Review Article https://doi.org/10.12973/ijem.11.1.27

International Journal of Educational Methodology
Volume 11, Issue 1, 27 - 42.
ISSN: 2469-9632
https://www.ijem.com/

A Meta-Systematic Review of the Conceptual, Methodological, and Reporting Quality of Systematic Reviews of Research on Educational Leadership and Management in Africa

Anthony J. Onwuegbuzie
University of Cambridge, UNITED KINGDOM /
University of Johannesburg, SOUTH AFRICA

Ricardo Sabates*
University of Cambridge, UNITED KINGDOM

Received: December 2, 2024 ▪ Revised: January 9, 2025 ▪ Accepted: January 29, 2025

Abstract: Formal publications in the area of educational leadership and management (EDLM) can be traced back more than a century. The 1950s saw the emergence of reviews of these EDLM works. Unfortunately, these reviews were limited because they were not systematic. Sometime after the establishment of the Campbell Collaboration as the standard for systematic reviews in the social sciences in 2000, scholars began conducting systematic reviews of the EDLM literature, including the African EDLM literature. However, little is known about the quality of this latter corpus of systematic reviews. Therefore, in this article, a systematic review of systematic reviews — a meta-systematic review — was conducted on the African EDLM literature. This meta-systematic review revealed 42 systematic reviews. The quality of these reviews was assessed using Hallinger's analytical rubric for conducting systematic reviews, which led to the identification of strengths and weaknesses of these reviews, as well as a set of exemplary reviews. These findings offer actionable insights for policymakers and practitioners by highlighting the need for targeted training and resources to improve systematic reviews' conceptual, methodological, and reporting quality, thereby enhancing their utility for evidence-based decision-making and educational policy development.

Keywords: *Africa, educational leadership and management, meta-systematic review, quality criteria, systematic review.*

To cite this article: Onwuegbuzie, A. J., & Sabates, R. (2025). A meta-systematic review of the conceptual, methodological, and reporting quality of systematic reviews of research on educational leadership and management in Africa. *International Journal of Educational Methodology*, *11*(1), 27-42. https://doi.org/10.12973/ijem.11.1.27

Introduction

Conducting the literature review represents a vital component of the research process in all empirical research studies (Onwuegbuzie & Frels, 2016). Its importance stems from the fact that it provides the most comprehensive and effective way of becoming familiar with previous findings as well as conceptual, theoretical, practical, and methodological frameworks used by researchers pertaining to a given topic of interest, thereby laying a foundation for research that yields meaningful and valid findings and interpretations (Onwuegbuzie & Frels, 2016). The systematic review represents an important type of literature review (Onwuegbuzie & Frels, 2016). Broadly speaking, a systematic review can be defined as "a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review" (Moher et al., 2010, p. 336). A systematic review typically includes a description of the selected body of works, and involves an integration of the findings stemming from each work to some degree.

Quality of Systematic Reviews

Hong and Pluye (2019) referred to the process of assessing the quality of systematic reviews as representing *critical appraisal*. According to these authors, the goal of critical appraisal is "to identify the strengths and weaknesses of studies, to determine how much confidence to have in the findings, and to ensure that the recommendations and conclusions properly reflect the quality of evidence reviewed" (p. 449). Hong and Pluye (2019) identified three dimensions of quality: conceptual quality, methodological quality, and reporting quality. Conceptual quality refers to the systematic review being characterized by a clarity of the underlying concept/construct for the purpose of facilitating theoretical insight, as well as a translucent understanding of the concept/phenomenon considering the depth of description presented (Toye

* **Corresponding author:**
Ricardo Sabates, Faculty of Education, University of Cambridge, 184 Hills Road, Cambridge, CB2 8PQ, United Kingdom. ✉ rs867@cam.ac.uk

et al., 2013). Simply put, conceptual quality refers to *insightfulness*—that is, to what extent does the study provide a clear, rich, and thick understanding of the phenomenon of interest?

Contrastingly, methodological quality refers to how the systematic review is conducted, specifically relating to the methodology and methods used and the extent to which biases were minimized (Rouleau et al., 2023) at every stage of the systematic review process (Whiting et al., 2016). In a nutshell, methodological quality is related to the construct of *trustworthiness* (i.e., To what extent is the systematic review sufficiently adequate for the results to be trustworthy?)—an umbrella term used to denote methodological quality criteria such as internal validity, reliability, internal credibility, truth value, applicability, consistency, neutrality, dependability, and credibility. Finally, reporting quality concerns the extent, if any, to which a systematic review report provides sufficient information about the design, conduct, analysis, and interpretation of the primary works that comprise it (Rouleau et al., 2023). Simply put, reporting quality refers to *transparency*, *accuracy*, and *completeness* (Simera et al., 2010).

Tools for Measuring the Quality of Systematic Reviews

A wide variety of critical appraisal tools and approaches have been developed (Hong & Pluye, 2019). In the vast majority of cases, the focus of each critical appraisal tool has been on one of the three measures of quality (i.e., conceptual quality, methodological quality, and reporting quality). For instance, with regard to conceptual quality, in their review of the conceptual quality of observational longitudinal studies, Beauregard et al. (2011) used two criteria, namely, analytical breadth and depth, to clarify a construct. In terms of methodological quality, several assessment tools have been developed (Rouleau et al., 2023). With respect to reporting quality, for example, The Enhancing the Quality and Transparency Of Health Research Network (The EQUATOR Network, 2019) website presents more than 400 reporting guidelines and checklists to assist reviewers in improving transparency in the reporting of various types of health research, including systematic reviews.

Unfortunately, although numerous tools have been developed to assess the conceptual, methodological, and reporting quality of systematic reviews, they have been applied in a fragmented manner. In particular, with very few exceptions—such as Rouleau et al.'s (2023) typology that is centered on the assessment of various types of systematic reviews, namely, systematic reviews of quantitative research, qualitative research, and mixed methods research—these quality assessments address research representing only a single tradition. For instance, as observed by Rouleau et al. (2023), although there are numerous tools for assessing the methodological quality of systematic quantitative reviews (e.g., Pollock et al., 2020), there is a lack of guidance for systematic qualitative reviews and systematic mixed methods reviews.

Further evidence of the disjointedness of tools for assessing the quality of systematic reviews stems from the fact that the overwhelming majority of these tools focus exclusively on *one* of the three dimensions of quality. Yet, these three quality dimensions are inextricably linked, interdependent, and "are intertwined in the critical appraisal process" (Hong & Pluye, 2019, p. 455). As an example, inadequate reporting will prevent an appropriate appraisal of both the methodological and conceptual qualities of systematic reviews (Hong & Pluye, 2019; Simera et al., 2010).

The disjointedness of tools for assessing the quality of systematic reviews also has arisen from the fact that these tools are commonly developed and used in the field of health and medicine, given the extensive use of systematic reviews in healthcare decision-making. Some of these tools have been applied to systematic reviews conducted in other fields, such as Risk of Bias in Systematic Reviews (ROBIS; Whiting et al., 2016). However, these adapted tools typically address only one of the three dimensions of quality in systematic reviews and/or address systematic reviews representing only one of the three research traditions (i.e., quantitative, qualitative, or mixed methods research). Further promoting the disjointedness of tools for assessing the quality of systematic reviews, the majority of systematic reviews tend to involve predominantly higher income countries in general and the United States and United Kingdom in particular (see, for e.g., Ehsan et al., 2019).

Purpose of Study

Because of the disjointedness of tools for assessing the quality of systematic reviews, what is needed is a more unified approach to assessing this quality. Moreover, multidimensional approaches to assessing quality are warranted. These approaches should be able to be used in fields other than the health fields. To this end, the goal of the present study was to demonstrate the utility of adapting a generic tool for assessing the quality of systematic reviews—regardless of the field and regardless of whether the systematic review was a systematic quantitative review, a systematic qualitative review, a systematic mixed methods review, or a systematic review of two or more primary research traditions. Further, in contrast to the field of health wherein the majority of systematic reviews are conducted, the current investigation involved the field of education in general and the discipline of educational leadership in particular. Also, this study occurred within a non-higher income context—specifically, the sub-Saharan Africa context.

