzing and reporting phases, researchers synthesize findings and write their systematic review paper. For the analyzing and reporting phases, existing guidelines advise authors to assess the quality of the primary studies to either exclude low-quality studies or to weigh them differently, to report descriptive information about the primary studies, to identify patterns, trends, and gaps in primary research, as well as to provide implications for theory, research, and practice (Alexander, 2020; Reed et al., 2005; West & Martin, 2023).

The phases of the DISCAR process have been considered central for conducting systematic reviews (Alexander, 2020; Page et al., 2021a). However, as mentioned above, existing guidelines for the phases of the DISCAR process in educational sciences are based on the individual subjective experiences of the authors, and evidence-based best practices for the phases of the DISCAR process are missing in educational sciences. Moreover, current research and technological developments have suggested that guidelines for conducting systematic reviews should not only focus on the DISCAR process, but that further aspects such as team collaboration, ethical considerations, and technologies used for conducting systematic reviews also need to be considered to ensure the methodological rigor of systematic reviews (Suri, 2020; Tsafnat et al., 2014; Zawacki-Richter, 2020).

2.3 Team Collaboration

Conducting a systematic review is a labor-intensive task. Therefore, systematic reviews are usually conducted by a research team rather than a single researcher (Borah et al., 2017; Zawacki-Richter, 2020). However, team collaboration when conducting systematic reviews may be challenging, as team members may have different views on educational concepts, the relevance of specific primary studies, or the data to be extracted from the primary studies. Therefore, reliable training procedures are crucial to ensure consistency in the systematic review process (Okoli & Schabram, 2010). However, little is known about how researchers in educational sciences establish such training procedures and ensure effective team collaboration.

2.4 Ethical Considerations

Contrary to primary study researchers, researchers conducting systematic reviews do not collect personal data from participants. Therefore, institutional ethics approval is usually optional for systematic reviews (Suri, 2020). However, with the increasing role of systematic reviews in educational research and practice, ethical considerations, such as biases, influences through funding sources, or lack of transparency and comprehensiveness, must be given greater attention (Page, et al., 2021b; Suri, 2020). Suri (2020) offered guidelines for ethical decision-making when conducting systematic reviews and summarized various ethical considerations related to biases in systematic reviews (e.g., biases through languages). However, little is known about other ethical considerations related to systematic

reviews, the extent to which researchers in educational sciences are aware of ethical considerations related to systematic reviews, and whether and how they try to deal with these considerations.

2.5 Technologies

There are several technologies for supporting and accelerating the systematic review process (Scott et al., 2021; Tsafnat et al., 2014). For example, researchers may use reference management software, such as Citavi (<https://www.citavi.com/de>) or EndNote (<https://endnote.com/>) to manage literature search results. Further, advances in the field of artificial intelligence have enabled the use of machine learning to assist and (partially) automate specific phases of the systematic review process, such as searching for literature or screening studies (Blaizot et al., 2022; Brunton et al., 2017; de la Torre-López et al., 2023; Scott et al., 2021; Tsafnat et al., 2014). However, two recent meta systematic reviews on educational technology and artificial intelligence in education revealed that a substantial amount of research syntheses did not reveal technology usage in their manuscripts (Bond et al., 2023, 2024). Therefore, further research is needed to understand how researchers in educational sciences use technologies when conducting systematic reviews, and which technologies are perceived as most useful.

2.6 Benefits, Challenges, and the Future Outlook of Systematics Reviews

In addition to the aspects discussed above, experienced systematic review researchers' perspectives on the benefits, challenges, and the future outlook of systematic reviews may help develop evidence-based best practices for conducting systematic reviews in educational sciences. Such insights may highlight positive and negative aspects of systematic reviews as well as critical issues that should be addressed in guidelines for systematic reviews (Alexander, 2020; Mallett et al., 2012; Reed et al., 2005).

Mallett et al. (2012) highlighted several benefits of conducting systematic reviews. For example, systematic reviews force researchers to conduct a comprehensive and systematic search of literature that may extend beyond the literature already known to the researchers and their networks. Therefore, systematic reviews may help decrease researcher bias and provide a more thorough overview of current literature than narrative literature reviews (Mallett et al., 2012). Further, publishing a systematic review requires researchers to present their review process in a transparent and detailed manner that allows for reproducibility and thorough evaluation of the systematic review findings (Mallett et al., 2012; Page et al., 2021a).

However, conducting a systematic review may be challenging (Alexander, 2020). According to Alexander (2020), researchers conducting systematic reviews in educational sciences may face framing (e.g., finding an appropriate research question), procedural (e.g., defining search parameters), consolidating and summarizing (e.g., charting relevant information), as well as interpreting and communicating (e.g., detecting trends and patterns) challenges. Further, Reed et al. (2005) proposed a list of several challenges (e.g., finding relevant literature) related to systematic reviews of educational intervention studies.

However, these lists of benefits and challenges have not been obtained through empirical research, and the completeness and relevance of these lists remain unclear (Alexander, 2020; Mallett et al., 2012; Reed et al., 2005). Moreover, despite the increasing attention to systematic reviews, there has been little critical reflection on the use and appropriateness of

systematic reviews as a research methodology in educational sciences as well as the future outlook of systematic reviews in educational sciences (Hammersley, 2020; Mallett et al., 2012). Methodological approaches to systematic reviews and technologies to support the systematic review process are constantly evolving. For example, the PRISMA statement, first published in 2009 (Liberati et al., 2009), was updated in the year 2020 due to innovations related to systematic reviews (Page et al., 2021a, 2021b). Therefore, valid guidelines for conducting systematic reviews in educational sciences should consider benefits, challenges, and the future outlook of systematic reviews in educational sciences.

3 Research Questions

Based on the research discussed above, the following research questions were formulated for our interview study:

1. What are best practices for the designing, including/excluding, screening, coding, analyzing, and reporting phases (i.e., the DISCAR process) of systematic reviews in educational sciences?
2. What are best practices for team collaboration, ethical considerations, and technologies when conducting systematic reviews in educational sciences?
3. What are experienced systematic review researchers' perspectives on the benefits, challenges, and future outlook of systematic reviews in educational sciences?

4 Methods

4.1 Research Design and Interview Protocol

Semi-structured interviews were chosen to answer our research questions. The interview protocol was developed based on the DISCAR process (West & Martin, 2023) and additional aspects (e.g., team collaboration, ethical considerations, technologies) identified from current research (e.g., Blaizot et al., 2022; Suri, 2020; Zawacki-Richter, 2020). The interview protocol was discussed within the research team and adapted in iterative processes based on the discussions. The interview protocol consisted of 16 questions to collect in-depth qualitative data about the experienced researchers' process of conducting a systematic review. All questions were open-ended and included follow-up questions when there was a need for clarification. The interview protocol is included in the Appendix.

