Speaking of transparency: Are all Artificial Intelligence (AI) literature reviews in education transparent?

Ahmed Tlili\textsuperscript{A}  
Ronghui Huang\textsuperscript{B}  
Muhammad Yasar Mustafa\textsuperscript{C}  
Jialu Zhao\textsuperscript{D}  
Aras Bozkurt\textsuperscript{E}  
Lin Xu\textsuperscript{F}  
Huanhuan Wang\textsuperscript{G}  
Soheil Salha\textsuperscript{H}  
Fahriye Altinay\textsuperscript{I}  
Saida Affouneh\textsuperscript{J}  
Daniel Burgos\textsuperscript{K}

\textbf{Keywords}  
Artificial intelligence; education; literature review; transparency assessment.

\textbf{Abstract}  
Literature reviews are considered a core research approach in developing new theories and identifying trends and gaps in a given research topic. However, the transparency level of literature reviews might hinder the quality of the obtained findings, thus limiting their implications. As transparency is one of the core elements when implementing Artificial Intelligence (AI), this study assesses the transparency level of literature reviews on AI in education. Specifically, this study used a systematic review to collect and analyze information about reports of methodological decisions and research activities in 61 literature review papers. The obtained findings highlighted that 51.9\% of the conducted reviews on AI in education are descriptive. Additionally, the transparency level of the conducted literature reviews was low; 40\% of the reviews were in Q1 and 32\% in Q2. Particularly, the quality assessment step had the lowest transparency level. The findings of this research can advance the educational technology field by underscoring the methodological gaps when conducting a literature review on AI in education and hence enhance the transparency and trustworthiness of the obtained findings.

\textbf{Correspondence}  
ahmed.tlili23@yahoo.com \textsuperscript{A}

\textbf{Article Info}  
Received 17 July 2023  
Received in revised form 26 July 2023  
Accepted 26 July 2023  
Available online 31 July 2023  
\textbf{DOI}: https://doi.org/10.37074/jalt.2023.6.2.15
Introduction

Since 2000, the term ‘Technology-enhanced Learning’ (TEL) has appeared more frequently in the educational landscape (Al-Ataby, 2020). A complex and intertwined relationship arose between education and technology, and the use of technology in education evoked pedagogical, social, political, and economic effects (Guilherme, 2017). In other words, the technology-intensive 21st century carries various educational implications. Specifically, the use of technology in learning and teaching can effectively facilitate the accomplishment of teaching tasks, improve learning outcomes, and increase classroom interaction and communication. The last 30 years, the advent of Artificial Intelligence (AI) technology penetrated the educational domain as well (Tahiru, 2021), and Artificial Intelligence in Education (AIED) emerged as a developing research field. In AIED, a machine is designed to mimic a human system to support human learning and teaching (Tahiru, 2021; Conati et al., 2018). Therefore, AIED has cognitive, adaptive, decision-making, problem-solving, modelling, and other capabilities to help perform different educational tasks more effectively, including reviewing and grading students’ assignments, providing flexible and personalized learning experiences, and implicitly modelling students’ profiles (Chen et al., 2020a; Essalmi et al., 2017; Pan et al., 2021; Tlili et al., 2022a).

To provide comprehensive insights into AI’s use in education, several literature reviews have been conducted (e.g., Ouyang et al., 2022; Wang et al., 2023). Such studies are intended to provide a holistic perspective by analyzing the approaches and synthesizing the research findings across scholarly papers. Rowe (2014) stated that review papers can be grouped into four categories based on the type of contribution to theory, namely, describing, understanding, theory testing, and explaining a phenomenon (see Table 1).

Regardless of the category of a given review paper, a transparent, as well as a systematic process, among other factors, contributes to the formation of high-quality review papers and the production of new perspectives on the research field. Therefore, transparency is described as a meticulous and thorough reporting of methodological choices made during the review process (Templier & Pare, 2018). Explicit disclosure may strengthen the work and its conclusions’ credibility. Further, it also helps to ensure the internal and external reliability of the review processes. Transparency creates methodological rigour and repeatability of studies. Its importance in scientific research is increasingly emphasized across disciplines in the social and natural sciences (Tuval-Mashiach, 2017; McIntosh et al., 2017). Paré et al. (2016) further highlighted two limitations of non-transparent literature reviews: (1) lack of clarity when discussing the methodology of the study and (2) structural constraints of the publishing environment on producing extensive information regarding systematicity in the process of research.