The present study involved what has been referred to by some reviewers as a *meta-systematic review*. Broadly speaking, a meta-systematic review is a type of research synthesis in which the researcher(s) aims to identify, to analyze, to interpret, and to evaluate a corpus of systematic reviews that have been conducted on a particular research domain, topic, or research question(s)—optimally in a comprehensive and rigorous manner. In particular, the identification of

the systematic reviews should be as systematic and transparent as should occur in a standard systematic review of single works on a specific topic. After analyzing and interpreting the selected systematic review works, the findings of the systematic reviews are synthesized and summarized to provide an overall landscape of the evidence base on the underlying topic. More specifically, meta-systematic reviews can help to identify gaps in the evidence, to highlight inconsistencies or conflicting findings, and to provide a comprehensive overview of the current state of knowledge on a particular topic. Meta-systematic reviews also can be used to guide future research and to inform practice and policy decisions. An important aspect of a meta-systematic review is to evaluate the quality of the corpus of systematic reviews and to appraise critically the methods used in each systematic review. Specifically, the purpose of this study was to undertake a meta-systematic review to examine the quality of systematic reviews of the African educational leadership and management (EDLM) literature over time.

This meta-systematic review of African EDLM knowledge production was guided by the following research question: What are the patterns of strength and weakness in the quality of systematic reviews conducted by reviewers of African EDLM literature over time? It was hoped that this article would contribute to the literature by identifying the patterns of quality in the systematic reviews of the African EDLM literature, with the goal of providing evidence-based recommendations for improving the quality of systematic reviews conducted in this area and beyond.

Conceptual Framework

The conceptual framework underpinning this meta-systematic review was the "three-level analytical rubric" developed by Hallinger (2014). This analytical rubric comprises explicit statements that describe eight sets of quality indices. Each quality index is assessed via a 3-point rating scale, as follows: 0 = the quality criterion is not met, 1 = the quality criterion is partially met, and 2 = the quality criterion is fully met (see Figure 1). As stated by Hallinger (2014), this analytical rubric assists in providing insights into the relative strength and weakness of the systematic reviews based on the eight quality criteria. A particularly unique strength of this analytical rubric is that it focuses, at least in part, on all three dimensions of quality, namely, *Conceptual* (i.e., Statement of Purpose; Conceptual Framework), *Methodological* (i.e., Sources and Search Procedures; Data Extraction; Data Analysis), and *Reporting* (i.e., Presentation of Findings; Limitations of the Review; Implications of the Review). The eight sets of scores can be combined into a total score that ranges from 0 (i.e., 8 criteria x score of 0 on each quality criterion) to 16 (i.e., 8 criteria x score of 2 on each quality criterion). This total score provides an indicator of the quality of the systematic review.

Table 1. A Rubric for the Critical Evaluation of Systematic Reviews

		Does Not Meet Standard	**Partially Meets Standard**	**Meets Standard**
#	**Criteria/Level**	**0**	**1**	**2**
1	Statement of Purpose	There is no clear definition of the research problem or questions behind the review.	The reviewer has articulated a topical focus, but this is not clearly defined in terms of research goals, outcomes, or questions.	The research problem and specific research goals or questions are clearly articulated with appropriate rationale for its importance.
2	Conceptual Framework	There is no conceptual framework used in the review and no justification for its omission.	The review applies a conceptual framework, but it lacks either articulation or justification.	An explicit conceptual framework to guide the review is articulated and justified or a clear rationale is offered for why a conceptual framework is not used.
3	Sources and Search Procedures	There is no discussion of source selection procedures or rationale.	Either the sources used in the review are not described and justified, or the procedures used to identify the specific set of sources are unclear.	Sources and procedures used to identify them are clearly described and justified.
4	Data Extraction	Procedures for extracting and evaluating information from the studies are not discussed and are unclear to the reader.	Procedures for extracting evaluating information from the studies are implicit but can be ascertained by the reader.	Procedures for extracting evaluating information from the studies are clearly stated.
5	Data Analysis	Procedures for analyzing and synthesizing data from the studies are unknown to the reader.	Procedures for analyzing and synthesizing information from the studies are implicit but can be ascertained by the reader.	Procedures for analyzing and synthesizing information from the studies are clearly stated and executed.

Table 1. Continued

#	Criteria/Level	Does Not Meet Standard 0	Partially Meets Standard 1	Meets Standard 2
6	Presentation of Findings	Presentation of findings does not clarify how the results advance our understanding of the research problem.	Presentation of findings emphasizes analysis more than synthesis and/or only partially clarifies how the results advance our understanding of the research problem.	Synthesizes findings across the studies and clearly communicates what was learned and how this advances understanding of the research problem.
7	Limitations of the Review	No explicit discussion of how the findings are limited by the methodology of the review.	Limitations of the review are mentioned but not directly linked to the interpretation of results.	Limitations of the review are described and linked to interpretation of results.
8	Implications of the Review	No explicit discussion of implications.	Discussion of implications could be vague, overstated, or incomplete (i.e., omits implications for a relevant audience).	Comprehensive set of implications is described for all relevant audiences of the review (e.g., scholars, policymakers, and/or practitioners).

Adapted from Hallinger (2014)

Methodology

Sources and Search Procedures

PRISMA (i.e., Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Moher et al., 2010) were used in the execution of the search process. As declared on the PRISMA Statement (Page et al., 2021), PRISMA provides detailed steps via a checklist (PRISMA, 2020) for the purpose of helping reviewers provide a standardized reporting of systematic reviews and meta-analyses. Scopus (circa 2004) was used for the meta-systematic review because it is deemed to provide an adequate balance between coverage and trustworthiness of the sources identified (Hallinger (2019b).

PRISMA was used to identify the full set of systematic reviews that were included in Scopus between 1960 and 2022, which represented every full year in the Scopus database at the time the study was conducted. Specifically, an initial keyword search was conducted using the following string: (TITLE-ABSKEY ("Educational Leadership") OR TITLE-ABS-KEY ("educational management") OR TITLE-ABS-KEY ("school leadership") OR TITLE-ABS-KEY ("school administration") AND TITLE-ABS-KEY ("systematic review"). The keywords "leadership," "management," and "administration" were used to reflect the fact that the field of leadership and management is pluralist (Bush, 2007). The goal in using the keywords (systematic review" was to identify works wherein the author(s) declared their reviews as representing some form of systematic review. The guiding (i.e., inclusion) criteria was that the systematic reviewer(s) had to make it explicit that African EDLM scholarship was included in the search.

A second round of searches was conducted by using the same keyword string, wherein "systematic review" was replaced with "science mapping." A third round of searches was conducted by replacing keyword string "systematic review" with "bibliometric." A fourth round of searches was conducted on the names that appeared in the documents selected in the previous three rounds to determine whether these authors had published other systematic reviews on the EDLM literature that had not yet been identified. Each of these four rounds was repeated 54 times by replacing "Africa", in turn, with each of the 54 independent African countries identified by the United Nations.

Data Extraction

Once the systematic reviews had been identified via Scopus, the pdf files of the full work of each source was obtained via the library interface at the university of the researchers. Each pdf file was downloaded and stored on a computer subdirectory (i.e., subfolder).

Data Analysis

Each article (i.e., systematic review) was coded using the eight criteria contained in Hallinger's (2014) analytical rubric, as described earlier. This coding yielded a *r* (number of extracted systematic reviews) x 8 (number of criteria) matrix—hereafter referred to as the Systematic Review Quality Matrix. Prior to coding the extracted corpus of systematic reviews, to ensure high inter-rater reliability, the present authors verified the coding undertaken by Hallinger (2014), using his three-level analytical rubric, of each of the 38 "reviews of research" (p. 539) (i.e., systematic reviews). Hallinger (2014) presented the results from his coding (i.e., quality matrix) in Table 2 of his article (pp. 556-558). The present authors were able to verify Hallinger's (2014) coding results in Table 2 with 100% accuracy. Once this 38-study quality matrix had been verified, the present authors were in a position to create their own quality matrix from the corpus of systematic reviews that they had extracted.

