Are open educational resources (OER) and practices (OEP) effective in improving learning achievement? A meta-analysis and research synthesis

Ahmed Tlili1, Juan Garzón2, Soheil Salha3, Ronghuai Huang1, Lin Xu1, Daniel Burgos4, Mouna Denden5,6, Orna Farrell7, Robert Farrow8, Aras Bozkurt9, Tel Amiel10, Rory McGreal11, Aída López-Serrano4 and David Wiley12

Abstract
While several studies have investigated the various effects of open educational resources (OER) and open educational practices (OEP), few have focused on its connection to learning achievement. The related scientific literature is divided about the effects of OER and OEP with regards to their contribution to learning achievement. To address this tension, a meta-analysis and research synthesis of 25 studies (N = 119,840 participants) was conducted to quantitatively investigate the effects of OER and OEP on students’ learning achievement. The analysis included course subject, level of education, intervention duration, sample size, geographical distribution, and research design as moderating variables of the obtained effects. The findings revealed that OER and OEP have a significant yet negligible (g = 0.07, p < 0.001) effect. Additionally, the analysis found that the obtained effect can be moderated by several variables, including course subject, level of education and geographical distribution. The study findings can help various stakeholders (e.g., educators, instructional designers or policy makers) in understanding what might hinder OER and OEP effect on learning achievement, hence accommodating better learning outcomes and more effective interventions.

Keywords: Open educational resources, Open educational practices, Learning achievement, Meta-analysis, Meta-synthesis

Introduction
Open educational resources and practices
The term Open Educational Resources (OER) was first coined at UNESCO’s 2002 Forum on Open Courseware, and it was defined in the recent UNESCO Recommendation on OER as “learning, teaching, and research materials in any format and medium that reside in the public domain or are under copyright that have been released under an open license that permit no-cost access, reuse, repurpose, adaptation, and redistribution...
by others” (UNESCO, 2019). Several studies have then reported the advantages of OER in reducing learning costs (Hilton, 2016), increasing accessibility to educational resources even for students with disabilities (Zhang et al., 2020a), and enhancing learning quality (Yuan & Recker, 2015; Weller et al., 2015; Zhang et al., 2020b). Wiley (2014) further outlined five key characteristics, also known as the 5Rs, of using OER, namely: (1) retain—each person has the right to make and own copies of the published resource; (2) reuse—each person has the right to use the educational resources content in different ways depending in the learning context (e.g., formal or informal learning); (3) revise—each person has the right to revise the educational resource for different purposes (e.g., adapting it to a learning context or enhancing it); (4) remix—each person has the right to create a new educational resource by combining one or more learning contents together; and (5) redistribute—each person has the right to share with others copies of the original revised or remixed educational resource. The 5Rs can support innovation in teaching and learning since OER can be created, used, shared and repurposed differently to traditional copyrighted educational materials.

Building on the idea of innovation in educational resources and the idea of openness in education (Bozkurt et al., 2023), the Open e-Learning Content Observatory Services (OLCOS) functions as a Transversal Action under the European eLearning Programme and is committed to advancing the creation, sharing, and global utilization of OER (OLCOS, 2007). In 2007, OLCOS conducted a roadmap study that emphasized the significance of integrating innovative teaching methods with OER (OLCOS, 2007). The project underscores that merely delivering OER within traditional teacher-centered frameworks might not sufficiently prepare individuals for educational success. It advocates for the incorporation of innovative educational practices alongside OER, and notably introduced the concept of Open Educational Practices (OEP). Based on this perspective, OEP can be defined as OER-enabled pedagogies, or “the set of teaching and learning practices that are only possible or practical in the context of the 5R permissions which are characteristic of OER” (Wiley & Hilton III, 2018, p. 135; cf. Bali et al., 2020). Ehlers (2011, p. 4) defined OEP as “practices which support the (re)use and production of Open Educational Resources through institutional policies, promote innovative pedagogical models, and respect and empower learners as co-producers on their lifelong learning paths.” In a comprehensive review, Huang et al. (2020) identified five dimensions for the possible implementation of OEP, namely: OER, open teaching, open collaboration, open assessment and facilitating technologies. Some research suggests that these practices can help enhance learning quality, access, and effectiveness in universities. With the positive potential of OER and OEP in education, their adoption in education has rapidly increased for the past years. A significant moment in the history of open education came with the UNESCO (2019) Recommendation on OER which provides strategic policy support for the uptake and monitoring of OER. Accordingly, the UNESCO recommendation calls upon member states to develop national policies for the adoption of OER, which include activities, such as creating guidelines and strategies to incorporate OER within educational institutions or facilitating the generation and sharing of OER materials among educators. This recommendation draws considerable attention and investments to OER and OEP projects without certainty about their positive effects. At present, with the great potential of OER and OEP in education, a majority
who remain unaware of the transformative potential of open practice; some educators consider OEP to be one of the most significant teaching forms of the twenty-first century (Shear et al., 2015) while others are oblivious of its existence. It is also important to note that OEP is not an orthodoxy so much as a concept that can be realized in a multitude of different ways.

Research gap and study objectives
Dotson and Foley (2017) emphasized that changing the curriculum content (i.e., from proprietary to open) does not produce a change in students’ learning achievement. Harvey and Bond (2022) also argued that there is a need to investigate if a change in the learning content licensing has an impact on students’ learning achievement. Despite a growing body of evidence regarding the effectiveness of OER and OEP in learning, open research studies have focused on other variables (e.g., affordability, accessibility). Less attention has been paid to whether OER and OEP can enhance students’ learning achievement compared to traditionally copyrighted materials (Robinson, 2015). For instance, Hilton (2016) conducted a systematic review of articles focusing on OER issues and learning achievement and perception, written between 2002 and August of 2015. The researcher found that only seven of sixteen studies focused on learning achievement. The researcher conducted another systematic review of twenty-nine OER-focused articles, written between September 2015 and December 2018, and only nine new learning achievement studies were obtained (Hilton, 2020). This reflects the decline in attention being paid to OER/OEP and learning achievement since 2002. Moreover, the literature about the effects of OER and OEP in enhancing students’ learning achievement is divided, where some studies reported positive effects (e.g., Colvard et al., 2018), no effects (e.g., Fortney, 2021; Grissett & Huffman, 2019) or even negative effects (e.g., Gurung, 2017), implying that some students who used traditionally copyrighted materials had better effects than those who used OER.