Motivated by this background, and since transparency has been also one of the key dimensions that should be considered when implementing AI in general or in education particularly (Tlili et al., 2021; Larsson & Heintz, 2020), this study answers the following research question: What is the transparency level of the conducted literature reviews on AI in education? Specifically, this study conducts a systematic review to identify literature reviews on AIED in the literature, and then assess their transparency level following the Rowe (2014) and Pare et al. (2016) classification and transparency assessment metrics in review articles. In other words, this present study analyzes how transparent the authors from the literature were when adopting a given methodology for their literature reviews on AIED. The findings of this study can contribute to the AIED field by highlighting the transparency gaps of the conducted AIED literature reviews, hence consider them in the future when conducting a literature review. This can ensure more reliable, reusable and trustworthy findings on AIED that can advance the field. Despite the importance of the topic, no previous research, to the best of our knowledge, has conducted a similar analysis.

Table 1. The classification of review papers and their contributions to the theory.

<table>
<thead>
<tr>
<th>Overarching goal (Adopted from Rowe, 2014)</th>
<th>Types of literature reviews commonly accepted goals, and frequently researched questions (Adopted from Pare et al., 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describing</td>
<td>Narrative review: A narrative review summarizes what has already been reported/published on a specific topic and what do we know or have discovered about this topic? Descriptive review: A descriptive review analyzes the trends or patterns in available theories, hypotheses, techniques, or conclusions.</td>
</tr>
<tr>
<td>Understanding</td>
<td>Scoping review: A scoping review provides a broad overview of scientific knowledge on a particular topic. This form of review also permits identifying the gaps in the literature as well as future potential research directions. Critical review: A critical review is a critical evaluation that identifies flaws, contradictions, disagreements, or inconsistencies on a certain topic.</td>
</tr>
<tr>
<td>Testing theory</td>
<td>Meta-analysis: A meta-analysis is a method of combining statistical data from many quantitative research to provide relevant findings on a given issue. Qualitative systematic: A qualitative systematic literature review compiles statistical/empirical data on a certain topic and displays it in a narrative style. Umbrella review: The umbrella review examines and combines the qualitative systematic reviews and meta-analyses to produce the highest level of proof.</td>
</tr>
<tr>
<td>Explaining</td>
<td>Theoretical review: A theoretical review helps to develop new logical structure and framework by extending current ideas. The study objectives are stated, although there are usually no formal research questions. Realist review: A realist review determines what works for whom, under what conditions, in what areas, and also how.</td>
</tr>
</tbody>
</table>

Method

This study assesses the transparency of the conducted literature reviews on AIED. To identify these literature reviews, the recommended reporting items for systematic reviews and meta-analyses (PRISMA) criteria were followed (Page et al., 2021). A PRISMA technique is one of the standardized peer-reviewed methodologies that employ a guideline checklist to ensure the quality and reliability of the revision process.
Search strategy and selection criteria

An extensive search was undertaken in the following databases, namely: Web of Science, Scopus, Taylor & Francis, and Science Direct, IEEE Xplore, as they are very popular in the field of educational technology (Tlili et al., 2022b; Wang et al., 2023). Particularly, the following search string was used: (AI OR Artificial Intelligence OR machine learning OR deep learning or natural language processing) AND (education OR learning) AND (literature review OR systematic review OR meta-analysis or state-of-art). The search query was applied to titles and abstracts, and the search keywords were partially adopted from Zawacki-Richter et al. (2020). After searching the appropriate databases, two authors individually analyzed the extracted papers by titles, abstracts, and textual on the inclusion and exclusion criteria reported in Table 2. During this phase, to reach a final consensus, disagreements between the authors were resolved through discussion or arbitration by a third author who has experience in AI research.