This quality matrix generated scores that had a Cronbach's alpha reliability coefficient of .82 (95% confidence interval [CI] = .73, .89). The score reliability coefficient exceeds Nunnally and Bernstein's (1994) cut-point of .70, thereby indicating that the quality matrix scores yielded excellent reliability.

From this quality matrix, descriptive statistics (i.e., frequency counts) were used to determine the conceptual, methodological, and reporting strengths and weaknesses of each individual systematic review, as well as the set of systematic reviews (i.e., via descriptive examination of patterns in the criterion scores on the rubric achieved across the corpus of systematic reviews). This analysis of the strengths and weaknesses enabled the researchers to identify trends in the three quality domains wherein the systematic reviews either fulfilled, partially fulfilled, or did not fulfill good practices of systematic reviews.

Results

The PRISMA chart (see Figure 2) revealed that the four rounds yielded 115 Scopus-indexed systematic review documents on African EDLM from 1960 to 2022. However, after reading their abstracts and the full text of these works, 81 documents were determined to be unsuitable and deleted from the list (representing a false positive rate of 70.43%), leaving 34 documents. A series of follow-up Scopus review searches replacing "Africa," in turn, with each of the 54 independent African countries provided an additional 13 documents. Of these 13 documents, 5 were unsuitable and, therefore, deleted from the list. Therefore, our meta-systematic review revealed that the total number of African EDLM-based systematic review documents for the 1960-2022 period was 42.

Table 1 presents the Systematic Review Quality Matrix pertaining to the 42 articles (i.e., systematic reviews). It can be seen from Table 1 that, with respect to the eight quality criteria, three criteria were particularly strongly met by the systematic reviewers: (a) Stating Implications (M = 1.93), Communicating Findings (M = 1.90), and Stating the Purpose (M = 1.88). As indicated by their consistently higher scores on the analytical rubric, these criteria were met with much greater frequency. These three criteria were followed, respectively, by Clarifies Method of Data Extraction (M = 1.71) and Clarifies Method of Data Analysis (M = 1.64). Both of these criteria also were met relatively strongly. The systematic reviewers met the desired standard at a moderate level with respect to the following three criteria, respectively: Justifies Search Procedures and Sources (M = 1.48), Conceptual Framework (M = 1.33), and States Limitations of the Review (M = 1.24). For these three criteria, on average, their scores were closer to 1 (i.e., criterion partially met) than to 2 (i.e., criterion met). Interestingly, States Limitations of the Review (SD = 0.93) and Conceptual Framework (SD = 0.93), by far, were most variable with respect to quality, with Stating Implications (SD = 0.26) being the least variable.

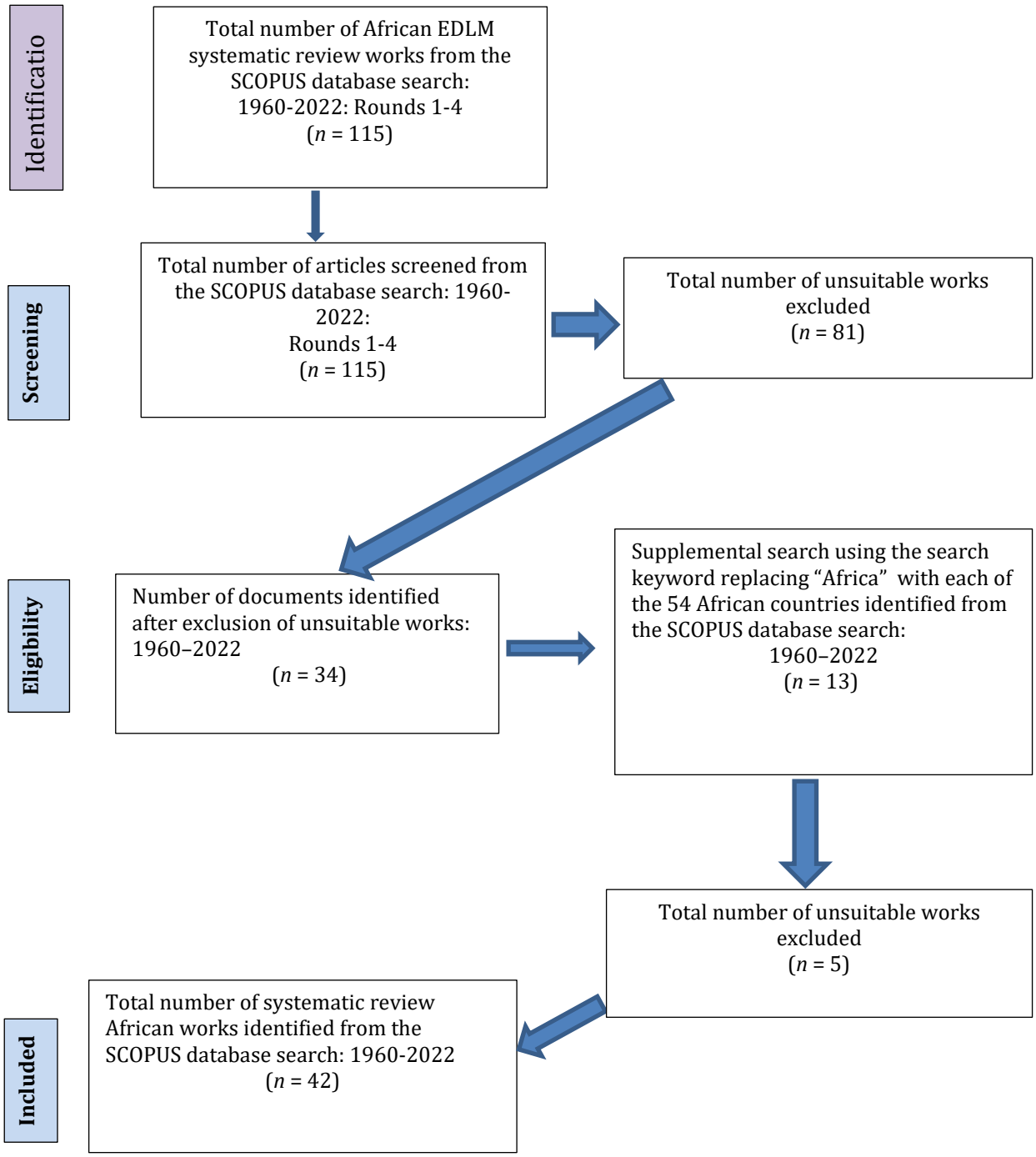

Figure 2: PRISMA Flow Chart for Meta-Systematic Review Detailing Steps in the Identification and Screening of Scopus-indexed, Africa-Included EDLM Systematic Review Works

Table 1: Examination of Systematic Reviews of African EDLM Scholarship Via 8-Item Criteria