4.2 Participants

Convenient and snowball sampling were used to identify eligible researchers to participate in our study. The research team (who was located in Germany and the United States) used their personal and professional connections to invite qualifying systematic review researchers to participate in the study. In addition, participants were asked at the end of the interview if they knew someone who was eligible for our study. The eligibility criterion was having more than one systematic review published to ensure rich knowledge about conducting systematic reviews. Participation was voluntary, and institutional research approval was received before participant recruitment. Twelve participants (7 female and

5 male researchers) provided consent to be interviewed. Participants were located across Europe and the United States and had conducted between three and 15 systematic reviews ($M=5.50$, $SD=3.23$). The participants included postdocs, senior researchers, senior lecturers, assistant professors, and professors from various educational disciplines (e.g., special education, educational leadership, educational technology).

4.3 Data Collection

All participants were interviewed via the Zoom web conferencing application (<https://zoom.us/>) between December 2022 and April 2023. The interview protocol was shared with participants once the interviews were scheduled to allow them to prepare for the interview. Each interview was conducted by two research team members and lasted between 34 and 55 min ($M=43.89$, $SD=6.30$).

4.4 Data Analysis

All interviews were transcribed. The automated Zoom transcription tool (<https://zoom.us/>) was used to assist generating the first draft of the interview transcript. The transcripts were then manually cleaned by the research team. The transcripts were analyzed using QCamap (<https://www.qcamap.org/ui/de/home>). Both inductive and deductive coding techniques (Mayring, 2015) were used. The phases of the DISCAR process as well as the additional aspects specified in the research questions (e.g., team collaboration, ethical considerations, benefits, challenges) were used as deductive main categories. Within these main categories, subcategories were added during the coding process using an inductive coding technique. These subcategories included best practices for conducting systematic reviews (e.g., consider journal requirements) or perspectives on the benefits, challenges, and the future outlook of systematic reviews in educational sciences (e.g., helpful for advancing own research) that the participants mentioned during the interviews. To develop a codebook, two of the researchers coded two interviews independently (interrater reliability: Krippendorff's $\alpha=0.67$). Differences between the two coders were resolved by discussion. Then, the codebook was reviewed within the whole research team and used to code the remaining interviews. Inductive codes were added to the codebook if new themes emerged from the remaining interviews. The two researchers discussed and reconciled the codes through discussion where researchers explained their perspectives and refined code definitions or categories. Regular peer debriefing was conducted during data analysis.

5 Results

5.1 Best Practices for the DISCAR Process

An overview of the findings for our first research question are presented in Table 2.

5.1.1 Designing

Participants encouraged researchers to be guided by their personal research goals when designing a systematic review: "Research questions basically come from [...] the common

Table 2 Best practices for the DISCAR process

DISCAR phase	Best practices
Designing	<ul style="list-style-type: none"> <i>Identify a meaningful research question</i> Focus on personal research goals Identify a relevant research gap Identify inconsistent findings in primary studies that need to be clarified Update existing systematic reviews Evolve research question exploratory <i>Use existing guidelines</i> PICOS/PICOT PRISMA Other guidelines for systematic review Other systematic reviews
Including/excluding	<ul style="list-style-type: none"> <i>Define appropriate inclusion/exclusion criteria</i> Based on research question Based on the method Based on publication type Based on publication date Based on language <i>Identify appropriate search parameters</i> Be clear about inclusion/exclusion criteria Specify key terms and their synonyms Identify relevant databases Use librarian support Consult previous systematic reviews Conduct initial searches and adapt search parameters Limit database search to title and abstract if appropriate Conduct manual searches in relevant journals and conference proceedings Conduct forward and backward searches
Screening	<ul style="list-style-type: none"> Remove duplicates before screening Screen titles, then abstracts, then full texts Screen for specific keywords Involve multiple researchers in the screening phase
Coding	<ul style="list-style-type: none"> Apply deductive coding if an appropriate theory or framework is available Apply inductive coding to complement deductive codes or if deductive coding is not possible Assess quality of primary studies Develop a codebook Involve multiple researchers in the coding phase
Analyzing	<ul style="list-style-type: none"> Discuss patterns within the research team Answer the research question(s) Update an existing model, theory, or framework Develop a new model, theory or framework
Reporting	<ul style="list-style-type: none"> Consider journal requirements Apply reverse outlining Provide rationale for methods (e.g., for inclusion/exclusion criteria) Start results section with descriptive information Present findings by research question Start with general findings and move to more specific aspects Try to get a big picture Use visualizations and tables Provide a short summary for each chapter Combine results and discussion Keep implications for the discussion

PRISMA = Reporting Items for Systematic Reviews and Meta-analysis (Liberati et al., 2009; Page et al., 2021a, 2021b). PICOS/PICOT = population, intervention, comparator, outcome, study design, time frame (Liberati et al., 2009; Riva et al., 2012)

goal that we define within the lab” (interview 5). To define specific research questions within their area of interest, participants searched for gaps or inconsistent findings in previous literature. Some participants mentioned that research questions could also be derived from previous systematic reviews that needed to be updated. One participant also emphasized exploratory approaches to identify research questions: “My research questions are sometimes concretized only at the full text stage. I don’t do systematic reviews in a straight line” (interview 10).

To formulate their research questions and frame their systematic review, some participants used the PICOS/PICOT framework. This framework aims to help researchers include all important information in their research questions, including information about the population (P), intervention (I), comparator (C), outcome (O), study design (S), and the time frame (T; for more details on the PICOS/PICOT framework, see Liberati et al., 2009; Riva et al., 2012). Moreover, participants mentioned using other common guidelines, such as the PRISMA statement (Liberati et al., 2009; Page et al., 2021a, 2021b), to plan and document their systematic review process. In this regard, participants suggested that designing a systematic review in educational sciences should not be too different from health sciences or other disciplines: “I always go with PRISMA [...] I don’t think there are very large differences in education” (interview 4). Participants also stated that they considered other systematic reviews to get ideas for research questions, theoretical frameworks, or methods: “I have used those previous papers as models for helping me learn” (interview 12).

5.1.2 Including/Excluding

Participants pointed out that inclusion and exclusion criteria should be based on the research question(s) of the systematic review. The inclusion and exclusion criteria should lead to the selection of those studies (e.g., in terms of population, intervention, outcome) that can answer the specified research question(s). Moreover, inclusion and exclusion criteria may be based on the method of the primary studies, the publication type, publication date, or the language. Participants emphasized that researchers should provide a rationale for their inclusion and exclusion criteria as well as consider the potential consequences of their chosen criteria. For example, participants mentioned that it might be appropriate to consider only peer-reviewed articles to guarantee quality standards, but potential limitations and consequences of this decision should not be ignored. Further, participants stated that in educational sciences, it was common to reduce search results by limiting the search to specific databases or a specific time frame in order to keep the workload reasonable, even though the systematic review may then lack of comprehensiveness: “In medicine, if you would like to do a systematic review, you should scan every item on the planet. In education, it is much lighter, you can go with two database searches [...] However, they are not very comprehensive” (interview 4). Moreover, participants mentioned that the literature search could be limited to a specific time frame if a specific reason (e.g., a previously conducted systematic reviews, specific technological developments) justified that only a specific time period was considered: “I try to base on the critical event or phenomena in this field [...] We focus on MOOCs analytics studies. MOOCs became popular in 2012, and the first learning analytics conferences started in 2011” (interview 9). However, interpreting the systematic review findings should not go beyond the specified time frame to avoid incorrect conclusions.