Table 2. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal article</td>
<td>Conference proceedings, dissertations, novels, book series, and book chapters</td>
</tr>
<tr>
<td>Literature review</td>
<td>Not a literature review article or an article that does not explain how the literature review was conducted.</td>
</tr>
<tr>
<td>Papers focusing on AI in education</td>
<td>Papers not focusing on AI or discussing AI, not in education</td>
</tr>
<tr>
<td>Accessible online papers</td>
<td>Not accessible online</td>
</tr>
<tr>
<td>Papers in English</td>
<td>Papers not in English</td>
</tr>
</tbody>
</table>

The search yielded a total of 1,367 articles, where 1,330 articles remained after removing duplicates. The screening of titles and abstracts resulted in the removal of 1,009 articles. The remaining 321 papers were considered and assessed as a full text. 260 of these articles failed to meet the criteria for inclusion. As a result, 61 research articles were suitable to be included in this study (see Table 3).

Table 3. 61 included studies in this systematic literature review.

<table>
<thead>
<tr>
<th>No.</th>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hanman &amp; Liu, 2023</td>
<td>AI: new source of competitiveness in higher education</td>
</tr>
<tr>
<td>2</td>
<td>Su et al., 2023</td>
<td>Artificial intelligence (AI) literacy in early childhood education: The challenges and opportunities</td>
</tr>
<tr>
<td>3</td>
<td>Bamshad &amp; et al., 2022</td>
<td>A systematic review of the role of learning analytics in enhancing feedback practices in higher education</td>
</tr>
<tr>
<td>4</td>
<td>Bilgic et al., 2022</td>
<td>Exploring the roles of artificial intelligence in surgical education: A scoping review</td>
</tr>
<tr>
<td>5</td>
<td>Chu &amp; et al., 2022</td>
<td>Artificial intelligence-based robots in education: A systematic review of selected SSCIT publications</td>
</tr>
<tr>
<td>6</td>
<td>Dai &amp; Ke, 2022</td>
<td>Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review</td>
</tr>
<tr>
<td>7</td>
<td>Kurnasiksa et al., 2022</td>
<td>Artificial Intelligence and Surgical Education: A Systematic Scoping Review of Interventions</td>
</tr>
<tr>
<td>8</td>
<td>Laupichler et al., 2022</td>
<td>Artificial intelligence literacy in higher and adult education: A scoping literature review</td>
</tr>
<tr>
<td>9</td>
<td>Maier &amp; Klotz, 2022</td>
<td>Personalized feedback in digital learning environments: Classification framework and literature review</td>
</tr>
<tr>
<td>10</td>
<td>Murtaza et al., 2022</td>
<td>AI-Based Personalized E-Learning Systems: Issues, Challenges, and Solutions</td>
</tr>
<tr>
<td>11</td>
<td>Su &amp; Yang, 2022</td>
<td>Artificial intelligence in early childhood education: A scoping review</td>
</tr>
<tr>
<td>12</td>
<td>Su et al., 2022</td>
<td>A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region</td>
</tr>
<tr>
<td>13</td>
<td>Tan et al., 2022</td>
<td>A systematic review of artificial intelligence techniques for collaborative learning over the past two decades</td>
</tr>
<tr>
<td>14</td>
<td>Xin et al., 2022</td>
<td>Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education</td>
</tr>
<tr>
<td>15</td>
<td>Xu &amp; Ouyang, 2022</td>
<td>A systematic review of AI role in the educational system based on a proposed conceptual framework</td>
</tr>
<tr>
<td>16</td>
<td>Zafiri et al., 2022</td>
<td>Artificial Intelligence Applications in K-12 Education: A Systematic Literature Review</td>
</tr>
<tr>
<td>17</td>
<td>Ahmad et al., 2021</td>
<td>Artificial intelligence and its role in education</td>
</tr>
<tr>
<td>18</td>
<td>Almen &amp; Alharghi, 2021</td>
<td>Explainable Student Performance Prediction Models: A Systematic Review</td>
</tr>
<tr>
<td>20</td>
<td>Balaji et al., 2021</td>
<td>Contributions of machine learning models towards student academic performance prediction: A systematic review</td>
</tr>
<tr>
<td>21</td>
<td>Brockwell et al., 2021</td>
<td>Artificial intelligence and reflections from educational landscape: A review of AI stories in half a century</td>
</tr>
<tr>
<td>22</td>
<td>Gonsalves-Calayud et al., 2021</td>
<td>Artificial intelligence for student assessment: A systematic review</td>
</tr>
<tr>
<td>23</td>
<td>Hahn et al., 2021</td>
<td>A Systematic Review of the Effects of Automatic Scoring and Automatic Feedback in Educational Settings</td>