No	Author	Year	Purpose	Conceptual Framework	Search and Sources	Data Extraction	Data Analysis	Communicates Findings	States Limitations	States Implications	Total Rubric Score
1	Richardson and McLeod	2009	2	0	1	2	2	2	0	2	11
2	Anohah and Suhonen	2015	2	0	0	0	0	1	1	2	6
3	McEwan	2015	0	0	1	2	2	2	0	2	9
4	Bush and Glover	2016a	2	2	1	0	0	1	0	1	7
5	Bush and Glover	2016b	0	0	0	0	0	1	0	1	2
6	Wang and Bowers	2016	2	2	1	2	2	2	2	2	15
7	Hallinger and Bridges	2017	2	2	1	2	2	2	2	2	15
8	Mertkan, Arsan, Inal Cavlan, and Onurkan Aliusta	2017	2	1	1	2	2	2	0	2	12
9	González-Campos, Aspeé-Chacón, Sessarego-Espeleta, González-Suárez, and Gómez-Osorio	2018	2	2	2	2	2	2	1	2	15
10	Hallinger	2018	2	2	2	2	2	2	2	2	16
11	Belfiore, Iovino, and Tafuri	2019	2	2	1	2	2	2	0	2	13
12	Hallinger	2019a	2	2	2	2	2	2	2	2	16
13	Hallinger	2019b	2	2	**2**	2	2	2	2	2	16
14	Hallinger and Kovačević	2019	2	2	**2**	2	2	2	2	2	16
15	Hallinger and Kovačević	2021	2	2	**2**	2	2	2	2	2	16
16	Nguyen, Harris, and Ng	2020	2	2	1	2	2	2	2	2	15
17	Pažur and Kovač	2019	2	0	0	0	0	1	0	2	5
18	Tian and Huber	2021	2	2	2	2	2	2	2	2	16
19	Tian and Huber	2020	2	2	2	2	2	2	2	2	16
20	Cruz-González, Rodríguez, and Segovia	2021	2	2	**2**	2	2	2	0	2	14
21	Hallinger	2020	2	2	**2**	2	2	2	2	2	16
22	Hallinger, Gümüş, and Bellibaş	2020	2	2	**2**	2	2	2	2	2	16
23	Jambo and Hongde	2020	2	0	2	2	1	2	0	2	11
24	Gümüş, Arar, and Oplatka	2021	2	1	2	2	2	2	2	2	15
25	Lipscombe, K., Tindall-Ford, S., and Lamanna	2023	2	2	**2**	2	2	2	0	2	14
26	Molina-Astoraymeand Cabanillas-Carbonell	2020	2	0	**2**	2	2	2	0	2	12
27	Webster, Glascoe, Moore, DauenhaueEgan, Russ, Orendorff, and Buschmeie	2020	2	2	**2**	2	2	2	1	2	15

Table 1. Continued

No	Author	Year	Purpose	Conceptual Framework	Search and Sources	Data Extraction	Data Analysis	Communicates Findings	States Limitations	States Implications	Total Rubric Score
28	Moyo, Perumal, and Hallinger	2020	2	0	1	1	2	2	2	2	12
29	Poekert, Swaffield, Demir, and Wright	2020	2	1	**2**	2	2	2	2	2	15
30	Cuéllar, Queupil, Cuenca, and Ravest	2020	2	2	**2**	2	2	2	0	2	14
31	Eckert, Sopory, Day, Wilkins, Padgett, Novak, Noyes, Allen, Alexander, Vanderford, and Gamhewage	2018	2	2	1	2	2	2	2	2	15
32	Hammad and Hallinger	2017	2	2	1	2	2	2	2	2	15
33	Kuzhabekova	2021	2	2	1	2	2	2	2	2	15
34	Hallinger and Kovacevic	2022	2	2	**2**	2	2	2	2	2	16
35	Sudirman and Gemilang	2020	2	1	1	2	2	2	0	2	12
36	Bush, Kirezi, Ashford, and Glover	2022	2	0	0	0	0	2	0	1	5
37	Nwosu, Matashu, and Buabeng	2022	2	2	**2**	2	0	2	2	2	14
38	Limone and Toto	2022	2	0	**2**	2	1	2	2	2	13
39	Gómez-Leal, Holzer, Bradley, Fernández-Berrocal, and Patti	2022	2	0	**2**	2	2	2	2	2	14
40	Costan, Gonzales, Gonzales, Enriquez, Costan, Suladay, Atibing, Aro, Evangelista, Maturan, Selerio, and Ocampo	2021	1	2	**2**	2	1	2	1	2	13
41	Tamadoni, Hosseingholizadeh, and Bellibaş	2024	2	2	**2**	2	2	2	2	2	16
42	Sun and Leithwood	2015	2	0	1	1	2	2	2	2	12
	Total		79	56	62	72	69	80	52	81	
	Mean		1.88	1.33	1.48	1.71	1.64	1.9	1.24	1.93	
	Standard Deviation		0.45	0.9	0.67	0.67	0.73	0.3	0.93	0.26	

.

Figure 3 displays a correspondence analysis plot, wherein the 42 systematic review studies are mapped onto the space that displays the seven quality criteria. It can be seen from this figure that the vast majority of the systematic reviews are located around the following five criteria: Clarifies Method of Data Extraction, States Limitations of the Review, Clarifies Method of Data Analysis, Communicating Findings, and Justifies Search Procedures and Sources.