The question of the relative efficacy of OER or OEP remains open. The main rationale, therefore for this study, is to examine whether or not OER and OEP can enhance learning achievement. Two systematic reviews (Hilton, 2016, 2020) attempted to investigate the above-mentioned phenomenon, however they were purely qualitative. The results from these two reviews did not effectively reveal the effects of OER and OEP on learning achievement. One study by Clinton and Khan (2019) conducted a meta-analysis related to this topic, however it investigated only the effect of open textbooks on post-secondary students’ learning achievement in the USA and Canada. Consequently, the previously obtained results do not reflect a comprehensive and an in-depth investigation of the effect of OER and OEP on learning achievement.

This present investigation aims at a more in-depth coverage of the current literature by including a range of types of OER (e.g., textbooks, videos, etc.) in many countries and at many educational levels. Smith (2013) highlighted the importance of researching improvements in achievement and attainment of OER, urging for further investigation into interventions that could result in significant enhancements in educational outcomes. In the same vein, Hilton (2020) has further suggested conducting sophisticated meta-analyses, where effect sizes across studies are calculated, to understand the measurable effect of OER on learning achievement. In response, this study employs a
systematic analysis of the OER/OEP literature to comprehensively investigate whether the data supports the hypothesis that the use of OER and OEP can improve students’ learning achievement in a range of subjects. Therefore, to address this research gap, this study consisted of a meta-analysis and research synthesis of the relevant literature to provide quantitative evidence on the effects of OER and OEP on learning achievement. Meta-analysis, utilizing statistical methods, was employed to accurately measure the effect of a given intervention and the associated moderators of this effect (Rosenthal & DiMatteo, 2001).

Additionally, several studies reported that the effects of OER and OEP on learning achievement might vary due to different confounders, such as demographic information, the type of the course delivered, educational level (grade), intervention duration, among others (e.g., Hilton, 2016, 2020). Therefore, the present study takes a forward step towards analyzing if these variables might moderate the effect of OER and OEP on learning achievement. Specifically, this study addressed the following research questions:

- RQ1. What is the effect of OER and OEP on students’ learning achievement?
- RQ2. How does the effect of OER and OEP on students’ learning achievement vary according to the educational subject?
- RQ3. How does the effect of OER and OEP on students’ learning achievement vary according to the educational level?
- RQ4. How does the effect of OER and OEP on students’ learning achievement vary according to the intervention duration?
- RQ5. How does the effect of OER and OEP on students’ learning achievement vary according to the sample size?
- RQ6. How does the effect of OER and OEP on students’ learning achievement vary according to geographical distribution of students?
- RQ7. How does the effect of OER and OEP on students’ learning achievement vary according to the research design?

**Methodology**

This study identifies the effects of using OER and OEP on learning achievement through meta-analysis. To secure the selection of the most relevant literature to be meta-analyzed, the researchers of the current study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Page et al., 2021). Additionally, the researchers followed recommendations outlined by Kitchenham and Charters (2007). This procedure suggests three stages, namely: planning, conducting, and reporting the review. Although these guidelines were originally proposed for conducting systematic reviews, they have been successfully employed in meta-analyses (e.g., Garzón et al., 2019). All the processes related to the selection and codification of the studies were carried out by two coders.

**Planning the review**

To ensure having only relevant studies (recall) within this meta-analysis, hence obtaining a high precision rate (Ting, 2010), “open educational resources” and “open educational practices” were used as search keywords. Particularly, the term abbreviations, namely
OER and OEP, were not considered as search keywords because in scientific writings, the full name of a term is provided prior to using its abbreviation. The search process was undertaken in the following databases: Web of Science, Scopus, Taylor and Francis and ERIC. These databases were selected because they are popular in the field of educational technology (Bedenlier et al., 2020; Wang et al., 2023). ERIC, particularly, focuses on educational science, especially OER (Otto et al., 2021) and Scopus is known as the largest database for scholarly publications. The publication interval was from 2012 up to 2023. The starting year of 2012 was selected as the initial date because it marked the release of the “UNESCO Paris OER Declaration”, which urged governments to promote the use of OER, and called for publicly funded educational materials to be released in a freely reusable form. As a result, several OER initiatives were launched worldwide which catalyzed the development of the OER field. Due to the novelty of the topic, conference papers and doctoral dissertations were considered to be included into the research corpus, as suggested by several studies (e.g., Chen et al., 2020; Denden et al., 2022).

The search was conducted on April 4, 2023, at which date, researchers were able to identify 643 studies (Web of Science: 117, Scopus: 38, Taylor and Francis: 262, and ERIC: 226). After eliminating duplicates ($n = 324$), a total of 319 publications were selected for further analysis. The first filter was based on each article’s title and keywords. This process allowed us to identify and remove 75 papers that were not relevant to the purpose.

![PRISMA flowchart for the search protocol](image-url)

**Fig. 1** PRISMA flowchart for the search protocol
of this present study. Then, the abstract of the remaining 244 papers was read and analyzed comprehensively. This process allowed us to remove 135 papers that were not relevant. Finally, we analyzed the remaining 109 studies based on the following criteria: (1) empirical studies, (2) studies that specifically used OER or OEP, (3) studies that provided sufficient information (i.e., mean, median, standard deviation) to calculate the effect size.

Therefore, a study was excluded if (1) it was not empirical research, (2) it did not focus on using OER or OEP, (3) it was qualitative or review research, (4) it did not provide sufficient information to calculate the effect size, or (5) it was not written in English. This process limited the corpus for investigation to 25 papers (23 journal papers, 1 conference paper and 1 PhD dissertation) to be further examined and included in the analysis. At the end of this process, the reference section of each paper was then reviewed. However, this process did not provide additional studies. Figure 1 shows the PRISMA flowchart (Page et al., 2021) of the study selection process, where inter-rater reliability in each phase was above 0.7, which is considered very good (Cohen, 1960).