</tr>
<tr>
<td>24</td>
<td>Huang et al., 2021</td>
<td>A Review on Artificial Intelligence in Education</td>
</tr>
<tr>
<td>25</td>
<td>Huang &amp; Chang, 2021</td>
<td>A review of opportunities and challenges of chatbots in education</td>
</tr>
<tr>
<td>26</td>
<td>Huang &amp; Tu, 2021</td>
<td>Roles and research trends of artificial intelligence in education: A bibliometric mapping analysis and systematic review</td>
</tr>
<tr>
<td>27</td>
<td>Kahwali et al., 2021</td>
<td>AI-enabled adaptive learning systems: A systematic mapping of the literature</td>
</tr>
<tr>
<td>28</td>
<td>Kahn &amp; Wasowski, 2021</td>
<td>Constructionism and AI: A history and possible futures</td>
</tr>
<tr>
<td>29</td>
<td>Kharbat et al., 2021</td>
<td>Identifying gaps in using artificial intelligence to support students with intellectual disabilities from education and health perspectives</td>
</tr>
<tr>
<td>30</td>
<td>Kibucukhojev et al., 2021</td>
<td>Some Aspects of AI Technologies in Education</td>
</tr>
<tr>
<td>31</td>
<td>Lee et al., 2021</td>
<td>Artificial Intelligence in Undergraduate Medical Education: A Scoping Review</td>
</tr>
<tr>
<td>32</td>
<td>Liang et al., 2021</td>
<td>Roles and research focus of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach</td>
</tr>
<tr>
<td>33</td>
<td>Liu &amp; Aitazi, 2021</td>
<td>Artificial Intelligence (AI) and Translation Teaching: A Critical Perspective on the Transformation of Education</td>
</tr>
<tr>
<td>34</td>
<td>Loan &amp; Tsai, 2021</td>
<td>A Review of Using Machine Learning Approaches for Practical Education</td>
</tr>
<tr>
<td>35</td>
<td>Maghrabi et al., 2021</td>
<td>Personalized Education in the Artificial Intelligence Era: What to Expect Next</td>
</tr>
<tr>
<td>36</td>
<td>Ng et al., 2021</td>
<td>Conceptualizing AI literacy: An exploratory review</td>
</tr>
<tr>
<td>37</td>
<td>Nigam et al., 2021</td>
<td>A Systematic Review on AI-Based Practicing Systems: Past, Present and Future</td>
</tr>
<tr>
<td>38</td>
<td>Seherencio et al., 2021</td>
<td>Systematic literature review on machine learning and student performance prediction: Critical gaps and possible remedies</td>
</tr>
<tr>
<td>39</td>
<td>Seura et al., 2021</td>
<td>The Potential of AI in Health Higher Education to Increase the Students’ Learning Outcomes</td>
</tr>
<tr>
<td>40</td>
<td>Taha, 2021</td>
<td>AI in education: A systematic literature review</td>
</tr>
<tr>
<td>42</td>
<td>Valencia-Casas, 2021</td>
<td>Artificial intelligence and education: A pedagogical challenge for the 21st century</td>
</tr>
<tr>
<td>43</td>
<td>Zhao et al., 2021</td>
<td>A Review of Artificial Intelligence (AI) in Education from 2010 to 2020</td>
</tr>
</tbody>
</table>
Following the PRISMA guidelines, the study selection process is presented in Figure 1.

Figure 1. PRISMA Flowchart of the systematic review process.

Coding scheme

To assess the transparency of each identified literature review (among the 61 studies), this study uses the recommendation by Paré et al. (2016) on transparency and systematicity, which includes 17 questions split into six categories (see Table 4). Each paper was coded according to the information on the transparency characteristics to assess the level of its transparency. Specifically, for each question in Table 4, if the information exists, "Y" standing for "Yes" was assigned; otherwise, "N" standing for "No" was assigned. Particularly, all items with value = "Y" were counted and divided by the number of items in their group to calculate the transparency level in each group (e.g., for S01, we divided by 3, for S02, we divided by 4, for S04, we divided by 2, etc.). The researchers grouped the subtotal by the set of groups (six groups) throughout the assessment schema to calculate the overall assessment level. The data-gathering procedure was carried out throughout the article to reduce the risk of incomplete information, which is not mentioned in the methodology section. To reduce the opportunity for bias, an electronic data extraction form was designed (Tlili et al., 2022b), where two coders filled it in according to the coding scheme (see Table 5). To further ensure the reliability of the coding results, weekly meetings during the whole coding process were organized between the coders to discuss their coding progress, where disagreements were discussed and resolved by consensus.