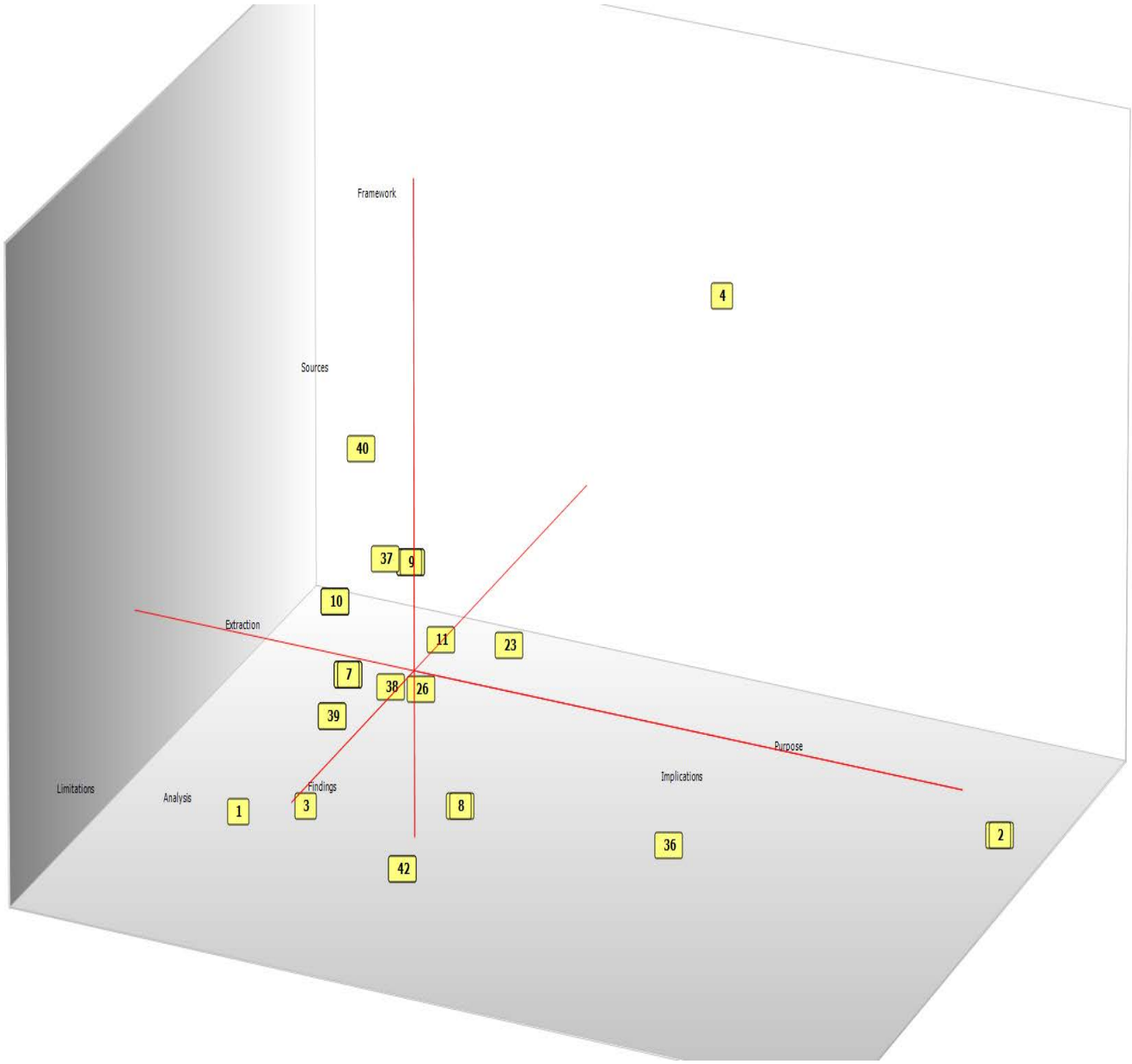

Figure 3: Correspondence Plot of the Eight Quality Criteria

On a positive note, none of the criteria yielded a mean score below 1. Even more positively, the scores generated by these systematic reviewers were statistically significantly higher than that for the systematic reviewers in Hallinger's (2014) study with respect to the following five criteria: Stating the Purpose (mean = 1.88 vs. 1.66; t = 3.29, p = .002; Cohen's [1988] d = 0.74), Justifies Search Procedures and Sources (M = 1.48 vs. 0.97; t = 2.71, p = .008; d = 0.61), Clarifies Method of Data Extraction (M = 1.71 vs. 0.74; t = 5.35, p < .001; d = 1.20), Clarifies Method of Data Analysis (M = 1.64 vs. 0.68; t = 5.36, p < .001; d = 1.20), and Communicating Implications (M = 1.93 vs. 1.47; t = 2.56, p = .012; d = 0.57). The effect sizes associated with these five statistically significant differences were moderate (i.e., Communicating Implications) to very large (i.e., Clarifies Method of Data Extraction, Clarifies Method of Data Analysis). The other three differences—that is, Conceptual Framework (M = 1.33 vs. 1.13; t = 0.98, p = .33; d = 0.22), Communicating Findings (M = 1.90 vs. 1.87; t = 0.51, p = .61; d = 0.11), and States Limitations of the Review (M = 1.24 vs. 0.92; t = 1.54, p = .13; d = 0.34)—were not statistically significant, although, in each case, the mean scores in the present study were larger. Indeed, using Onwuegbuzie and Levin's (2005) Binomial Test of Trend, the probability that 8 out of the 8 criteria yielding higher scores for the systematic reviewers in the present study was .004, indicating that the trend was statistically significant, with a maximum effect size of 1.00 (i.e., 8/8). Therefore, the eight sets of quality criterion scores were consistently higher for the present set of systematic reviewers.

Separate mean scores were computed for each of the three individual dimensions of quality, namely, Conceptual (i.e., Statement of Purpose; Conceptual Framework), Methodological (i.e., Sources and Search Procedures; Data Extraction; Data Analysis), and Reporting (i.e., Presentation of Findings; Limitations of the Review; Implications of the Review). These means revealed that that all three dimensions were very similar (see Table 2), with the Reporting dimension being slightly higher. However, they are not statistically significantly different from each other.

Table 2. Means and Standard Deviations Pertaining to the Three Quality Dimensions of quality, Conceptual, Methodological, and Reporting (n = 42)

Quality Dimension	*M*	*SD*
Conceptual	1.61	0.56
Methodological	1.61	0.61
Reporting	1.69	0.41

A series of Spearman rho correlations revealed that the number of authors did not predict any of the quality criterion scores. Contrastingly, year of publication of the systematic reviews was positively and statistically significantly related to both the Sources and Search Procedures (r_s = .52, p < .001) and Presentation of Findings (r_s = .36, p < .02) criteria, as well as the Methodological dimension (r_s = .31, p < .05). This finding suggests that, with respect to these three indices, in general, the most recent systematic reviews were more likely to meet these criteria.

Of the 42 systematic reviews, 11 (26.2%) obtained the maximum score of 16. In other words, they met all eight criteria for a systematic review. Hallinger (2014) labelled such studies as representing "exemplary reviews" (p. 566). In Hallinger's (2014) study, 8 of the 38 (21.1%) systematic reviews were categorized as being exemplary. Although approximately 5% higher, the proportion of exemplary reviews was not statistically significantly higher (Fisher's Exact Test p = .61). Encouragingly, a further 10 systematic reviews met seven of the eight criteria, indicating that exactly one half (n = 21) of the reviews met at least seven of the criteria.

Finally, a concerning finding is that, although all but one of the systematic reviews in the present investigation was published at least 5 years after the PRISMA guidelines were published, only 17 of these 41 reviews (41.5%) included PRISMA displays of some form. Of these reviews, 7 of them represented exemplary studies. Interestingly, Zorzela et al. (2014) reported that "the overall, unweighted, proportion of reviews with good reporting" was 56% (95% confidence interval [CI] = 55% to 57%) (p. 1), which is higher than that observed in the present study.

Conclusion

Like all systematic reviews, the present meta-systematic review has several limitations. First and foremost, the corpus of African EDLM literature was extracted for this study via the Scopus database. Unfortunately, although Scopus is the largest abstract and citation database of peer-reviewed literature, in particular, it does not index all the journals (Meho & Yang, 2007); also, it has less depth of coverage, only covering works from 1960, as opposed to Web of Science, which has indexing coverage from the year 1900 to the present (Meho & Yang, 2007). As a result, Scopus inevitably yields false negatives. That is, it is likely that using the Scopus database led to the omission of key African EDLM works.

Another limitation stems from the maximum word count allowed for this manuscript by the journal editor. The 10,000-word maximum (including references, appendices, tables, and figures) made it extremely challenging to produce a meta-systematic review that met all eight quality criteria, especially bearing in mind that merely presenting the references of the 42 systematic reviews identified via the meta-systematic review took up 1,274 words. This word limit also prevented any discussion of findings from the 42 systematic reviews. Consequently, this discussion will be provided in a follow-up work.

Comparing the quality matrix in the present study to Hallinger's (2014) quality matrix revealed that the present corpus of systematic reviews consistently demonstrated more of the characteristics associated with good practices of systematic reviews. That is, in general, the eight sets of quality criterion scores are consistently higher for the present set of systematic reviewers. More specifically, computing the mean score across all eight criteria representing all systematic reviews—yielding what we refer to as an *overall quality index*—revealed that the overall quality index for the present inquiry (1.64) was statistically significantly (t = 3.90, p < .001) higher than was the overall quality index for Hallinger's (2014) systematic review studies (1.18). Hallinger's (2014) studies ranged from 1961 to 2012. In the present inquiry, there was one systematic review study published in 2009, with the remaining studies published between 2015 and 2022, inclusive. Therefore, it is clear that the quality of systematic reviews in the area of EDLM in the last 10 years has increased from previous years, at least in terms of the African EDLM literature.

This positive conclusion regarding the improvement in quality of systematic reviews during the last decade likely reflects the influence of the Campbell Collaboration. This organization emerged in 2000 as the standard for systematic reviews in the social sciences as a means of informing decision-making and contributing to evidence-based practices (Petrosino

et al., 2001). However, there still remains considerable room for improvement in all three quality dimensions (i.e., conceptual, methodological, and reporting) pertaining to systematic reviews on African EDLM knowledge production.

Addressing Contextual Challenges in Conducting High-Quality Systematic Reviews in African EDLM Research

African researchers conducting systematic reviews of EDLM face numerous systemic challenges that hinder their ability to meet global standards for methodological rigor. One of the most pressing issues is the limited access to financial, technological, and infrastructural resources. Many African universities and research institutions lack subscriptions to comprehensive databases such as Scopus or Web of Science, which are critical for conducting exhaustive literature searches. This often forces researchers to rely on open-access resources, leading to incomplete or biased datasets (Bush & Glover, 2016b). Additionally, limited funding for academic research constrains the ability of scholars to access advanced training in systematic review methodologies or to hire multidisciplinary teams, which are often essential for conducting rigorous reviews (Bush & Glover, 2016b; Hallinger, 2014).

A recurring challenge is the lack of capacity-building opportunities for African EDLM researchers. Training programs in systematic review methodologies, particularly those emphasizing advanced techniques for data synthesis and bias reduction, remain scarce in the region (Oronje et al., 2022). Although global initiatives such as the Campbell Collaboration provide useful frameworks, localised application of these methods often requires contextual adaptation, which many researchers are not equipped to undertake (Hallinger, 2014; Oronje et al., 2022). Capacity-building efforts further are hindered by high teaching loads and administrative responsibilities that limit the time that researchers can dedicate to methodological training or conducting systematic reviews. As a result, many systematic reviews conducted in the region struggle to meet global standards of rigor and transparency (Hallinger, 2014).