To identify appropriate search parameters, participants emphasized the importance of clear inclusion and exclusion criteria, as search parameters should be based on these

criteria. Further, participants recommended specifying all key terms that are central to the research question(s) and all possible synonyms. To identify all relevant terms and databases, participants recommended using librarian support or consulting previous systematic reviews in the research area. Participants also conducted initial searches and adapted their search parameters in iterative processes: “We debate our strategies and keywords, we try them and we look at what we get” (interview 4). One participant also pointed out that she usually limited database searches to title and abstract to get the most relevant publications and to limit the number of search results. Moreover, participants emphasized that database searches should be complemented with manual as well as forward and backward searches.

5.1.3 Screening

Most participants started the screening phase with removing duplicates (e.g., using a reference management software) and then moved on to screening titles, abstracts, and full texts. Some participants mentioned that they focused on specific keywords (e.g., specific outcomes of interest) to accelerate the screening process: “We have some keywords and I start looking for those keywords, basically in the abstract, titles, and the keywords themselves” (interview 1). Moreover, participants emphasized that multiple researchers should be involved in the screening process: “We usually take [...] the first twenty or thirty articles, and [...] each team member codes or screens them separately. We discuss our different decisions [...] as long as we have good interrater reliability” (interview 6).

5.1.4 Coding

Participants applied both deductive and inductive coding procedures. Some participants mentioned that they usually coded the quality of the primary studies to either exclude low-quality studies from their systematic review or to estimate how the quality of the primary studies might affect their findings: “We always go with few indicators of the quality of the papers [...] Whether they reported sample size, effect size, implications, limitations” (interview 5). Other researchers emphasized other strategies for ensuring the quality of the included studies, such as focusing on peer-reviewed work. Moreover, participants emphasized that the coding process should be guided by a codebook describing all relevant variables and that multiple researchers should be involved in the coding process. While some participants explained that they double-coded all studies, other participants mentioned that they calculated interrater reliability for a few studies only: “We code the same articles, maybe 10%” (interview 9).

5.1.5 Analyzing

To analyze and synthesize primary studies, some participants mentioned that they usually started with identifying and discussing patterns within their research team: “We talked about the patterns that we were seeing across the studies” (interview 12). Participants emphasized that the primary goal in analyzing and synthesizing primary studies should be to answer the research question(s). Thus, patterns in primary studies should be interpreted based on the research questions. To answer the research question(s), a systematic review may update an existing or develop a new model, theory, or framework, depending on whether the coding has been conducted deductively or inductively. If the coding has been conducted inductively, researchers should use the overarching model, theory, or

framework to synthesize studies: “I think the framework that you adopt and use in your review can be really helpful in terms of how to synthesize these papers” (interview 11). If the coding has been conducted inductively, researchers should aim to build a new model, theory, or framework summarizing the trends emerging from the primary studies: “We usually try to understand what the factors are, that are emerging from traditional research and to synthesize them into something bigger, into a bigger model that we can use in further research” (interview 5).

5.1.6 Reporting

For the reporting phase, different participants mentioned different strategies. For example, one participant recommended considering the requirements of the target journal before starting to write the systematic review paper. The same participant mentioned that she usually applied reverse outlining: “Figuring out your headings or making a reverse outline is super helpful for me” (interview 12). Another participant emphasized the importance of a clear rationale for the chosen method (e.g., the inclusion and exclusion criteria).

When presenting systematic review findings, several participants mentioned that they usually started their results section with descriptive information about the included primary studies: “I always begin with study characteristics, like the publication type, the journals, the participants, the study design” (interview 3). Some participants also recommended organizing findings by research question and moving from general to more specific findings. Moreover, participants emphasized the need to get a big picture of the included studies, that is, researchers should not only report the studies included in the systematic review but synthesize them into an overarching model, theory, or framework: “Story your results, avoid mechanical reporting. So, it’s really important to find an interesting storyline that you can use to draw this interesting picture” (interview 6). Several participants mentioned that visualizations and tables might help readers understand the systematic review findings: “I typically use figures and tables to present the trends or the themes of the studies more clearly” (interview 9). Further, one participant stated that he usually used short summaries for different subsections of the results part. While one participant mentioned that she combined the results and discussion part of a systematic review paper, another participant preferred keeping the implications for the discussion part.

5.2 Best Practices for Team Collaboration, Ethical Considerations, and Technologies

An overview of the findings for the second research question are presented in Table 3.

5.2.1 Team Collaboration

Participants emphasized the importance of having reliable team members to share the heavy workload of conducting a systematic review:

Usually, one or two of us design the research query and decide which database you are going to search. Then, we usually have two people who are looking into the first set of studies for inclusion and exclusion and coding and so on. So, we usually have four or five people involved in the whole process. (interview 9)

Participants divided the workload based on projects, time availability, or experiences of team members: “Usually, the least experienced goes to things that do not need very

Table 3 Best practices for team collaboration, ethical considerations, and technologies

Aspect	Best practices
Team collaboration	Divide workload among team members (e.g., based on project, time availability, experiences) Consult team members for questioning and confirming methods Train yourself and your team members
Ethical considerations	<i>Problems with team collaboration</i> Make sure team members agree on their role <i>Bias based on search and inclusion strategies</i> Do comprehensive searches and consider including gray literature Work with a multilingual team Reveal potential bias in your paper <i>Bias based on subjective screening and reading of primary studies</i> Critically reflect and reveal your own positioning Calculate interrater reliability <i>Lack of credibility and replicability</i> Be transparent and follow reporting standards Give credit to prior work and make appropriate claims <i>Ethical issues in primary studies</i> Exclude study from systematic review Contact journal
Technologies	<i>Systematic review software</i> Covidence ^a EPPI Reviewer ^b Rayyan ^c <i>Reference management software</i> Citavi ^d EndNote ^e Zotero ^f Mendeley ^g <i>Tools for searching</i> ResearchRabbit ^h <i>Tools for screening, coding, and analyzing</i> Microsoft Excel ⁱ Spreadsheets <i>Tools for coding</i> Nvivo ^j ATLAS.ti ^k <i>Collaboration tools</i> Box ^l Dropbox ^m Google Drive ⁿ