Conducting the review
This stage included the coding scheme for the data extraction process. In an effort to minimize the potential for bias, an online electronic data extraction form was designed (Kitchenham & Charters, 2007). To answer the aforementioned research questions, the following information in each study was coded: (1) OER type: The type of resource used for teaching, such as textbooks, videos, etc.; (2) course subject: the subject that was taught when using OER and OEP, such as mathematics, psychology, etc.; (3) educational level: the student grade in which OER and OEP were used, such as primary, bachelors, etc.; (4) the length of time over which OER and OEP were used (i.e., course duration); (5) sample size: the number of participants in each study. According to Cheung and Slavin (2016), the sample size was divided into small, where the number of participants is less than or equal to 250, and large, where the number of participants is larger than 250; (6) region: the region (country) where the experiment was conducted; and (7) research design: the followed research design when conducting the experiment.

Calculation of the effect size
Comprehensive Meta-Analysis V.4 (Borenstein, 2022) software was used to conduct the present meta-analysis. Additionally, Hedges’ g was used to calculate the effect sizes (Hedges, 1981). The motivation behind using Hedges’ g instead of Cohen’s d effect size was that the differential sample size between studies may bias the estimated effect size. This bias affects studies having a sample size smaller than 20, in which case Hedges’s g presents more reliable estimates than Cohen’s d (Hedges & Olkin, 1985). Eleven studies followed the pretest–posttest-control (PPC) research design. In this research design, students are randomly assigned to experimental and control treatments and are evaluated before and after the treatment. As stated by Morris (2008), this design provides better results regarding the accuracy of d values and control of threats to internal validity. The remaining fourteen studies, on the other hand, followed the posttest only with control (POWC) design, where students are assigned to experimental and control treatments and assessed only once, after the treatment (i.e., learning using OER or OEP).
According to the guidelines provided by Thalheimer and Cook (2002) for interpreting effect size, an effect size is negligible if $-0.15 < g < 0.15$; small if $0.15 \leq g < 0.40$; medium if $0.40 \leq g < 0.75$; large if $0.75 \leq g < 1.10$; very large if $1.10 \leq g < 1.45$; and, huge if $1.45 \leq g$. Additionally, to test if there was any heterogeneity in the variation of effect sizes within the reviewed studies, $Q$ and $I^2$ were evaluated (Konstantopoulos & Hedges, 2019). Specifically, a preplanned analysis was conducted to investigate if the field of education, the level of education, or the learning setting influenced the overall average effect size.

**Publication bias**

Three methods were used to assess publication bias: classic Rosenthal’s fail-safe N, Orwin’s fail-safe N, and the trim-and-fill method. Rosenthal’s (1979) fail-safe number aims to determine the number of studies with nonsignificant results of unpublished data needed to nullify the mean effect size. A fail-safe number larger than $5k + 10$ (where $k$ is the original number of studies included in the meta-analysis) is robust. This means that the effect size of unpublished studies is not likely to affect the average effect size of the meta-analysis. However, this method assumes that the mean effect size in the missing studies is zero (Borenstein et al., 2021). To overcome this issue, Orwin (1983) proposed a more stringent method to identify how many missing studies would bring the overall effect to a specific non-zero value. This method permits selecting a value that represents the smallest effect of substantive importance and identifying how many missing studies it would take to bring the overall effect below this value. Alternatively, the trim-and-fill method was proposed by Duval and Tweedie (2000) with the intention of identifying publication bias by means of a funnel plot wherein the studies are represented by dots. If the dots are distributed on both sides of a vertical line representing the average effect size, there is no publication bias. Conversely, if most of the dots are located at the bottom of the funnel or on one side of the vertical line, publication bias is present (Borenstein et al., 2010).

**Results**

**Description of the included sample**

Table 1 presents the included 25 studies in this present meta-analysis. Most of the studies ($n = 19$) were conducted with bachelor students. Additionally, OER and OEP were used mostly to teach psychology ($n = 6$), mathematics ($n = 5$) or also varied courses ($n = 6$). Among the 25 studies, 10 studies used small sample size (less than or equal to 250) and 15 studies used large sample size (larger than 250). Hedges’s $g$ was also calculated. A positive Hedges’s $g$ indicates that students using OER and OEP had better achievement than those who used traditionally copyrighted resources, and vice versa. Table 1 shows that 10 studies had negative Hedges’s $g$ value.