Another challenge stems from the difficulty of adapting global systematic review tools to the African context. Many frameworks, such as PRISMA and Hallinger's (2014) analytical rubric, are developed with higher-income countries in mind and might not account for the socio-cultural, economic, and political complexities of African educational systems. This creates a gap between global best practices and the contextual realities faced by African researchers. Without tools and guidelines specifically tailored to the African context, researchers often face difficulties in producing systematic reviews that are both methodologically sound and locally relevant (Hallinger, 2014; Moyo et al., 2020).

African researchers also face barriers due to fragmented research networks, which limit opportunities for collaborative and interdisciplinary research. Systematic reviews often require collaboration across fields to ensure comprehensive analyses; yet, the absence of robust research networks constrains the sharing of expertise, resources, and data (Bush & Glover, 2016b). This isolation not only reduces the methodological quality of systematic reviews, but also hampers their ability to address complex, cross-cutting issues in education (Moyo et al., 2020). Building stronger networks among African EDLM scholars significantly could enhance the quality and impact of systematic reviews by fostering collaboration, mentorship, and the pooling of resources (Moyo et al., 2020; Oronje et al., 2022).

Finally, linguistic diversity and the dominance of English in academic publishing present significant obstacles. Systematic reviews frequently exclude non-English publications, which sidelines research conducted in African languages or widely spoken languages such as French, Portuguese, and Arabic (Bush & Glover, 2016b). This exclusion limits the evidence base and diminishes the inclusivity and representativeness of findings (Moyo et al., 2020). Addressing language and publication barriers is essential to ensuring that systematic reviews fully capture the breadth of African EDLM scholarship (Bush & Glover, 2016b; Moyo et al., 2020).

In light of these challenges, several recommendations are proposed to improve the quality of systematic reviews conducted by African EDLM researchers. Investment in capacity-building initiatives, such as regionally focused training programs and workshops tailored to the specific needs of African scholars, is critical (Oronje et al., 2022). Increased funding for academic research, access to global databases, and the development of localized tools for conducting systematic reviews are necessary steps toward overcoming these barriers (Hallinger, 2014). Furthermore, fostering regional and international collaborations can bridge resource gaps and promote methodological excellence. By addressing these contextual challenges, African researchers can produce high-quality systematic reviews that not only advance academic knowledge, but also inform evidence-based educational policies and practices across the continent (Hallinger, 2014).

Practical Strategies for Enhancing the Quality of Systematic Reviews in African EDLM

In order to enhance the quality of systematic reviews conducted by African EDLM researchers, we propose the following actionable and context-specific implementation strategies aligned with our recommendations:

1. Invest in Capacity-Building Initiatives. To build capacity among African researchers, a multi-tiered approach is recommended, as follows:

Step 1: Identify key skill gaps in systematic review methodologies by conducting needs assessments in educational institutions. These assessments should involve surveys and focus groups with researchers, administrators, and policymakers.

Step 2: Develop targeted training programs tailored to African contexts, such as that conducted by Oronje et al. (2022). These programs should include workshops on advanced techniques such as data synthesis, meta-analysis, and bias assessment, delivered by experienced trainers familiar with local research challenges.

Step 3: Partner with international and regional organizations, such as the Campbell Collaboration or African research consortia—such as Wellcome Trust's African Institutions Initiative, Consortium for Advanced Research Training in Africa (CARTA), Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), and African Population and Health Research Center (APHRC)—to provide mentorship opportunities and access to global resources. These partnerships should prioritize sustainable, long-term skill development rather than one-time interventions.

Step 4: Establish training hubs within leading universities to serve as regional centers of excellence, facilitating continuous learning and knowledge-sharing among researchers.

2. Increase Funding and Access to Resources. Addressing resource limitations requires a combination of institutional, national, and international efforts:

Step 1: Advocate for increased national funding for research by engaging with policymakers and showcasing the value of high-quality systematic reviews in evidence-based policy formulation.

Step 2: Negotiate discounted or subsidized access to global databases such as Scopus and Web of Science through collective bargaining agreements involving universities and governments.

Step 3: Develop open-access repositories and platforms that aggregate local and regional research outputs, ensuring broader access to African scholarship for systematic reviews.

Step 4: Introduce competitive research grants specifically for systematic review projects, prioritizing multidisciplinary teams that address key regional educational challenges.

3. Foster Regional and International Collaborations. Collaboration across disciplines and regions significantly can enhance the quality and relevance of systematic reviews:

Step 1: Establish interdisciplinary research networks by hosting regional conferences and online platforms for African EDLM researchers to connect and to share expertise.

Step 2: Encourage partnerships among African institutions and established research organizations globally to facilitate knowledge exchange, resource sharing, and co-authorship opportunities.

Step 3: Create collaborative funding schemes that incentivize joint research projects involving multiple African universities and international partners.

4. Localize Systematic Review Frameworks. Adapting global tools to the African context requires customization and the involvement of partners:

Step 1: Form working groups of African EDLM researchers to evaluate existing frameworks such as PRISMA and Hallinger's (2014) rubric for their applicability in the African context.

Step 2: Modify these frameworks to include culturally relevant parameters and contextual factors such as linguistic diversity, resource constraints, and localized research priorities.

Step 3: Pilot the adapted frameworks through systematic reviews in diverse African contexts, gathering feedback to refine the tools further.

Step 4: Publish and disseminate these adapted frameworks widely, accompanied by detailed user guides for African researchers.

5. Address Linguistic and Publication Barriers. Inclusive systematic reviews must account for the linguistic diversity and publication realities of African research:

Step 1: Promote the translation of key research outputs into English and other widely spoken languages, supported by translation grants and institutional partnerships.

Step 2: Encourage journal editors to accept and to publish systematic reviews in multiple languages, enhancing accessibility and inclusivity.

Step 3: Create repositories of non-English research outputs, enabling researchers to include these studies in their systematic reviews.

Step 4: Provide training for researchers on how to identify and to incorporate non-English studies into systematic reviews effectively.

By implementing these strategies systematically, educational institutions and researchers can overcome key challenges and significantly improve the quality of systematic reviews in the African EDLM context. These step-by-step guidelines

not only provide a clear roadmap, but also ensure that the proposed recommendations are practical, sustainable, and impactful. By addressing critical gaps such as capacity-building, resource accessibility, and methodological rigor, these strategies pave the way for more robust and reliable evidence to inform educational policies and practices.

Final Thoughts

The conclusion that there is a need for continued improvement has important implications for college-/university-level faculty, authors, and journal editors. At the college/university level, we recommend that college-/university-level faculty develop courses, or at least workshops/seminars, on systematic reviews for graduate students wherein the importance of conceptual, methodological, and reporting quality of systematic reviews is emphasized. As many students as possible then could be encouraged to enroll in these courses/workshops/seminars. Also, it is important that mentors and thesis/dissertation advisors/supervisors stress continually to their students the importance of conducting quality systematic reviews.

In terms of authors, systematic reviewers should be encouraged to read as many exemplary systematic reviews as possible, such as the 11 exemplary reviews identified in Table 1. Such exemplary reviews could help reviewers to develop strategies for producing quality reviews. Additionally, when conducting systematic reviews might consider using tools such as Hallinger's (2014) analytical rubric (see Table 1). Finally, with regard to the journal review process, editors might consider using several strategies at their disposal that might help improve the quality of systematic reviews that they publish in their journals. In particular, authors could be given instruments like the analytical rubric and then asked to declare (e.g., by checking an option that declares) that they have attempted to address in their manuscripts the three dimensions of quality to the greatest extent possible. Further, journal editors might consider asking reviewers/editorial board members to use the analytical rubric when reviewing systematic review manuscripts. Also, journal editors should not include the references in systematic reviews as part of the maximum word count because this means that authors will be unduly penalized for the good and transparent practice of presenting the references of the extracted works, with this penalty increasing as a function of the number of works extracted via the search.