Table 4. Transparency assessment in review articles (Paré et al., 2016).

<table>
<thead>
<tr>
<th>Review steps to be assessed</th>
<th>Elements to be assessed in each step</th>
</tr>
</thead>
<tbody>
<tr>
<td>S01 – Review planning</td>
<td>1. Are the objectives of review articles well argued?</td>
</tr>
<tr>
<td></td>
<td>2. Is the review type and procedures well described and justified in the study? (For reviews articles, we established frameworks or recommendations.)</td>
</tr>
<tr>
<td></td>
<td>3. Is the procedure or protocol described and published?</td>
</tr>
<tr>
<td>S02 – Search strategy</td>
<td>4. Is the search technique (for example, databases with coverage dates) well-defined?</td>
</tr>
<tr>
<td></td>
<td>5. Are the criteria for inclusion and exclusion explicitly mentioned?</td>
</tr>
<tr>
<td></td>
<td>6. Is there a complete electronic search technique for at minimum one database? (search terms and keywords)</td>
</tr>
<tr>
<td></td>
<td>7. Is there information about reference management tools and techniques, as well as other research processes?</td>
</tr>
<tr>
<td>S03 – Study selection</td>
<td>8. Is there a description of the screening and selection processes of the study mentioned?</td>
</tr>
<tr>
<td></td>
<td>9. Are there enough details about the included studies?</td>
</tr>
<tr>
<td></td>
<td>10. (if applicable) Is there a list of excluded studies, together with the reasons for the exclusion mentioned?</td>
</tr>
<tr>
<td></td>
<td>11. Is there a flow diagram depicting the study selection process?</td>
</tr>
<tr>
<td>S04 – Quality assessment</td>
<td>12. Are the results of each study’s quality assessment presented?</td>
</tr>
<tr>
<td></td>
<td>13. Is there a description of the methods for integrating assessments into analyses?</td>
</tr>
<tr>
<td>S05 – Data extraction strategy</td>
<td>14. Are the methods and processes for data extraction mentioned?</td>
</tr>
<tr>
<td></td>
<td>15. Are the data extracted forms or items presented?</td>
</tr>
</tbody>
</table>
To assess the transparency level among the six criteria/dimensions (see Table 4), the following information in Table 5 was coded. This information can help to provide comprehensive and deep insights related to the 17 items within the six steps (see Table 4).

<table>
<thead>
<tr>
<th>Table 5. Coding scheme.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Code</strong></td>
</tr>
<tr>
<td>Article keywords</td>
</tr>
<tr>
<td>Electronic databases</td>
</tr>
<tr>
<td>Keywords and search terms</td>
</tr>
<tr>
<td>Years included</td>
</tr>
<tr>
<td>Audience</td>
</tr>
<tr>
<td>Type of primary sources</td>
</tr>
<tr>
<td>Number of studies</td>
</tr>
<tr>
<td>General purpose</td>
</tr>
<tr>
<td>Research questions or Hypothesis</td>
</tr>
<tr>
<td>Review paper</td>
</tr>
<tr>
<td>Author’s review paper type</td>
</tr>
<tr>
<td>Coders’ review paper type</td>
</tr>
<tr>
<td>Scope of the research question</td>
</tr>
<tr>
<td>Search strategy</td>
</tr>
<tr>
<td>Articles explicitly mention the selection processes</td>
</tr>
</tbody>
</table>