^a<https://www.covidence.org/>, ^b<https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2914> ^c<https://www.rayyan.ai/>
^d<https://www.citavi.com/de/> ^e<https://endnote.com/> ^f<https://www.zotero.org/> ^g<https://www.mendeley.com/>
^h<https://www.researchrabbit.ai/> ⁱ<https://www.microsoft.com/de-de/microsoft-365/excel/> ^j<https://www.nvivo.de/> ^k<https://atlasti.com/atlas-ti-desktop/> ^l<https://www.box.com/> ^m<https://www.dropbox.com/de/> ⁿ<https://www.google.com/intl/de/drive/>

much critical in-depth analysis” (interview 4). Moreover, participants mentioned that team members played important roles in questioning and confirming ideas and methods: “I also do involve them in the consultation. Like to confirm the questions and also how we will do these reviews” (interview 1). Participants emphasized the importance of providing comprehensive training to team members, so that everyone has the same understanding of the systematic review process, the planned search, screening, and coding

procedures, as well as the key constructs related to the research question(s). Therefore, participants recommended encouraging all team members to attend workshops on conducting systematic reviews and holding regular team meetings: “We all have meetings and we take a very rigorous look at how we will approach the methodology” (interview 4). In this regard, participants recommended solving disagreements between team members by in-depth discussions: “We check whether we agree. Then, we come back and have discussions and say ‘why do you include this?’ And then, the agreement comes little by little” (interview 7).

5.2.2 Ethical Considerations

Five ethical considerations related to systematic reviews were identified in the interview study. First, participants mentioned ethical considerations related to team collaboration, such as team members not fulfilling their duties: “Well, we had a case where one of the leading reviewers dropped two days before submission” (interview 1). To avoid problems related to team collaboration, participants recommended that all team members should agree in advance on their role within the systematic review process: “Make sure that everyone in the team [...] has agreed on their role” (interview 12).

Second, participants mentioned ethical considerations related to bias caused by the search and inclusion strategies. For example, if only peer-reviewed articles or publications in a specific language are considered, this may cause bias in systematic review findings. Therefore, participants advised to carry out comprehensive searches in multiple databases and to think carefully about the advantages and disadvantages of including gray literature. Participants also pointed out that working in a multilingual team could enable the inclusion of publications in different languages: “I feel like I’m missing a lot of great work in other languages, and that I don’t know how to resolve that unless I have a team of people who can speak all the languages in the world” (interview 12). However, as time and human resources are limited, participants mentioned that bias caused by restrictive search and inclusion strategies could never be completely prevented. Therefore, researchers should point out limitations and potential biases in their systematic review paper: “It is a balancing act that you have to be aware of and that you have to reflect on when you write your systematic review” (interview 6).

Third, participants mentioned that bias in systematic reviews might also be caused by subjective screening and reading of the included primary studies. Therefore, researchers should critically reflect on their own position in relation to the research topic and reveal potential bias in their paper: “We must critically and ethically reflect on and disclose our own positioning in relation to the topic or intervention” (interview 6). Moreover, participants emphasized that multiple researchers should be included in the screening and coding process to reduce bias.

Fourth, participants mentioned ethical considerations related to systematic reviews that are not credible or cannot be replicated. Participants recommended being as transparent as possible when writing a systematic review paper and to publish the paper as well as the research methods (Open Access):

I think that we really need to be as transparent as possible. We need to provide our full data extraction coding tool, we need to provide the databases, we need to provide the full search string. Just be as transparent as possible, also in terms of ethical considerations. I always try to publish Open Access. (interview 3)

Moreover, participants mentioned that researchers should follow clear reporting standards (e.g., PRISMA, Liberati et al., 2009; Page et al., 2021a, 2021b) and carefully scrutinize the implications they derived from their systematic review findings: “Each step in the review process should be transparent, rigorous and documented in the protocol” (interview 6).

Fifth, participants pointed out that ethical considerations in systematic reviews might also stem from ethical issues, such as plagiarism, in the included primary studies. Participants were unsure how to deal with such ethical issues. One participant advised either excluding the primary study from the systematic review or contacting the journal that had published the study: “So, one option would be to exclude this paper without bringing this up somehow. Another option would be to contact the journal where the paper has been published to get them aware that there may be some issues” (interview 1).

5.2.3 Technologies

Various technologies were used in different phases of the systematic review process. Some participants used specific systematic review software for managing, screening, and coding primary studies. For example, one participant perceived EPPI reviewer (<https://eppi.ioe.ac.uk/cms/>) as beneficial for the coding phase, as it provides automatic clustering and classification of studies: “Especially for the coding phase [...] based on the codes, it generates automatic reports” (interview 6). Other participants emphasized Rayyan (<https://www.rayyan.ai/>):

It’s free and it gives everybody access and enables us, for example, to search for keywords, to see the title of the articles, to have a track record of what everybody has agreed on, to label the article: that’s why we’ve got it in, that’s why we’ve got it out. (interview 4)

For managing search results and removing duplicates, participants also used reference management software. Moreover, one participant mentioned using ResearchRabbit (<https://www.researchrabbit.ai/>) for searching studies and identifying research questions: “There are some tools coming out recently, such as ResearchRabbit [...] I couldn’t find them useful for conducting the core study. More like understanding what is there and helping to frame those big questions” (interview 5). Moreover, several participants used Microsoft Excel and spreadsheets for screening, coding, and analyzing as well as qualitative data analysis software for coding studies: “We usually use Excel to create a codebook. In each column, we have for example journal name, author name, research methods, the research setting, sample size, or something like that” (interview 9). Participants also emphasized tools that enable users to share files with others (e.g., Dropbox, Google Drive) in order to collaborate with team members in the different phases of the systematic review process.

5.3 Perspectives on Benefits, Challenges, and the Future Outlook of Systematic Reviews

An overview of the findings for the third research question are presented in Table 4.

Table 4 Perspectives on benefits, challenges, and the future outlook of systematic reviews

Aspect	Findings
Benefits	Systematic overview of the research area Replicability Helpful for advancing own research Especially helpful for early career researchers (e.g., PhD students) Are cited frequently Potential to impact practice
Challenges	Workload Narrowing the scope Few studies published on the topic Coming up with meaningful findings Setting guidelines for team collaboration Topics outside the researcher's field of expertise Long publication time Keeping up to date with the latest papers
Outlook	Increasing number of systematic reviews Need to increase methodological rigor Increasing role of (artificial intelligence-based) technologies for conducting systematic reviews along with human analysis

5.3.1 Benefits

Several participants mentioned that systematic reviews offered great benefits because they helped getting a systematic and replicable overview of a specific research area. Participants also pointed out that systematic reviews could help researchers advance their own research by identifying research gaps and providing a rationale for new primary studies: "We explore the literature in a systematic way and we can identify gaps and priority areas in a research field. Thereby, we can provide a very strong rationale for our own primary research" (interview 6). In this regard, participants mentioned that systematic reviews could be especially beneficial for early career researchers, such as PhD students. Moreover, participants pointed out that systematic reviews were usually cited frequently: "There is a real benefit just from [...] earning citations" (interview 11). Finally, participants mentioned that systematic reviews could help shape practice by providing solid recommendations based on a large number of empirical studies.