We have attempted to provide some recommendations for the various members of the professional community who play a role at various stages in the quality of manuscripts that authors write. However, our list of suggestions is by no means exhaustive. Indeed, we encourage members of these different sectors to develop other strategies for improving the quality of systematic reviews. As the number of quality systematic reviews in the field of African EDLM scholarship and beyond increases, it is hoped that a culture of rigorous systematic reviews will be enhanced that better informs decision-making and contributes to evidence-based practices.

Funding

Funding for this project was provided by MasterCard Foundation under their Leaders in Teaching Programme (Rwanda).

References

(* denotes a work identified from the meta-systematic review)

*Anohah, E., & Suhonen, J. (2015). Modelling mobile learning policy in computing education for developing countries in Africa: A backward mapping approach. *International Journal of Mobile Learning and Organisation, 9*(3), 201-217. https://doi.org/10.1504/IJMLO.2015.074203

Beauregard, N., Marchand, A., & Blanc, M.-E. (2011). What do we know about the non-work determinants of workers' mental health? A systematic review of longitudinal studies. *BMC Public Health, 11*, Article 439. https://doi.org/10.1186/1471-2458-11-439

*Belfiore, P., Iovino, S., & Tafuri, D. (2019). Sport management and educational management: A bibliometric analysis. *Sport Science, 12*(1), 61-64. https://ricerca.uniparthenope.it/handle/11367/75871

Bush, T. (2007). Educational leadership and management: Theory, policy, and practice. *South African Journal of Education, 27*(3), 391-406. https://hdl.handle.net/10520/EJC32136

*Bush, T., & Glover, D. (2016a). School leadership and management in South Africa: Findings from a systematic literature review. *International Journal of Educational Management, 30*, 211-231. https://doi.org/10.1108/IJEM-07-2014-0101

*Bush, T., & Glover, D. (2016b). School leadership in West Africa: Findings from a systematic literature review. *Africa Education Review, 13*(3-4), 80-103. https://doi.org/10.1080/18146627.2016.1229572

*Bush, T., Kirezi, J., Ashford, R., & Glover, D. (2022). School leadership and gender in Africa: A systematic overview. *Research in Educational Administration and Leadership, 7*(4), 680-712. https://doi.org/10.30828/real.1159040

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

*Costan, E., Gonzales, G., Gonzales, R., Enriquez, L., Costan, F., Suladay, D., Atibing, N. M., Aro, J. L., Evangelista, S. S., Maturan, F., Selerio, E., Jr., & Ocampo, L. (2021). Education 4.0 in developing economies: A systematic literature review of implementation barriers and future research agenda. *Sustainability, 13*(22), Article 12763. https://doi.org/10.3390/su132212763

*Cruz-González, C., Rodríguez, C. L., & Segovia, J. D. (2021). A systematic review of principals' leadership identity from 1993 to 2019. *Educational Management Administration and Leadership, 49*(1), 31-53. https://doi.org/10.1177/1741143219896053

*Cuéllar, C., Queupil, J. P., Cuenca, C., & Ravest, J. (2020). A systematic review on multiculturalism and educational leadership: Similarities and contrasts in knowledge production across societies. *Multicultural Education Review, 12*(4), 235-249. https://doi.org/10.1080/2005615X.2020.1842655

*Eckert, S., Sopory, P. Day, A., Wilkins, L., Padgett, D., Novak, J., Noyes, J., Allen, T., Alexander, N., Vanderford, M., & Gamhewage, G. (2018). Health-related disaster communication and social media: Mixed-method systematic review. *Health Communication, 33*(12), 1389-1400. https://doi.org/10.1080/10410236.2017.1351278

Ehsan, A., Klaas, H. S., Bastianen, A., & Spini, D. (2019). Social capital and health: A systematic review of systematic reviews. *SSM - Population Health, 8*, Article 100425. https://doi.org/10.1016/j.ssmph.2019.100425

EQUATOR Network. (2019). *Enhancing the Quality and Transparency Of health Research.* https://www.equator-network.org/

*Gómez-Leal, R., Holzer, A. A., Bradley, C., Fernández-Berrocal, P., & Patti, J. (2022). The relationship between emotional intelligence and leadership in school leaders: A systematic review. *Cambridge Journal of Education, 52*(1), 1-21. https://doi.org/10.1080/0305764X.2021.1927987

*González-Campos, J., Aspeé-Chacón, J., Sessarego-Espeleta, I., González-Suárez, H., & Gómez-Osorio, R. (2018). Análisis de la producción científica en política y gestión educativa publicada en SCIELO 2012-2015 [Analysis of the scientific production in politics and educational management published in SCIELO 2012-2015]. *Revista Electronica Educare, 22*(3). https://bit.ly/4hfQAvR

*Gümüş, S., Arar, K., & Oplatka, I. (2021). Review of international research on school leadership for social justice, equity and diversity. *Journal of Educational Administration and History*, *53*(1), 81-99. https://doi.org/10.1080/00220620.2020.1862767

Hallinger, P. (2014). Reviewing reviews of research in educational leadership: An empirical assessment. *Educational Administration Quarterly, 50*(4), 539-576. https://doi.org/10.1177/0013161X13506594

*Hallinger, P. (2018). Surfacing a hidden literature: A systematic review of research on educational leadership and management in Africa. *Educational Management Administration and Leadership, 46*(3), 362-384. https://doi.org/10.1177/1741143217694895

*Hallinger, P. (2019a). A systematic review of research on educational leadership and management in South Africa: Mapping knowledge production in a developing society. *International Journal of Leadership in Education, 22*(3), 316-334. https://doi.org/10.1080/13603124.2018.1463460

*Hallinger, P. (2019b). Science mapping the knowledge base on educational leadership and management in Africa, 1960–2018. *School Leadership and Management, 39*(5), 537-560. https://doi.org/10.1080/13632434.2018.1545117

*Hallinger, P. (2020). Science mapping the knowledge base on educational leadership and management from the emerging regions of Asia, Africa and Latin America, 1965–2018. *Educational Management Administration and Leadership, 48*(2), 209-230. https://doi.org/10.1177/1741143218822772

*Hallinger, P., & Bridges, E. M. (2017). A systematic review of research on the use of problem-based learning in the preparation and development of school leaders. *Educational Administration Quarterly, 53*(2), 255-288. https://doi.org/10.1177/0013161X16659347

*Hallinger, P., Gümüş, S., & Bellibaş, M. Ş. (2020). 'Are principals instructional leaders yet?' A science map of the knowledge base on instructional leadership, 1940–2018. *Scientometrics, 122,* 1629-1650. https://doi.org/10.1007/s11192-020-03360-5

*Hallinger, P., & Kovačević, J. (2019). A bibliometric review of research on educational administration: Science mapping the literature, 1960 to 2018. *Review of Educational Research, 89*(3), 335-369. https://doi.org/10.3102/0034654319830380

*Hallinger, P., & Kovačević, J. (2021). Science mapping the knowledge base in educational leadership and management: A longitudinal bibliometric analysis, 1960 to 2018. *Educational Management Administration and* Leadership, *49*(1), 5-30. https://doi.org/10.1177/1741143219859002

*Hallinger, P., & Kovačević, J. (2022). Mapping the intellectual lineage of educational management, administration and leadership, 1972–2020. *Educational Management Administration and Leadership*, *50*(2) 192-216. https://doi.org/10.1177/17411432211006093

*Hammad, W., & Hallinger, P. (2017). A systematic review of conceptual models and methods used in research on educational leadership and management in Arab societies. *School Leadership and Management, 37*(5), 434-456. https://doi.org/10.1080/13632434.2017.1366441

Hong, Q. N., & Pluye, P. (2019). A conceptual framework for critical appraisal in systematic mixed studies reviews. *Journal of Mixed Methods Research, 13*(4), 446-460. https://doi.org/10.1177/1558689818770058

*Jambo, D., & Hongde, L. (2020). The effect of principal's distributed leadership practice on students' academic achievement: A systematic review of the literature. *International Journal of Higher Education, 9*(1), 189-198. https://doi.org/10.5430/ijhe.v9n1p189