**Results and discussion**

**Descriptive summary of the AIED systematic reviews**

Table 6 shows the goal of the included 61 literature reviews on AIED. These articles were published between 2012 and 2023. Additionally, based on the overarching goal of a literature review (Table 1), 57.36% of the AIED reviews had a primary purpose of describing a phenomenon with little or no addition to the theory (Rowe et al., 2012), as 54.09% of them were descriptive, while only 3.27% were narrative. Table 6 shows that the second highest type of review article is the critical review (31.14%) which comes under understanding. In contrast, review papers other than descriptive and critical reviews are underrepresented, calling for more research in this context to cover those less applied types of literature review. It is noteworthy that the descriptive type’s dominance probably stems from the relatively recent history of AI technologies. It is also important to note that the novelty effect of AI technologies expectedly urges researchers to investigate the phenomenon through descriptive review designs (Johnson et al., 2022). However, this tendency creates an imbalance and is a potential drawback to gaining a deeper understanding of AIED.

**Table 6. The classification of systematic review articles (type and year) 2012–2023.**

**Transparency assessment of the literature reviews**

This section assesses the transparency criteria of each step (the six steps, see Table 4) of the reviewed articles. Each step is discussed in a subsequent section, while the overall transparency level is discussed in the final section.
Developing a review plan (S01)

This section assesses the transparency of the step of developing a review plan. The most important aspects of guaranteeing systematicity are planning the strategy, identifying the problem, proclaiming the purpose and research questions, and selecting and explaining the review type (Paré et al., 2016). Creating a review plan further improves the review process’s systematicity and serves as the foundation for more extensive reporting of methodological decisions made during the study process (Templier & Pare, 2018). As demonstrated in Table 7, 81% (n = 50) of the review articles explicitly mentioned the targeted audience (researchers, practitioners, policy-makers, teachers, analysts, and students). Particularly, more than 70% of the reviewed articles centred on teachers as the audience when discussing AIED. Further, as shown in Table 6, 100% (n = 61) of the reviewed articles clearly stated their objectives or purposes by explicitly mentioning the research questions or hypotheses. Additionally, 81% (n = 50) of the studies clearly mentioned the scope of the research question (see Table 7).

We classified the literature reviews according to the type and breadth of their research questions based on their stated objectives and research questions (when available). All articles (n = 61) declared explicitly that their work was a review paper (i.e., the authors stated directly that their study was a review paper) and the authors’ review paper type. However, only 50% (n = 31) of the articles mentioned the coders’ review paper type (review types assigned by coders). Additionally, in the studies included, justifications for the review type selection were not identified. Furthermore, in the descriptive reviews, the authors stated their review objectives, types, and protocols at higher rates than in narrative, critical, scoping, or meta-analysis reviews. The use of explicit frameworks and guidelines for undertaking a literature review helps to explain some of these findings (Snyder, 2019).

Based on the above results, it is very important that authors put more attention on the rationale for choosing to conduct one type of literature review and not the other based on the research questions to be answered. For instance, if the authors want to measure the impact of a specific educational intervention, conducting a meta-analysis would be the most adequate type for this objective. Adding such information could increase the transparency of the conducted literature review, and help readers understand the ultimate goal of conducting a given literature review generally, and on AIED particularly.

Searching the literature (S02)

As shown in Table 7, 55% of the articles reported their search methods, which include databases and timelines with a clear description. Specifically, 75% of the articles reported the review period, and 81% of the articles reported the search queries. Additionally, all articles (100%) reported the electronic databases used to identify the research corpus. Particularly, when analyzing the mentioned electronic databases (in all 61 articles), Scopus (n = 25) was the most preferred electronic database to identify and review AIED papers, followed by Web of Science (n = 15), Science direct (n = 15) and IEEE Xplore (n = 6). However, only 45% of the articles explicitly reported the search queries used during the process of building a research corpus, and only 8% of the articles reported the tools and methods used to manage bibliographic materials.

Based on the above results, it is seen that authors should elaborate more in terms of the used search keywords to conduct their systematic reviews on AIED. This is crucial, especially for a complex topic like AIED, where several technologies (e.g., deep learning, machine learning, natural language processing) and techniques (e.g., learning analytics, prediction and modelling) can be used interchangeably with AIED. Besides, it is seen that less attention has been put by authors to elaborate on the used tools to manage their bibliographic materials, and this could be because these tools are more managerial and do not impact the research quality in any way.