**Publication bias assessment**

Borenstein et al. (2010) stated that a symmetric funnel plot—when the dots (studies) are distributed on both sides of the vertical line (combined effect size)—implies that there is no publication bias. However, if most of the dots are situated at the bottom of the funnel or on one side of the vertical line, there is publication bias. Figure 2 shows that the dots in this study are distributed symmetrically around the vertical line. Additionally,
<table>
<thead>
<tr>
<th>#</th>
<th>Author(s) &amp; year</th>
<th>OER type</th>
<th>Region</th>
<th>Subject</th>
<th>Educational level</th>
<th>Intervention duration</th>
<th>Sample size</th>
<th>Hedges's g</th>
<th>SE</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Feldstein et al. (2012)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Business</td>
<td>Bachelor</td>
<td>1 year</td>
<td>Large</td>
<td>0.2</td>
<td>0.05</td>
<td>1394</td>
</tr>
<tr>
<td>2</td>
<td>Hilton et al. (2013)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Mathematics</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>−0.09</td>
<td>0.30</td>
<td>2043</td>
</tr>
<tr>
<td>3</td>
<td>Allen et al. (2015)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Chemistry</td>
<td>Bachelor</td>
<td>1 trimester</td>
<td>Large</td>
<td>0.02</td>
<td>0.07</td>
<td>377</td>
</tr>
<tr>
<td>4</td>
<td>Robinson (2015)</td>
<td>Interactive course</td>
<td>North America</td>
<td>Varied</td>
<td>Post-secondary</td>
<td>1 semester</td>
<td>Large</td>
<td>−0.12</td>
<td>0.03</td>
<td>3254</td>
</tr>
<tr>
<td>5</td>
<td>Hilton et al. (2016)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Varied</td>
<td>Bachelor</td>
<td>1 trimester</td>
<td>Large</td>
<td>0.12</td>
<td>0.03</td>
<td>2014</td>
</tr>
<tr>
<td>6</td>
<td>Kelly and Rutherford (2017)</td>
<td>Interactive course</td>
<td>North America</td>
<td>Mathematics</td>
<td>Secondary</td>
<td>1 month</td>
<td>Small</td>
<td>0.19</td>
<td>0.21</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>Chiorescu (2017)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Mathematics</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>0.20</td>
<td>0.14</td>
<td>159</td>
</tr>
<tr>
<td>8</td>
<td>Choi and Carpenter (2017)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Computer science</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>−0.44</td>
<td>0.12</td>
<td>175</td>
</tr>
<tr>
<td>9</td>
<td>Grewe and Davis (2017)</td>
<td>Interactive course</td>
<td>North America</td>
<td>History</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Small</td>
<td>1.10</td>
<td>0.20</td>
<td>58</td>
</tr>
<tr>
<td>10</td>
<td>Gurung (2017)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Bachelor</td>
<td>1 week and less</td>
<td>Large</td>
<td>−0.08</td>
<td>0.07</td>
<td>1228</td>
</tr>
<tr>
<td>11</td>
<td>Hendricks et al. (2017)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Physics</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>0.12</td>
<td>0.05</td>
<td>755</td>
</tr>
<tr>
<td>12</td>
<td>Westernmann Juárez and Venegas Muggli (2017)</td>
<td>Interactive course</td>
<td>Latin America</td>
<td>Mathematics</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Small</td>
<td>−0.10</td>
<td>0.04</td>
<td>31</td>
</tr>
<tr>
<td>13</td>
<td>Winitzky-Stephens and Pickavance (2017)</td>
<td>Document</td>
<td>North America</td>
<td>Varied</td>
<td>Bachelor</td>
<td>1 week and less</td>
<td>Large</td>
<td>−0.02</td>
<td>0.06</td>
<td>329</td>
</tr>
<tr>
<td>14</td>
<td>Basu Mallick et al. (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Varied</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>−0.14</td>
<td>0.02</td>
<td>4247</td>
</tr>
<tr>
<td>15</td>
<td>Clinton (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>0.23</td>
<td>0.09</td>
<td>204</td>
</tr>
<tr>
<td>16</td>
<td>Colvard et al. (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Varied</td>
<td>Bachelor</td>
<td>Not mentioned</td>
<td>Large</td>
<td>0.21</td>
<td>0.01</td>
<td>10,141</td>
</tr>
<tr>
<td>17</td>
<td>Jhangiani et al. (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Post-secondary</td>
<td>1 semester</td>
<td>Small</td>
<td>0.18</td>
<td>0.15</td>
<td>94</td>
</tr>
<tr>
<td>18</td>
<td>Lawrence and Lester (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Politics</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Large</td>
<td>−0.06</td>
<td>0.10</td>
<td>178</td>
</tr>
<tr>
<td>19</td>
<td>Medley-Rath (2018)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Sociology</td>
<td>Bachelor</td>
<td>1 week and less</td>
<td>Small</td>
<td>0.08</td>
<td>0.29</td>
<td>23</td>
</tr>
<tr>
<td>#</td>
<td>Author(s) &amp; year</td>
<td>OER type</td>
<td>Region</td>
<td>Subject</td>
<td>Educational level</td>
<td>Intervention duration</td>
<td>Sample size</td>
<td>Sample size (NE)</td>
<td>Sample size (NC)</td>
<td>Sample size (Total)</td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>20</td>
<td>Engler and Shedlosky-Shoemaker (2019)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Small</td>
<td>0.00</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>21</td>
<td>Grissett and Huffman (2019)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Bachelor</td>
<td>1 semester</td>
<td>Small</td>
<td>-0.16</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>22</td>
<td>Hardin et al. (2019)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Psychology</td>
<td>Bachelor</td>
<td>1 week and less</td>
<td>Large</td>
<td>0.22</td>
<td>507</td>
<td>1118</td>
</tr>
<tr>
<td>23</td>
<td>Shemy and Al-Habsi (2021)</td>
<td>Interactive course</td>
<td>Asia</td>
<td>Varied</td>
<td>Professional development</td>
<td>1 week and less</td>
<td>Small</td>
<td>2.29</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>24</td>
<td>Sulisworo and Basriyah (2021)</td>
<td>Video</td>
<td>Asia</td>
<td>Physics</td>
<td>Secondary</td>
<td>1 week and less</td>
<td>Small</td>
<td>0.20</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>25</td>
<td>Harvey and Bond (2022)</td>
<td>Interactive (text) book</td>
<td>North America</td>
<td>Mathematics</td>
<td>Secondary</td>
<td>Not mentioned</td>
<td>Small</td>
<td>0.27</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>
although some dots are outside the triangle of the funnel plot, most of them are in the upper part of Fig. 2 and not at the bottom. Therefore, it can be argued that the reliability of the present meta-analysis is not affected by publication bias.

**Overall effect size for learning achievement**

The meta-analysis yielded an overall effect size of $g = 0.07$, $p < 0.001$, indicating that OER and OEP had a negligible effect on students’ learning achievement (see Table 2). Specifically, Document ($g = -0.20$; 95% CI = −0.14 to 0.10; $n = 1$), Interactive (text) book ($g = 0.13$; 95% CI = 0.11 to 0.15; $n = 18$) and Interactive course ($g = -0.11$; 95% CI = −0.14 to −0.08; $n = 5$) had a negligible effect on students’ learning achievement. Video ($g = 0.20$; 95% CI = −0.0.33 to 0.73; $n = 1$) had a small effect on students’ learning achievement.

The $I^2$ statistic showed that 96.60% of variance resulted from between-study factors, implying that other variables might moderate the effect size of OER (as pointed out in the background of this study).