5.3.2 Challenges

Several participants stated that the main challenge of conducting systematic reviews was the heavy workload: "It needs a lot of time. I think this is the biggest challenge" (interview 2). Moreover, several participants mentioned challenges related to narrowing the scope if there were too many search results, or extending the searches if there were only a few studies published on the topic. Some participants also pointed out that it might be challenging to come up with meaningful findings: "If you really want to contribute something new with your review, you can't just report numbers of studies and what they have done" (interview 12). Another challenge relates to team collaboration: "It is really important to have clear instructions of who is doing what so that you don't end up doing everything on your own" (interview 1). One participant mentioned that it could be challenging if the systematic

review uncovered issues outside the researcher's field of expertise. Some participants mentioned that the peer-review process for systematic review papers could take a long time, making it hard to get a systematic review published. In this regard, it might be challenging to keep up with the latest papers, as new relevant primary studies might be published while a systematic review is being conducted and published.

5.3.3 Outlook

Most participants assumed that the number of systematic reviews in educational sciences would continue to increase. Participants saw a strong need to increase the methodological rigor of systematic reviews: "It is really important that we improve the quality of the work that is being published [...] There is too much duplication [...] and not enough forethought" (interview 3). Moreover, participants stated a trend towards more (artificial intelligence-based) technologies for supporting systematic reviews along with human analysis. Participants stated that artificial intelligence would play an increasing role in the search and synthesis of studies: "Things like machine learning are already able to help us conduct our searches and keep our searches up-to-date already" (interview 3). However, participants doubted that the systematic review process could be fully automated in the near future and highlighted the importance of human-machine-collaboration: "It's not completely automated and I don't think that's going to happen in the near future. I also think it's really important that there is that aspect of human synthesis and connection of different concepts" (interview 3).

6 Discussion

Despite the increase in systematic reviews in educational sciences, the methodological approach can be challenging (Alexander, 2020; Reed et al., 2005). Consequently, the field of education may face the risk of publishing systematic reviews which do not meet the needed methodological rigor (Alexander, 2020). Therefore, this interview study investigated best practices for conducting systematic reviews from the perspective of experienced systematic review researchers in educational sciences.

6.1 Interpretation of Results and Theoretical Implications

6.1.1 DISCAR Process

The findings of this interview study highlight the importance of a systematic, comprehensive, and transparent research process as outlined in the DISCAR process (West & Martin, 2023). For the designing phase, the participants emphasized the importance of including meaningful research questions which provide an original contribution to the scientific community and using existing guidelines such as PICOS/PICOT (Liberati et al., 2009; Riva et al., 2012) or PRISMA (Liberati et al., 2009; Page et al., 2021a, 2021b). Though used more commonly in health sciences research, our interview study shows that both the PRISMA and PICOS/PICOT frameworks have been found beneficial while conducting systematic reviews in educational sciences. In this regard, our interview findings did not reveal large differences between the designing phase of systematic reviews in educational sciences compared to health sciences or other disciplines.

For including/excluding, participants highlighted the need for appropriate inclusion and exclusion criteria and search parameters. Several of the best practices discussed for the including/excluding phase (e.g., specifying key terms and their synonyms, using librarian support) have also been discussed in previous guidelines (e.g., Alexander, 2020; Higgins & Green, 2008b). However, while existing guidelines have mainly focused on the completeness of results when defining strategies for including/excluding studies (e.g., Higgins & Green, 2008b), the participants in our interview study also emphasized feasibility related to the workload. For example, they recommended limiting database searches to title and abstract or specific years to reduce the number of search results, as comprehensive searches can lead to a very large number of search results due to unclear definitions and multi-faceted concepts in educational sciences (Zawacki-Richter, 2020).

For screening, participants recommended de-duplication, screening by title, abstract, and full text as well as screening for specific keywords and involving multiple researchers in the screening process, which do not differ from existing guidelines from health sciences or other disciplines (Alexander, 2020; Higgins & Green, 2008b; Liberati et al., 2009). Although the participants in our interview study mentioned that (artificial intelligence-based) technologies would increasingly shape the systematic review process, they only discussed manual screening. Other researchers (e.g., Brunton et al., 2017) discussed how text mining can be used in the study selection process by training computer software to screen studies in addition to manual screening.

For coding, participants discussed using both inductive and deductive coding and emphasized quality appraisal, developing a codebook, and involving multiple researchers. Quality appraisal has been considered an important aspect of the systematic review process and can focus on elements such as study design, study methods, and study relevance to the review question (Gough et al., 2017; Petticrew & Roberts, 2005). However, quality appraisal as well as the subsequent decisions made for the systematic review process are subject to the researchers' subjective perceptions and may introduce researcher bias. Therefore, the criteria for the quality appraisal as well as the handling of primary studies of poor quality must be precisely selected and clearly documented (Okoli & Schabram, 2010). Although tools for assessing the quality of primary studies exist (e.g., CASP, 2024; Higgins & Altman, 2008a), quality appraisal may be challenging, as these tools may not apply to all methods and publication types. Therefore, some of the participants in our interview study preferred ensuring quality of primary studies through appropriate inclusion and exclusion criteria (e.g., by focusing on peer-reviewed work). However, this approach also might have limitations, as peer-review cannot always be considered an indicator of quality and important contributions might be ignored (Paez, 2017).

For analyzing, participants mainly discussed best practices, such as discussing patterns according to research questions or updating models, theories, or frameworks that have already been addressed in previous guidelines for systematic reviews (e.g., Alexander, 2020; Hallinger, 2013; West & Martin, 2023). For reporting, participants used different strategies (e.g., reverse outlining, providing descriptive information), with some of these strategies contradicting each other (e.g., combining results and discussion section vs. keeping implications for the discussion). While all these strategies may help researchers report the systematic review findings in a meaningful manner, it is up to the researchers to decide which strategies are most suitable depending on their review questions and expertise. Previous guidelines for reporting systematic reviews have mainly focused on reporting standards and formality issues such as standards for presenting the screening phase or systematic review findings (e.g., Page et al., 2021a, 2021b). In contrast, our interview study provides

best practices (e.g., apply reverse outlining, start results section with descriptive information) for achieving these reporting standards.

To summarize, with a few exceptions in the including/excluding and reporting phase, the best practices for the DISCAR process identified in our interview study are very similar to existing guidelines from health sciences or other disciplines (Alexander, 2020; Higgins & Green, 2008b; Kitchenham, 2004). This may indicate that the conduct of systematic reviews in educational sciences is not significantly different from other disciplines. However, our findings may also indicate that best practices for the field of educational sciences have not been established yet, as researchers predominantly rely on existing guidelines from other disciplines.