Selecting studies (S03)

The selection process of articles provides important information on the sources used for interpretation and analysis, as well as the techniques used to pick these sources, by disclosing information about these aspects (Tricco et al., 2011). Consequently, it highlights the papers that are not relevant to the researchers’ search interests. A
comprehensive and detailed screening approach decreases
the risk of bias when it comes to including or excluding
articles for further research. Researchers also give evidence
on the usefulness of these studies for generating relevant
results and addressing the study questions by giving
thorough information about the included and excluded
articles.

As presented in Table 7, 78% (n = 48) of the review articles
mentioned the number of studies included in the review
process. It is also critical for researchers and practitioners
to have precise and organized information regarding exclusion
methods since this allows them to assess the criteria’s
soundness and scientific rigour. Moreover, reporting
inclusion and exclusion criteria provide insights that lead to
generating research findings through reviewing the related
literature and providing the base criteria for replicating the
study and benchmarking the research process. Specifically,
70% of the articles mentioned their inclusion/exclusion
criteria. 50% of the articles revealed the list of the profile
of included articles, and just 34% of the articles provided
information about the inclusion and exclusion criteria.
Particularly, 50% (n = 31) of the publications presented the
list of studies that were included and 45% (n = 28) of the
studies graphically published data.

Inclusion/exclusion criteria are an important step to help
readers understand how a given study might or might not be included within a given literature review. Based on
the obtained results, it is seen that most studies did not
elaborate on this step, as well as the final list of included
studies. Consequently, this makes those conducted AIED
literature reviews a black box, where it is not clear what
was included and why. This also hinders understanding the
obtained results, hence making full use of them to advance
the AIED field, as the literature review input (i.e., included
studies and how they are selected) is absent.

Assessing quality (S04)

This information can guarantee that only high-quality
sources are obtained as a method of improving the quality
of the findings and outcomes (Bandara et al., 2015). As
shown in Table 7, only 8% of the articles thoroughly reported
the quality assessment results for specific resources, and 6%
of the articles reported information on quality assessment
methods. On the other hand, 73% of the studies clearly
mentioned the nature of the primary source. However, only
55% of the review studies mentioned the quality appraisal,
where they compared the covered articles collectively. These
findings are alarming, as assessing quality is related to the
robustness of the research conducted.

Therefore, to increase the adoption and use of the
obtained results given by some AIED literature reviews,
the quality assessment should be highlighted. This is
because researchers might always be hesitant to rely on
some findings that they are not sure of their quality. This
is even more pertinent in the AIED field as designing AI-
based educational systems is very tricky and requires careful
attention to not accidently harm (e.g., biased interventions)
users (e.g., learners, educators) instead of supporting them.

Extracting data or key aspects from included studies (S05)

Whittemore et al. (2014) stated that the accurate reporting
of individual research is vital to improving the quality of
any knowledge synthesis technique. This aims to identify,
organize, and carry out agreed-upon methods for collecting
data from primary sources to mitigate the risks of omission,
misclassification, or misrepresenting crucial information
(Paré et al., 2016; Webster et al., 2002).

As shown in Table 7, only 16% of the articles reported
their data extraction techniques and methods, whereas 59% of
the articles disclosed the specific items or data extraction
types for structured data collection. Furthermore, 95% of the
reviewed articles included the items or information needed
to extract data from their primary sources (e.g., descriptive,
narrative, evidence-based qualitative, evidence-based
quantitative, conceptual or theoretical paper, critical paper,
bibliometric, mixed methods, or literature review). Some
articles utilized a list of items to display this information
(e.g., Al-Azawei et al., 2016), whereas others used structured
approaches with more specific information (e.g., Arbaugh,
2014).

Based on the obtained results, authors should elaborate
more on their extraction techniques, especially their coding
scheme, to increase the replicability of their research (i.e.,
literature reviews) by others.

Synthesizing and interpreting data and formulating
conclusions (S06)

As shown in Table 7, 60% of the studies indicated the major
constructs or outcomes, while 59% described the analytical
and synthesis methodologies. In both cases, the number
of descriptive reviews outnumbered the remainder of the
review articles.

This result might reveal that it is always easier to elaborate
on some descriptive analysis, while it is not the case when
the analysis is more advanced. For instance, when discussing
meta-analysis review papers, there is a need to go beyond
the simple description of the process and elaborate on the
motivation of selecting a given technique, for instance,
related to measuring effect size (Cohen’s d and Hedges’
g) or publication bias based on the different samples or
research methods used in each included study within the
review process.