*Kuzhabekova, A. (2021). A postcolonial feminist exploration of the scholarship on women and educational leadership with a bibliometric approach. *Educational Management Administration and Leadership*, 1-21. https://doi.org/10.1177/17411432211003885

*Limone, P., & Toto, G. A. (2022). Psychological strategies and protocols for promoting school well-being: A systematic review. *Frontiers in Psychology, 13*, Article 914063. https://doi.org/10.3389/fpsyg.2022.914063

*Lipscombe, K., Tindall-Ford, S., & Lamanna, J. (2023). School middle leadership: A systematic review. *Educational Management Administration and Leadership, 51*(2), 270-288. https://doi.org/10.1177/1741143220983328

*McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research, 85*(3), 353-394. https://doi.org/10.3102/0034654314553127

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105-2125. https://doi.org/10.1002/asi.20677

*Mertkan, S., Arsan, N., Inal Cavlan, G., & Onurkan Aliusta, G. (2017). Diversity and equality in academic publishing: The case of educational leadership. *Compare, 47*(1), 46-61. https://doi.org/10.1080/03057925.2015.1136924

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *International Journal of Surgery, 8*(5), 336-341. https://doi.org/10.1016/j.ijsu.2010.02.007

*Molina-Astorayme, J., & Cabanillas-Carbonell, M. (2020, October 21). *Predicting academic performance using automatic learning techniques: A review of the scientific literature*. Proceedings of the 2020 IEEE Engineering International Research Conference. Article number 9254065. https://hdl.handle.net/11537/26929

*Moyo, Z., Perumal, J., & Hallinger, P. (2020). Struggling to make a difference against the odds: A synthesis of qualitative research on women leading schools in Zimbabwe. *International Journal of Educational Management, 34*(10), 1577-1594. https://doi.org/10.1108/IJEM-01-2020-0015

*Nguyen, D., Harris, A., & Ng, D. (2020). A review of the empirical research on teacher leadership (2003–2017): Evidence, patterns and implications. *Journal of Educational Administration, 58*(1), 60-80. https://doi.org/10.1108/JEA-02-2018-0023

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

*Nwosu, L. I., Matashu, M., & Buabeng, A. T. (2022). A call to strengthen instructional leadership to support learner achievement during and post COVID-19: A systematic literature review approach. *International Journal of Learning, Teaching and Educational Research, 21*(7), 219-240. https://doi.org/10.26803/ijlter.21.7.12

Onwuegbuzie, A. J., & Frels, R. (2016). *Seven steps to a comprehensive literature review: A multimodal and cultural approach*. Sage.

Onwuegbuzie, A. J., & Levin, J. R. (2005). Strategies for aggregating the statistical nonsignificant outcomes of a single study. *Research in the Schools, 12*(1), 10-19. https://bit.ly/40Xb0nm

Oronje, R. N., Mukiira, C., Kahurani, E., & Murunga, V. (2022). Training and mentorship as a tool for building African researchers' capacity in knowledge translation. *PLOS ONE, 17*(3), Article e0266106. https://doi.org/10.1371/journal.pone.0266106

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., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

*Pažur, M., & Kovač, V. (2019). Democratic school leadership: Analysis of current researches and open issues. [Demokratsko Školsko voĂenje: Analiza dosadaŠnjih istraŽivanja i otvorena pitanja] *Metodicki Ogledi, 26*(1), 33-60. https://doi.org/10.21464/mo.26.1.5

Petrosino, A., Boruch, R. F., Soydan, H., Duggan, L., & Sanchez-MecaFirst, J. (2001). Meeting the challenges of evidence-based policy: The Campbell Collaboration. *The Annals of the American Academy of Political and Social Science, 578*(1), 14-34. https://doi.org/10.1177/000271620157800102

*Poekert, P. E., Swaffield, S., Demir, E. K., & Wright, S. A. (2020). Leadership for professional learning towards educational equity: A systematic literature review. *Professional Development in Education, 46*(4), 541-562. https://doi.org/10.1080/19415257.2020.1787209

Pollock, M., Fernandes, R. M., Becker, L. A., Pieper, D., & Hartling, L. (2020). Overviews of reviews. In J. Higgins, J. Thomas, J. Chandler, M. S. Cumpston, T. Li, M. J. Page, & W. Welch, *Cochrane handbook for systematic reviews of interventions* (Version 6.0). Cochrane Training. https://bit.ly/40EMHcM

PRISMA. (2020). *PRISMA 2020 checklist*. https://bit.ly/4juIAst

*Richardson, J. W., & McLeod, S. (2009). Where should educational leadership authors publish to get noticed by the top journals in the discipline? *Educational Administration Quarterly, 45*(4), 631-639. https://doi.org/10.1177/0013161X09331770

Rouleau, G., Hong, Q. N., Kaur, N., Gagnon, M.-P., Côté, J., Bouix-Picasso, J., & Pluye, P. (2023). Systematic reviews of systematic quantitative, qualitative, and mixed studies reviews in healthcare research: How to assess the methodological quality of included reviews? *Journal of Mixed Methods Research, 17*(1), 51-69. https://doi.org/10.1177/15586898211054243

Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K. F., & Altman, D. G. (2010). Transparent and accurate reporting increases reliability, utility, and impact of your research: Reporting guidelines and the EQUATOR Network. *BMC Medicine, 8*, Article 24. https://doi.org/10.1186/1741-7015-8-24

*Sudirman, A., & Gemilang, A. V. (2020). Promoting work-based learning as a praxis of educational leadership in higher education. *International Journal of Learning, Teaching and Educational Research, 19*(3), 149-173. https://doi.org/10.26803/ijlter.19.3.9

*Sun, J., & Leithwood, K. (2015). Leadership effects on student learning mediated by teacher emotions. *Societies, 5*(3), 566-582. https://doi.org/10.3390/soc5030566

*Tamadoni, A., Hosseingholizadeh, R., & Bellibaş, M. Ş. (2024). A systematic review of key contextual challenges facing school principals: Research-informed coping solutions. *Educational Management Administration and Leadership*, *52*(1), 116-150. https://doi.org/10.1177/17411432211061439

*Tian, M., & Huber, S. G. (2020). Mapping educational leadership, administration and management research 2007–2016: Thematic strands and the changing landscape. *Journal of Educational Administration, 58*(2), 129-150. https://doi.org/10.1108/JEA-12-2018-0234

*Tian, M., & Huber, S. G. (2021). Mapping the international knowledge base of educational leadership, administration and management: A topographical perspective. *Compare: A Journal of Comparative and International Education, 51*(1), 4-23. https://doi.org/10.1080/03057925.2019.1585757

Toye, F., Seers, K., Allcock, N., Briggs, M., Carr, E., Andrews, J., & Barker, K. (2013). "Trying to pin down jelly": Exploring intuitive processes in quality assessment for meta-ethnography. *BMC Medical Research Methodology, 13*, Article 46. https://doi.org/10.1186/1471-2288-13-46

*Wang, Y., & Bowers, A. J. (2016). Mapping the field of educational administration research: A journal citation network analysis. *Journal of Educational Administration, 54*(3), 242-269. https://doi.org/10.1108/JEA-02-2015-0013

*Webster, C. A., Glascoe, G., Moore, C., Dauenhauer, B., Egan, C. A., Russ, L. B., Orendorff, K., & Buschmeier, C. (2020). Recommendations for administrators' involvement in school-based health promotion: A scoping review. *International Journal of Environmental Research and Public Health, 17*(17), Article 6249. https://doi.org/10.3390/ijerph17176249

Whiting, P., Savovi´c, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., Churchill, R., & ROBIS Group. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology, 69*, 225-234. https://doi.org/10.1016/j.jclinepi.2015.06.005

Zorzela, L., Golder, S., Liu, Y., Pilkington, K., Hartling, L., Joffe, A., Loke, Y., & Vohra, S. (2014). Quality of reporting in systematic reviews of adverse events: Systematic review. *BMJ, 348*, Article f7668. https://doi.org/10.1136/bmj.f7668