6.1.2 Team Collaboration, Ethical Considerations, and Technologies

Our interview study is the first empirical study that investigated systematic review researchers' perspectives on team collaboration, ethical considerations, and technology usage when conducting systematic reviews in educational sciences. For team collaboration, the participants in our interview study recommended the importance of dividing workload, consulting team members for specific expertise, and training. Systematic reviews in educational sciences are usually varied in scope and large scoped reviews need a team to carry it out effectively and in a timely manner (Zawacki-Richter, 2020). Moreover, to reduce subjective bias and to calculate interrater reliability, a research team is needed (West & Martin, 2023). However, challenges can arise with team collaboration (Okoli & Schabram, 2010) and, hence, the roles and responsibilities of each team member has to be established early in the project.

Moreover, participants emphasized the need for a stronger focus on ethical considerations in systematic reviews, such as biases and ethical issues in primary studies. Several biases were discussed based on searching, inclusion, and screening. These biases have already been recognized in previous research (e.g., Higgins & Green, 2008b; Liberati et al., 2009; Suri, 2020). Several best practices for reducing biases in systematic reviews were recommended, such as including gray literature to reduce publication bias, working with a multilingual team to reduce language bias, calculating interrater reliability, as well as reflecting and reporting on researcher positionality. In addition to biases in systematic reviews, further ethical considerations related to problems with team collaboration, lack of credibility and replicability, and ethical issues in primary studies were identified. The researchers interviewed in our study emphasized the importance of transparency and following reporting guidelines, which is critical for future researchers to replicate.

A number of technologies were discussed in our interviews which assist in conducting systematic reviews efficiently. Technologies were used for different steps of the systematic review process, such as screening, coding, or analyzing. Some of the participants used artificial-intelligence based technologies such as EPPI Reviewer (<https://eppi.ioe.ac.uk/cms/>) or ResearchRabbit (<https://www.researchrabbit.ai/>). However, although participants assumed an increasing role of artificial intelligence-based technologies in the systematic review process, the current use of such technologies to (partially) automate different steps of the systematic review process is rather less than in other disciplines (de la Torre-López et al., 2023; Scott et al., 2021). These findings are consistent to Bond et al.'s (2023, 2024) meta systematic reviews on educational technology and artificial intelligence in education where only a small amount of the included papers (3.7–12.1%) used artificial intelligence-based technologies to support the (systematic) review process.

6.1.3 Benefits, Challenges, and Outlook

The findings on the benefits, challenges, and future outlook of systematic reviews in educational sciences provide motivation for additional researchers to conduct systematic reviews, while being aware of the challenges when working on these projects. Generally, systematic reviews were seen as an appropriate research methodology in educational sciences, helping to advance research and practice. However, some issues such as the heavy workload, too many or too few primary studies on the topic, or coming up with meaningful findings still remain key challenges when conducting systematic reviews in educational sciences. According to the participants in our interview study, the number of systematic reviews in educational sciences will continue to increase in future, as will the number of technologies to support the systematic review process.

6.2 Limitations and Future Research

The sample of our interview study could have included more participants and further differentiate the different levels of experience in conducting systematic reviews. Moreover, our findings may be subject to self-selection bias (Heckmann, 1990), as no random sampling methods were applied. Further, our participants were located in Europe and the United States, which limits the generalizability of our findings. Future research should aim to sample the participants from larger geographic locations.

The interviews covered several different aspects (e.g., designing, screening, team collaboration, ethical considerations) that might have required more time and follow-up questions than provided in our interviews to obtain in-depth information. Since we wanted to keep the duration of the interviews within reasonable limits for the participants, some participants only provided superficial answers to the questions. Therefore, future research should focus on specific aspects of the systematic review process rather than investigating the whole systematic review process in one study. Moreover, the identified best practices do not represent an exhaustive list of all possible best practices but only the ones discussed by the researchers in our interview study. In addition, although we combined deductive and inductive coding, the predefined categories of our deductive coding system might have led to some new or unexpected findings being overlooked. Future studies could add to our list of best practices both through primary research and through analyzing systematic reviews that have been conducted. In this regard, future research should also conduct quantitative studies to identify the most relevant best practices. In addition, the challenges discussed in our interview study (e.g., workload) highlight aspects for which best practices still need to be established.

6.3 Implications for Systematic Reviews in Educational Sciences

Our interview study has implications for researchers who plan to conduct systematic reviews, instructors who teach the systematic review process to their students, as well as editors and reviewers who review and publish systematic reviews. Further, the findings from this study have implications for improving the quality of systematic reviews being conducted. The identified best practices (see Tables 2, 3 and 4) can be used by researchers as guidance when planning and conducting their systematic review as well as by journal editors and reviewers to evaluate the quality of systematic review papers.

While most of the best practices identified for the different phases of the DISCAR process (West & Martin, 2023) are similar to those reported in previous guidelines (e.g., Higgins & Green, 2008b; West & Martin, 2023), our interview study highlights those best practices that are considered particularly useful for educational sciences. When those best practices are implemented, they improve the overall readability, reproducibility, and scholarly value of the systematic review, making it a more effective tool for advancing knowledge. The best practices collectively enhance the rigor, quality and credibility of the systematic review, leading to more reliable and comprehensive findings and their contributions to the field. Some implications for the different phases of the systematic review process based on the best practices identified in our interview study include:

6.3.1 Designing

Identifying a meaningful research question, aligned with personal research goals and relevant gaps in the literature, can ensure that the systematic review addresses critical areas in the field (West & Martin, 2023). By focusing on clarifying inconsistent findings in primary studies, researchers can provide clarity where previous research has been inconclusive, providing a big picture of the mixed findings. Updating existing systematic reviews highlights the importance of maintaining the relevance of the review as new studies emerge (Siddaway et al., 2019). Evolving research questions in an exploratory manner enables the researcher to refine the focus during the systematic review process. Using existing frameworks, along with insights from other systematic reviews, allows researchers to build on established practices (Okoli & Schabram, 2010).

6.3.2 Including/Excluding

Researchers need to carefully define inclusion/exclusion. By setting various criteria such as research questions, study methods, publication type, date, and language, researchers can selectively include studies that are most relevant to their specific inquiry. This inclusion helps manage the scope of the review while ensuring that the chosen studies adequately address the research questions (Alexander, 2020; Okoli & Schabram, 2010). Being explicit about inclusion and exclusion criteria ensures transparency and reproducibility (Liberati et al., 2009; Page et al., 2020b). Specifying key terms and their synonyms enhances search accuracy, while consulting previous systematic reviews and using librarian support can help identify the most relevant databases and search strategies. Conducting preliminary searches allows for refining these parameters iteratively, ensuring the search is comprehensive yet manageable. Limiting searches to titles and abstracts when appropriate helps streamline the search process, but should be balanced with manual searches in key journals and conference proceedings to capture potentially overlooked studies. Forward and backward searches can uncover additional studies, ensuring a more exhaustive review (Page et al., 2021b).