The forest plot presents the variation of effect size across the 25 included studies (see Fig. 3). The black square represents each study’s weighted effect size, where a larger square size implies a larger effect size. The arrow underneath each square (effect size) represents the confidence interval of the associated effect size. The overall mean

![Funnel plot of standard error by Hedges' $g$](image)

Table 2  Effect of OER on students’ learning achievement

<table>
<thead>
<tr>
<th>Analysis</th>
<th>$N$</th>
<th>$g$</th>
<th>95% CI</th>
<th>$Z$</th>
<th>$p$</th>
<th>$I^2$</th>
<th>$\tau^2$</th>
<th>Effect size interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>25</td>
<td>0.07</td>
<td>[0.06, 0.09]</td>
<td>10.69</td>
<td>0.001*</td>
<td>96.60</td>
<td>0.04</td>
<td>Negligible</td>
</tr>
<tr>
<td>Document</td>
<td>1</td>
<td>−0.02</td>
<td>[−0.14, 0.10]</td>
<td>−0.33</td>
<td>0.744</td>
<td>0</td>
<td>0</td>
<td>Negligible</td>
</tr>
<tr>
<td>Interactive (text) book</td>
<td>18</td>
<td>0.13</td>
<td>[0.11, 0.15]</td>
<td>16.48</td>
<td>0.001*</td>
<td>95.86</td>
<td>0.04</td>
<td>Negligible</td>
</tr>
<tr>
<td>Interactive materials</td>
<td>5</td>
<td>−0.11</td>
<td>[−0.14, −0.08]</td>
<td>−7.51</td>
<td>0.001*</td>
<td>94.75</td>
<td>0.06</td>
<td>Negligible</td>
</tr>
<tr>
<td>Video</td>
<td>1</td>
<td>0.20</td>
<td>[−0.33, 0.73]</td>
<td>0.74</td>
<td>0.461</td>
<td>0</td>
<td>0</td>
<td>Small</td>
</tr>
</tbody>
</table>

$N$ = number of studies; $g$ = Hedges’ $g$ effect size; CI = confidence interval; $Z$ = $Z$ value for Hedges’ $g$, $p = p$ values of Hedges’ $g$, $I^2$ and $\tau^2$ are measures of effect size variability; *$p < 0.001$
Effect size ($g = 0.073$) is presented at the last row of the forest plot. Interestingly, it is seen that almost half of the studies had a negative effect size with different confidence intervals, implying that the use of traditionally copyrighted materials had a better impact on learning achievement compared to the use of OER and OEP. This further explains the obtained negligible effect of OER and OEP on students’ learning achievement (see Table 2).

Effect sizes of learning achievement for moderator variables

Course subject

Meta-regression was used to investigate any possible variations in the effect sizes of educational subjects (Liesa-Orus et al., 2023). According to Table 3, the meta-regression result indicates that the course subject model is associated with the effect sizes of the learning achievement under OER as the p-value is 0.05 (Borenstein 2022). Moreover, the statistics of the subject indicate that using OER in history ($p = 0.001$) is likely to relate to the effect size. Specifically, the coefficient indicates that the expected mean effect size for studies using OER in history is 1.14 points higher than the expected mean effect size for studies using OER in psychology, with standard error 0.33 and a confidence interval 0.49–1.78. In other words, OER used in history is likely to have a significantly better effect on learning achievement than OER used in psychology.

Educational level

Meta-regression was used to investigate any possible variations in the effect sizes of educational level (Chaudhary & Singh, 2022). According to Table 4, the meta-regression result indicates that the educational level model is associated with the effect sizes of the learning achievement under OER as the p-value is equal to 0.001.
Moreover, the statistics of educational level indicate that using OER in professional development \((p = 0.001)\) is likely to relate to the effect size. Specifically, the coefficient indicates that the expected mean effect size for studies using OER in professional development is 2.26 points higher than the expected mean effect size for studies using OER in bachelors, with standard error 0.48 and a confidence interval 1.31–3.20. In other words, OER used in professional development is likely to have a significantly better effect on learning achievement than OER used in bachelor.

**Table 3** Meta-regression results for the learning achievement of response from subject

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>95% lower</th>
<th>95% upper</th>
<th>Z-value</th>
<th>Two-sided p-value</th>
<th>(\chi^2) = 17.15, (df = 9, p = 0.05^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept Subject(^a)</td>
<td>−0.04</td>
<td>0.11</td>
<td>−0.26</td>
<td>0.18</td>
<td>−0.33</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>1 = Business</td>
<td>0.24</td>
<td>0.26</td>
<td>−0.28</td>
<td>0.75</td>
<td>0.90</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>2 = Chemistry</td>
<td>0.06</td>
<td>0.27</td>
<td>−0.47</td>
<td>0.59</td>
<td>0.21</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>3 = Computer science</td>
<td>−0.40</td>
<td>0.29</td>
<td>−0.97</td>
<td>0.16</td>
<td>−1.40</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>4 = History</td>
<td>1.14</td>
<td>0.33</td>
<td>0.49</td>
<td>1.78</td>
<td>3.46</td>
<td>0.001**</td>
<td></td>
</tr>
<tr>
<td>5 = Mathematics</td>
<td>0.02</td>
<td>0.17</td>
<td>−0.31</td>
<td>0.36</td>
<td>0.14</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>6 = Physics</td>
<td>0.18</td>
<td>0.23</td>
<td>−0.27</td>
<td>0.63</td>
<td>0.79</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>7 = Politics</td>
<td>−0.02</td>
<td>0.28</td>
<td>−0.57</td>
<td>0.53</td>
<td>−0.08</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>8 = Sociology</td>
<td>0.12</td>
<td>0.39</td>
<td>−0.64</td>
<td>0.87</td>
<td>0.30</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>9 = Varied</td>
<td>0.16</td>
<td>0.15</td>
<td>−0.14</td>
<td>0.46</td>
<td>1.06</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>

\(^*\) 0 = Psychology as the reference group

\(^b\) The Q statistics are for the set of variables of subject

\(^*p \leq 0.05; **p \leq 0.001\)

**Table 4** Meta-regression results for the learning achievement of response from educational level

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>95% lower</th>
<th>95% upper</th>
<th>Z-value</th>
<th>Two-sided p-value</th>
<th>(\chi^2) = 22.16, (df = 3, p = 0.001**)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept Educational level(^a)</td>
<td>0.03</td>
<td>0.07</td>
<td>−0.10</td>
<td>0.17</td>
<td>0.52</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>1 = Secondary</td>
<td>−0.02</td>
<td>0.21</td>
<td>−0.43</td>
<td>0.40</td>
<td>−0.08</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>2 = Post-secondary</td>
<td>−0.02</td>
<td>0.21</td>
<td>−0.44</td>
<td>0.39</td>
<td>−0.12</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>3 = Professional development</td>
<td>2.26</td>
<td>0.48</td>
<td>1.31</td>
<td>3.20</td>
<td>4.69</td>
<td>0.001**</td>
<td></td>
</tr>
</tbody>
</table>

\(^*\) 0 = Bachelor as the reference group

\(^b\) The Q statistics are for the set of variables of educational level

\(^*p \leq 0.05; **p \leq 0.001\)
Meta-regression was used to investigate any possible variations in the effect sizes of intervention duration (Shi et al., 2023). According to Table 5, the meta-regression result indicates that the intervention duration model is not associated with the effect sizes of the learning achievement under OER as the p-value is 0.99 (Borenstein 2022).