The overall level transparency (S07)

Table 8 summarizes the overall level of transparency of the
61 review articles on AIED arranged by type and quartiles,
with Q1 being the highest and Q4 the lowest. Specifically,
a study belongs to Q1 if its overall transparency level is
between 76% and 100%, Q2 if its overall transparency level
is between 51% and 75%, Q3 if its overall transparency level
is between 26% and 50%, and Q4 if its overall transparency
level is between 0% and 25%. When describing research
efforts, the findings reflect various levels of transparency;
40% of the articles were in Q1 and 32% in Q2. This implies
that the transparency level of the conducted literature reviews on AIED is low. Therefore, future research should consider the transparency factor of the conducted literature reviews on AIED, as this may provide detailed insights about the field and positively impact the scientific community more broadly (Vom Brocke et al., 2018).

Additionally, descriptive review articles were in the top two quartiles (Q1 and Q2), while critical review articles (11 out of 13) were in the last quartile (Q4), implying that descriptive review articles have the highest transparency level while critical review articles have the lowest. Contrary to our findings, Castro-Gil and Correa (2021) found that the lowest transparency level was in descriptive literature reviews on blended learning in higher education.

Furthermore, Table 8 shows that step four “quality assessment” had the lowest transparency level. This implies that the reported reviews on AIED did not explicitly discuss the quality of the reviewed articles. Consequently, this might hinder the quality of the reported findings related to AIED. In this context, to ensure quality assessment when conducting review articles, several studies focused only on reviewing SSCI/SCIE or top journals in the field (e.g., Crompton & Burke, 2018; Hwang & Tsai, 2011).

Table 8. Studies fulfilling the transparency assessment items by type of review paper and quartile.

<table>
<thead>
<tr>
<th>Review type and number of studies</th>
<th>Reviews</th>
<th>Steps of the review process (SRM-SRO) and transparency assessment items (S1-S7)</th>
<th>Quartiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Q1</td>
<td>Q2</td>
</tr>
<tr>
<td>Descriptive</td>
<td>33</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td>Narrative</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Narrative</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Critical</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Critical</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scoping</td>
<td>6</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Scoping</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Theoretical</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Theoretical</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>45</td>
<td>4</td>
</tr>
</tbody>
</table>

Conclusions, implications, and limitations

This study conducted a transparency assessment of AIED review articles. The obtained findings showed that the transparency level is considerably low. Specifically, researchers should focus more on elaborating on the quality assessment of the reviewed articles, as well as the included and excluded articles.

The findings of this study can contribute to the educational technology field from different perspectives. From a theoretical perspective, this study can enrich the ongoing debate about the dimensions to consider for applying a transparent systematic review generally, and on AIED particularly. From a methodological perspective, this study presents how to conduct a transparency assessment of articles, as well as how to enhance the methodological part of a given literature review to obtain valid and reproducible research by others. From a practical perspective, this study can contribute to the AIED field by highlighting to researchers and practitioners the weaknesses of the conducted AIED literature reviews. Enhancing these parts can contribute to enhancing the quality of the obtained findings related to AIED, hence providing more insights to the working community on AIED, as well as providing evidence-based practices or making decision processes related to AIED.

Despite the solid ground of this study, it has some limitations that should be acknowledged. For instance, the covered literature review articles in this study might be limited due to the search queries and electronic databases used. Therefore, interested researchers can further complement the research presented in this study. Future research can focus on going beyond assessing the transparency level to analyze how the conducted literature reviews tackled AIED (e.g., from which perspective, the targeted stakeholders, education level and context, etc.). This might reveal the trends of AIED, as well as the gaps that researchers and practitioners should focus on in the future.

References


Journal of Applied Learning & Teaching Vol.6 No.2 (2023)
Artificial intelligence (AI) is increasingly integrated into education and training contexts as a means to support learning, teaching, and assessment. In this context, we first introduce an AI-based model to support task-based learning. Then, we present a survey on the use of AI in education. Finally, we discuss future directions in AI for education.


doi:10.1016/j.techfore.2022.121636