6.3.3 Screening and Coding

Researchers are recommended to prioritize consistency during the collaborative screening and coding process involving multiple screeners and coders. Involving multiple researchers in the screening and coding phase promotes objectivity, reduces individual bias, and increases interrater reliability, leading to more accurate and reproducible results (Shaffril

et al., 2021; Siddaway et al., 2019). Researchers are recommended to deduplicate and then screen at various levels, title, abstract and full text (West & Martin, 2023). During coding, they are recommended to use both deductive and inductive coding methods. Applying deductive coding allows researchers to systematically validate existing concepts while inductive coding provides room for new themes and insights to emerge (Mayring, 2015). Researchers are recommended to develop a codebook and conduct quality appraisal of primary studies. Developing a codebook helps maintain consistency across the coding process, ensuring all researchers interpret and apply codes uniformly (West & Martin, 2023). Assessing the quality of primary studies ensures that only quality research informs the review's conclusions, thereby enhancing the credibility of the findings (Siddaway et al., 2019).

6.3.4 Analyzing

Researchers are recommended to not just provide a list of findings but contribute to a larger theoretical or conceptual understanding of the topic (Alexander, 2020). When team members engage in dialogue about the patterns emerging from the data, it helps refine interpretations, promote diverse perspectives, and ensure a more comprehensive understanding of the results. Answering the research question(s) is a central goal of any systematic review, and the patterns identified should directly address these questions (Alexander, 2020; West & Martin, 2023). Updating an existing model, theory, or framework has the implication of advancing knowledge by integrating new findings into established frameworks. This ensures that the review contributes to refining or extending their applicability based on new evidence (Alexander, 2020; Okoli & Schabram, 2010). Developing a new model, theory, or framework as a result of the systematic review has far-reaching implications by introducing new perspectives that may better explain the patterns found in the data and can shape future research and practice (West & Martin, 2023).

6.3.5 Reporting

The best practices recommended for reporting in systematic reviews can significantly enhance the clarity, and coherence of the research (Pager et al., 2021a, 2021b). By considering journal requirements, researchers ensure their work meets publication standards, increasing the likelihood of acceptance. Applying reverse outlining helps structure the paper logically, while providing clear rationales for inclusion/exclusion criteria strengthens the transparency and credibility of the review. Presenting results by research question and starting with general findings before moving to specifics allows readers to follow the research progression intuitively. Using visualizations and tables makes complex data more accessible and easier to interpret (Alexander, 2020), while short summaries help to reinforce key points. Combining results and discussion fosters a more integrated interpretation of findings, while keeping implications in the discussion section allows for a deeper reflection on how the results contribute to the field.

6.3.6 Team Collaboration

Systematic reviews in educational sciences require a collaborative team to distribute the workload effectively (Borah et al., 2017; Zawacki-Richter, 2020). Such a team may have expertise in different aspects of systematic reviews, including literature search, data

extraction, and synthesis. However, clear roles and responsibilities of each team member are required to avoid duplication of effort and ensure timely progress. Further, fostering open communication and collaboration among team members encourages scientific discussion, problem-solving, and collective learning. Regularly evaluating and adjusting team dynamics is essential to address any potential issues, maintaining optimal collaboration, and ensuring the smooth operation of the systematic review process (O'Dwyer & Wafford, 2021).

Systematic reviews in educational sciences require extensive training on systematic review methodology which is essential to ensure that all team members have a thorough understanding of the systematic review's objectives and methodological approach (Okoli & Schabram, 2010). Such training also includes using appropriate technologies, such as search engines, reference management software, and coding tools, which empower team members to manage the systematic reviews process efficiently and effectively.

6.3.7 Ethical Considerations

Systematic reviews in educational sciences should address ethical considerations through clear reporting standards and carefully scrutinize the implications drawn from the findings (Suri, 2020). Bias caused by restrictive search and inclusion strategies should be minimized by conducting comprehensive searches in multiple databases. Subjective screening and coding should be mitigated by involving multiple researchers and critically reflecting on potential biases.

6.3.8 Technologies

Systematic reviews in educational sciences should utilize adequate technologies for managing, screening, and coding primary studies (de la Torre-López et al., 2023; Tsafnat et al., 2014). Reference management software can help manage search results and remove duplicates. Further, qualitative data analysis software can support the coding of primary studies as well as summarizing findings. In addition, social collaboration tools can streamline communication and documentation among research team members. The use of artificial intelligence-based tools may be critically investigated in order to support the scientific rigor of systematic reviews. Accordingly, the researchers' synthesis and connection of concepts will remain essential in systematic reviews despite the use of artificial intelligence-based tools (de la Torre-López et al., 2023; Jardim et al., 2022).

6.4 Conclusion

Our study provides evidence-based best practices for conducting systematic reviews in educational sciences. By implementing and considering these best practices when conducting, reporting, or reviewing systematic reviews in educational sciences, researchers can enhance the quality and effectiveness of systematic reviews, contributing to a more robust and evidence-based foundation for research and practice. While the number of systematic reviews in educational sciences is expected to increase further, there is a strong need to guarantee credible and replicable systematic reviews for advancing the knowledge of the scientific community.

Appendix

Interview Protocol

In this interview, we focus on the broader secondary research area – systematic reviews of research, and not on meta-analyses. A systematic review is a type of secondary research that aims to identify, evaluate, and synthesize primary research in a transparent and analytical manner to draw conclusions based on the evidence (Martin et al., 2020).

1. Tell us about your experience conducting systematic reviews.
2. How do you design or frame a systematic review of research?
Sample follow-up question: How do you find appropriate research questions?
3. What makes the review original or how does it contribute to the scientific community?
Sample follow-up question: How do you identify the need for a new systematic review?
4. How do you use a framework or guideline to conduct your systematic reviews?
Sample follow-up question: How does this framework/guideline help you conduct systematic reviews?
5. What roles do your team members play during the systematic review process?
Sample follow-up question: What training does each person receive?
6. Describe some strategies you use for including and excluding studies in a systematic review of research.
Sample follow-up question: How do you choose an appropriate time frame for study inclusion?
7. Describe some strategies you use for searching studies in a systematic review of research to result in high quality outcomes.
Sample follow-up question: What resources do you use for searching?
8. Describe some strategies you use for screening studies during a systematic review of research.
9. Describe some strategies you use for coding studies during systematic review of research.
10. Describe some strategies you use for analyzing and synthesizing studies in systematic review of research.
11. Describe strategies on how you present or report findings in a systematic review of research?
12. What are some ethical considerations researchers have to make while conducting systematic reviews of research?
13. What technologies do you use in the various phases while conducting a systematic review of research?
14. What are some benefits of conducting systematic reviews of research?
15. What are some challenges you have faced in conducting systematic reviews of research?

16. What is your perspective on the outlook for future systematic reviews in education?

Acknowledgements We thank Julia Spuck for her assistance with interview transcriptions.