Table 5 Meta-regression results for the learning achievement of response from intervention duration

| Model                       | Coefficient | Standard Error | 95% lower  | 95% upper | Z-value | Two-sided p-value | $Q^2 = 0.61, df = 5, p = 0.99$ |
|-----------------------------|-------------|----------------|------------|-----------|---------|--------------------|
| Intercept                   | 0.05        | 0.11           | -0.16      | 0.26      | 0.45    | 0.65               |
| Intervention duration$^a$   |             |                |            |           |         |                    |
| 1 = One week and less       | 0.12        | 0.20           | -0.28      | 0.51      | 0.58    | 0.56               |
| 2 = One month               | 0.14        | 0.43           | -0.70      | 0.99      | 0.33    | 0.74               |
| 3 = One trimester           | 0.02        | 0.28           | -0.52      | 0.57      | 0.08    | 0.94               |
| 4 = One year                | 0.15        | 0.38           | -0.59      | 0.89      | 0.40    | 0.69               |
| 5 = Not mentioned           | -0.05       | 0.29           | -0.62      | 0.52      | -0.17   | 0.86               |

$^a$ 0 = One semester as the reference group
$^b$ The Q statistics are for the set of variables of intervention duration

Meta-regression was used to investigate any possible variations in the effect sizes of the region (Liesa-Orus et al., 2023). According to Table 7, the meta-regression result indicates that the region model is associated with the effect sizes of learning achievement under OER as p-value is 0.01 (Borenstein 2022). Moreover, the region statistics indicate that using OER in Asia (p = 0.001) is likely to relate to the effect size. Specifically, the

| Table 6 Meta-regression results for the learning achievement of response from sample size
| Model                       | Coefficient | Standard Error | 95% lower | 95% upper | Z-value | Two-sided p-value | $Q^2 = 3.09, df = 1, p = 0.08$ |
|-----------------------------|-------------|----------------|-----------|-----------|---------|--------------------|
| Intercept                   | 0.01        | 0.09           | -0.19     | 0.17      | -0.13   | 0.90               |
| Sample size$^a$             |             |                |           |           |         |                    |
| 1 = Small                   | 0.28        | 0.16           | -0.03     | 0.58      | 1.76    | 0.08               |

$^a$ 0 = Large as the reference group
$^b$ The Q statistics are for the set of variables of sample size

**Intervention duration**

Meta-regression was used to investigate any possible variations in the effect sizes of intervention duration (Shi et al., 2023). According to Table 5, the meta-regression result indicates that the intervention duration model is not associated with the effect sizes of the learning achievement under OER as the p-value is 0.99 (Borenstein 2022).

**Sample size**

Meta-regression was used to investigate any possible variations in the effect sizes of sample size (Cheung & Slavin, 2016). According to Table 6, the meta-regression result indicates that the sample size model is not associated with the effect sizes of the learning achievement under OER as p-value is 0.08 (Borenstein 2022).

**Geographical distribution**

Meta-regression was used to investigate any possible variations in the effect sizes of the region (Liesa-Orus et al., 2023). According to Table 7, the meta-regression result indicates that the region model is associated with the effect sizes of learning achievement under OER as p-value is 0.01 (Borenstein 2022). Moreover, the region statistics indicate that using OER in Asia (p = 0.001) is likely to relate to the effect size. Specifically, the
coefficient indicates that the expected mean effect size for studies using OER in Asia is 1.01 points higher than the expected mean effect size for studies using OER in North America, with standard error 0.33 and a confidence interval 0.36–1.65. In other words, OER used in Asia is likely to have a significantly better effect on learning achievement than OER used in North America.

**Research design**

Meta-regression was used to explore any possible associations in the effect sizes of the research design (Geissbühler et al., 2021). According to Table 8, the meta-regression result indicates that the research design model is not associated with the effect sizes of learning achievement under OER as \( p \)-value is 0.77 (Borenstein 2022).

Finally, to further investigate for possible covariance between confounding variables, a meta-regression that includes all of the individual confounding variables that yielded statistically significant results, namely subject, educational level and region, was conducted. Table 9 reveals that subject \( (p = 0.01) \) and educational level \( (p = 0.001) \) yielded a significant covariance between confounding variables.

**Discussions**

This meta-analysis aimed to comprehensively assess the effectiveness of Open Educational Resources (OER) and Open Educational Practices (OEP) in relation to learning achievement. The analysis of 25 independent studies revealed that the impact of OER and OEP on learning achievement is generally negligible. These quantitative findings support the conclusions drawn from qualitative (Hilton, 2016, 2020) and quantitative (Clinton & Khan, 2019) reviews that compare learning achievement between courses.
using open and commercial textbooks. Additionally, it is found that course subject, educational level and the region of students might moderate the effects of OER and OEP. The obtained findings of this study can be discussed and explained from the following perspectives.

**Improvement in access does not imply improvement in learning achievement: a holistic design is needed**

The use of OER and OEP is often considered an effective learning intervention due to its potential to provide equal access to educational resources for all students (Grimaldi et al., 2019). However, the results of this meta-analysis do not substantiate this hypothesis. Dotson and Foley (2017) also argue that the change of a curriculum content license from proprietary to open does not always lead to a change in students' learning achievement. In other words, we cannot expect an improvement in learning achievement by simply changing the license of a given educational resource from proprietary to open. It requires a more comprehensive approach that involves changing not only the license,