Funding Open Access funding enabled and organized by Projekt DEAL. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability Anonymized data will be made available on justified request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research*, *90*(1), 6–23. <https://doi.org/10.3102/0034654319854352>
- Blaizot, A., Veettil, S. K., Saidoung, P., Moreno-Garcia, C. F., Wiratunga, N., Aceves-Martins, M., Lai, N. M., & Chaikyakunapruk, N. (2022). Using artificial intelligence methods for systematic review in health sciences: A systematic review. *Research Synthesis Methods*, *13*, 353–362. <https://doi.org/10.1002/jrsm.1553>
- Bond, M., Khosravi, H., Bergdahl, N., Buntins, K., De Laat, M., Oxley, E., Negrea, V., Chong, S. W., & Händel, M. (2023). Digital evidence synthesis tools in educational technology research: A systematic mapping review. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.30594.25288>
- Bond, M., Khosravi, H., de De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, *21*, 1–42. <https://doi.org/10.1186/s41239-023-00436-z>
- Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, *7*(2), e012545. <https://doi.org/10.1136/bmjopen-2016-012545>
- Brunton, J., Graziosi, S., & Thomas, J. (2017). Tools and techniques for information management. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 154–180). Sage.
- CASP. (2024). CASP checklists. Retrieved December, 19, 2024, from <https://casp-uk.net/casp-tools-checklists/>
- de la Torre-López, J., Ramírez, A., & Romero, J. R. (2023). Artificial intelligence to automate the systematic review of scientific literature. *Computing*, *105*, 2171–2194. <https://doi.org/10.1007/s00607-023-01181-x>
- Dowd, A. C., & Johnson, R. M. (2020). Why publish a systematic review: An editor's and reader's perspective. In O. Zawacki-Richter, M. Kerres, S. - Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research: Methodology, perspectives, application* (pp. 69–87). Springer.
- Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). Sage.
- Gough, D., Oliver, S., & Thomas, J. (2017). Introducing systematic reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (2nd ed., pp. 1–18). 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, 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Greyson, D., Rafferty, E., & Slater, L. (2019). Systematic review searches must be systematic, comprehensive, and transparent: A critique of Perman et al. *BMC Public Health*, 19(153), 19–153. <https://doi.org/10.1186/s12889-018-6275-y>
- 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, application* (pp. 23–40). Springer.
- Heckman, J. J. (1990). Selection bias and self-selection. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Econometrics* (pp. 201–224). Springer.
- Higgins, J. P. T., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions*. Wiley.
- Higgins, J. P. T., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 187–235). Wiley.
- Jardim, P. S. J., Rose, C. J., Ames, H. M., Echavez, J. F. M., Van de Velde, S., & Muller, A. E. (2022). Automating risk of bias assessment in systematic reviews: A real-time mixed methods comparison of human researchers to a machine learning system. *BMC Medical Research Methodology*, 22, 1–22. <https://doi.org/10.1186/s12874-022-01649-y>
- Kitchenham, B. (2004). *Procedures for performing systematic reviews*. Keele University. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=29890a936639862f45cb9a987dd599dce9759bf5>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology*, 62, 1–34. <https://doi.org/10.1016/j.jclinepi.2009.06.006>
- Mallett, R., Hagen-Zanker, J., Slater, R., & Duvendack, M. (2012). The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness*, 4(3), 445–455. <https://doi.org/10.1080/19439342.2012.711342>
- Martin, F., Dennen, V. P., & Bonk, C. J. (2020). A synthesis of systematic review research on emerging learning environments and technologies. *Educational Technology Research and Development*, 68, 1613–1633. <https://doi.org/10.1007/s11423-020-09812-2>
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken [Qualitative content analysis: Foundations and techniques]* (12th ed.). Beltz.
- Needleman, I. G. (2022). A guide to systematic reviews. *Journal of Clinical Periodontology*, 29, 6–9. <https://doi.org/10.1034/j.1600-051X.29.s3.15.x>
- 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.
- O'Dwyer, L. C., & Wafford, Q. E. (2021). Addressing challenges with systematic review teams through effective communication: A case report. *Journal of the Medical Library Association*, 109(4), 643–647. <https://doi.org/10.5195/jmla.2021.1222>
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Working Papers on Information Systems*, 10(26), 1–49.
- Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10(3), 233–240. <https://doi.org/10.1111/jebm.12266>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021a). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, e105906. <https://doi.org/10.1016/j.ijsu.2021.105906>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021b). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *British Medical Journal*, 372, 1–36. <https://doi.org/10.1136/bmj.n160>
- Petticrew, M., & Roberts, H. (2005). *Systematic reviews in the social sciences: A practical guide*. Wiley. <https://doi.org/10.1002/9780470754887>

- Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, 90(1), 24–26. <https://doi.org/10.3102/0034654319877153>
- Reed, D., Price, E. G., Windish, D. M., Wright, S. M., Gozu, A., Hsu, E. B., Beach, M. C., Kern, D., & Bass, E. B. (2005). Challenges in systematic reviews of educational intervention studies. *Annals of Internal Medicine*, 142(12), 1080–1089. https://doi.org/10.7326/0003-4819-142-12_Part_2-200506211-00008
- Review of Educational Research. (2023). *Submission guidelines*. <https://journals.sagepub.com/author-instructions/RER>
- Riva, J. J., Malik, K. M., Burnie, S. J., Endicott, A. R., & Busse, J. W. (2012). What is your research question? An introduction to the PICOT format for clinicians. *The Journal of the Canadian Chiropractic Association*, 56(3), 167–171.
- Scott, A. M., Forbes, C., Clark, J., Carter, M., Glasiou, P., & Munn, Z. (2021). Systematic review automation tools improve efficiency but lack of knowledge impedes their adoption: A survey. *Journal of Clinical Epidemiology*, 138, 80–94. <https://doi.org/10.1016/j.jclinepi.2021.06.030>
- 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, 1319–1346. <https://doi.org/10.1007/s11135-020-01059-6>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Suri, H. (2020). Ethical considerations of conducting systematic reviews in educational research. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research: Methodology, perspectives, application* (pp. 41–54). Springer.
- Tsafnat, G., Glasziou, P., Choong, M. K., Dunn, A., Galgani, F., & Coiera, E. (2014). Systematic review automation technologies. *Systematic Reviews*, 3, 1–15. <https://doi.org/10.1186/2046-4053-3-74>
- West, R. E., & Martin, F. (2023). What type of paper are you writing? A taxonomy of review and theory scholarship distinguished by their summary and advocacy arguments. *Educational Technology Research and Development*, 72, 2443–2476. <https://doi.org/10.1007/s11423-023-10233-0>
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X1772397>
- Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (Eds.). (2020). *Systematic reviews in educational research: Methodology, perspectives and application*. Springer. <https://doi.org/10.1007/978-3-658-27602-7>
- Zawacki-Richter, O. (2020). Introduction: Systematic reviews in educational research. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research: Methodology, perspectives, application* (pp. V–XIII). Springer. <https://doi.org/10.1007/978-3-658-27602-7>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.