<table>
<thead>
<tr>
<th>Table 9</th>
<th>Meta-regression results for the learning achievement of response from subject, educational level and region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.04</td>
</tr>
<tr>
<td>Subject</td>
<td></td>
</tr>
<tr>
<td>1 = Business</td>
<td>0.24</td>
</tr>
<tr>
<td>2 = Chemistry</td>
<td>0.06</td>
</tr>
<tr>
<td>3 = Computer science</td>
<td>−0.40</td>
</tr>
<tr>
<td>4 = History</td>
<td>1.14</td>
</tr>
<tr>
<td>5 = Mathematics</td>
<td>0.16</td>
</tr>
<tr>
<td>6 = Physics</td>
<td>0.16</td>
</tr>
<tr>
<td>7 = Politics</td>
<td>−0.02</td>
</tr>
<tr>
<td>8 = Sociology</td>
<td>0.12</td>
</tr>
<tr>
<td>9 = Varied</td>
<td>0.05</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
</tr>
<tr>
<td>1 = Secondary</td>
<td>−0.17</td>
</tr>
<tr>
<td>2 = Post-secondary</td>
<td>0.00</td>
</tr>
<tr>
<td>3 = Professional development</td>
<td>2.03</td>
</tr>
<tr>
<td>Region</td>
<td></td>
</tr>
<tr>
<td>1 = Asia</td>
<td>0.25</td>
</tr>
<tr>
<td>2 = Latin America</td>
<td>Region</td>
</tr>
</tbody>
</table>

*The Q statistics are for the set of variables of subject
*p ≤ 0.05; **p ≤ 0.001
but also the used instructional approach, the way the educational resources are designed, etc. Based on the review of 25 included studies, it is found that ensuring an improvement in learning achievement is beyond the simple access to educational resources, and several elements should be considered, some of which are considered and discussed below, namely: OER quality, instructional, and learners’ individual factors.

**OER quality**

The quality of OER and effective implementation of OEP are crucial factors that significantly influence learning achievement. High-quality OER, characterized by accurate and up-to-date content, clear learning objectives, and appropriate instructional design, have been shown to positively impact student learning outcomes (Butcher, 2015). Learners who have access to well-designed OER that align with the curriculum and provide meaningful learning experiences are more likely to engage with the materials and effectively acquire knowledge and skills. However, it is important to acknowledge that not all OER resources meet the necessary standards of accuracy, coherence, and pedagogical effectiveness. Research has indicated significant variability in the quality of OER, resulting in inconsistent learning experiences and potentially limiting their impact on learning achievement (Weller, 2017). To address this, quality assurance processes, peer review, and evaluation mechanisms are essential to ensure that the content and resources meet established standards. While there was early skepticism and critique in regards to OER quality based on design and economic production models (see Kahle, 2008; Weller, 2010), there is nothing inherently different in regards to open and closed/proprietary content beyond the intellectual property rights. In other words, the quality criteria that apply to proprietary/closed resources also apply to OER and we should not expect quality differences in OER produced under the same production modes (e.g. by experienced publishers and designers). The results of this study indicate that no significant difference was found, which substantiates this assertion.

**Instructional approach**

Scoring improvements in learning achievement also depends on how students engage with OER and the effective implementation of OEP accordingly. On the one hand, despite the easy access to learning materials provided by OER, learners may not use them at all (Feldstein et al., 2012) or may not have sufficient time to engage with them (Westermann Juárez and Venegas Muggli, 2017). On the other hand, while OER offer the advantage of making learning more individualized, students may encounter a broader range of perspectives through OER but the content they learn may not align with objective measures of learning (Gurung, 2017). In the same vein, Zulaiha and Triana (2023) stated that a proper teaching method and learning strategy must accompany the OER to be effectively leveraged to improve students’ skills in order for OER to make a significant impact on the student learning. An older study by Slavin and Lake (2008) similarly found that the selection of instructional approach has a larger impact on learning achievement than the choice of curriculum content. Besides, numerous educators face challenges such as time constraints, insufficient skills and competences (e.g., digital), lack of understanding about what OER or OEP actually mean, and a lack of incentives to
engage in open practices. Consequently, the widespread adoption of OEP remains limited, potentially hindering its impact on learning achievement (Tlili et al., 2021; Zhang et al., 2020b).

**Learners' individual factors**

The impact of OER and OEP on learning achievement is influenced by individual learner characteristics, including prior knowledge and motivation. Tlili and Burgos (2022) emphasized the importance of providing personalized learning as students in open education might have different backgrounds and competencies. It is crucial to acknowledge that students may show diverse responses to open educational initiatives, and some may require extra support or guidance to fully reap the benefits of these resources. For example, studies have shown that students from lower socioeconomic backgrounds, who may lack essential skills in effectively utilizing OER, tend to attain lower learning outcomes compared to their peers (Robinson, 2015). Thus, recognizing and addressing the diverse needs of learners is important in enhancing the impact of OER and OEP on learning achievement.

**Adequate experimental design is crucial for accurately measuring learning achievement**

Beyond the OER and OEP selection and implementation, the study indicates that the applied experimental design might hinder the accurate measurements of OER and OEP effects on learning achievement. Based on the 25 reviewed studies, it is found that most of the studies used quasi-experiments given that random assignments are not always possible in open education. As a result, this might hinder measuring the measurable effects of OER and OEP on learning achievement (Griggs & Jackson, 2017; Gurung, 2017).

Additionally, separating the effects of OER and OEP from other effects is also a challenge in the conducted experiments. Wiley (2022) described several ways in which research that purports to show the impact of OER adoption on student learning actually shows the impact of other interventions that are associated with OER adoption (e.g., when faculty receive support from an instructional designer to redesign a course after adopting OER). Pawlyshyn et al. (2013) also reported an improvement in learning achievement when OER was adopted simultaneously with flipped classrooms. However, it is not clear whether this improvement was due to the use of OER or flipped classrooms. OER are often employed alongside other interventions which can make isolating their effect methodologically problematic. This challenge of correlating improvements in learning achievement with the use of OER and OEP was also reported by other researchers (Griggs & Jackson, 2017; Gurung, 2017).

Most of the reviewed studies used final exam scores or GPA (grade point average) to measure the learning achievement of incorporating OER and OEP. However, this method is questionable as the designed exam may vary depending on the taught course subject and requirements, leading to a variation in the measured learning achievement. It is, therefore, recommended to use standardized instruments when measuring learning achievement using OER and OEP (Hendricks et al., 2017; Hilton, 2020). This normalization might lead to competence validation or even credit recognition through alternative
credentials (e.g., Alternative Digital Credentials -ADC-), which is one of the open challenges around open education (Griffiths et al., 2022).

Confounding variables might lead to a variation of learning achievement

The present meta-analysis revealed that several confounding variables could affect the learning achievement of students when using OER and OEP. One of these variables is the course subject. This might be explained by the fact that some subjects have quality OER published online while others do not. For example, Lawrence and Lester (2018) highlighted a specific concern in the open content space for subjects like political science, which is the lack of available textbook options. Similarly, Choi and Carpenter (2017) found it challenging to find a suitable OER for their interdisciplinary Human Factors and Ergonomics course. The researchers discovered that OER for the Human Factors and Ergonomics course often provided in-depth content for individual topics, including extra information that is relevant to their subject of focus but not directly related to the course learning objectives. Furthermore, the limited number of OER options creates difficulties for instructors in applying some of their preferred pedagogical approaches.

The obtained findings also revealed that students’ geographic region can moderate the effect of OER and OEP on learning achievement. This could be explained with the fact that several regions, such as East Asia, have made remarkable progress in terms of raising awareness and adopting OER and OEP (Tlili et al., 2019), while others like the Arab region and sub-Saharan Africa are still behind (Tlili et al., 2020, 2022). This might result in divided regions in terms of students’ perception and acquired competencies to use OER and OEP, hence having varied learning achievement across regions.

Conclusions, implications and limitations

This study included a meta-analysis and research synthesis to investigate the effects of OER and OEP on students’ learning achievement. This analysis describes how this effect is moderated across different variables (i.e., course subject, level of education, intervention duration, sample size and geographical distribution). As discussed above, to the best of our knowledge, no previous study has conducted a similar analysis. Based on the findings, it can be argued that holistic OER learning design may be needed to optimize learning outcomes; that researchers should employ adequate experimental design when investigating the relationship between OER and learning achievement; and highlighted the potential role of confounding factors that can lead to a variation of learning achievement when using OER.

Implications

This study supports previous research in identifying no significant differences between the interventions using open and closed approaches (content or practice). This meta-analysis supports this conclusion following along existing literature on media/intermedia comparison studies (e.g. Clark, 1994; Salomon & Clark, 1977).

The conundrum for comparisons studies such as those included in this meta-analysis is thus: if a true experiment made to evaluate the influence of OER in learning achievement were to be designed, the only variable would be the OER itself, in other words, the intellectual property license of the content (considering this to be the defining
characteristic of OER). If this were possible, one could only expect that the affordances of open licensing would possibly point to the effects of reduced cost or ease of access to relate to achievement (e.g., Fischer et al., 2015). But if this is done, it would offer us minimal new insights beyond what we already could expect, in principle. It stands to reason that not having access to resources designed to be part of a course would reduce achievement (a comparison on whether students actually did or did not access and make use of resources in the treatment and control condition is another study entirely).

However, the truly intriguing and critical questions pertain to practical applications. If we do allow practice to vary, for example: if the OER afforded some different sort of practice (as OEP is defined in OER-enabled pedagogy) then we are really measuring something more holistic—the practice which includes the resource. As Salomon and Clark (1977, p. 102) conclude: “In short, when only the least significant aspects of instruction are allowed to vary, nothing of interest could, and did, result.”

This study then might point us to valuable avenues for further research. Perhaps course instructors and designers are attempting to faithfully replicate courses that make use of OER simply to test possible outcomes in achievement; here, clearly, we should expect no difference to emerge. Furthermore, instructors may not be really leveraging OER-enabled pedagogy or more expansive perspectives of OEP.

Additionally, this meta-analysis might help disencourage further comparison studies based on OER and achievement. It points us to the urgency of expanding the object of analysis beyond intrinsic characteristics of OER and focus on how principles of openness might significantly alter the nature of the practices and courses themselves, might lead to outcomes which are not measured simply by achievement gains, and additionally, might or might not cater to different types of students.

This present study can contribute to the literature from different perspectives. From a theoretical perspective, this study adds to the ongoing debate for the past twenty years about the effectiveness of OER and OEP by revealing what might moderate the effectiveness of OER and OEP. From a practical perspective, this study can contribute to Sustainable Development Goals (SDGs) (UN, 2021), specifically SDG 4 quality education, by highlighting the different variables (i.e., quality, the used pedagogical approach, etc.) that different stakeholders (i.e., educators, instructional designers, etc.) should consider when adopting OER and OEP for better learning achievement. Finally, from a methodological perspective, this study contributes to the literature by pointing out various experimental criteria (i.e., standardized measurements, design, etc.) that should be considered when designing research experiments to effectively measure the true effect of OER, hence providing more accurate results that could advance the field in this regard.

Limitations and future directions
It should be noted that the statistical power was not examined in this present meta-analysis, which is the case in the majority of published meta-analyses in the literature (Burçin, 2022; Dumas-Mallet et al., 2017; Thorlund & Mills, 2012). A significant barrier to the widespread implementation of statistical power in meta-analysis is the difficulty of understanding how it can be computed due to the various variables that should be considered in each study, as well as the heterogeneity of the conducted studies (Cafri et al., 2010; Ioannidis et al., 2014; Vankov et al., 2014). Additionally, there is a lack of an accessible
and easy-to-use software or R script that can help to compute statistical power (Griffin, 2021; Thomas & Krebs, 1997). In this context, various software, such as G*power, have been developed to calculate statistical power for primary research, allowing for widespread implementation of power analysis in primary research (Faul et al., 2007). However, despite the similarity in procedure, such analogous software options do not exist for meta-analysis. Consequently, to calculate statistical power for a given meta-analysis, researchers must manually perform the calculations, use an online calculator, or utilize a user defined script (e.g., Cafri et al., 2009). These methods can be limited in functionality and difficult to integrate into a reproducible workflow (Griffin, 2021).

Besides, despite the reliability of the obtained results having been validated through the bias assessment, this study has some other limitations that should be acknowledged. For instance, the obtained results might be limited to the used keywords and electronic databases. Additionally, the obtained analysis was only based on courses conducted in English; non-English course studies might reveal different results. Moreover, while the present meta-regression yielded valuable insights about the effect of OER and OEP on learning achievement as well as the moderating variables of this effect, the limited sample size of the included studies might impact the generalizability of the findings. Therefore, future researchers are encouraged to complement this work by covering more databases and analyzing non-English courses, hence providing a more comprehensive view of OER and OEP effects. Additionally, this present meta-analysis did not consider teacher variables (e.g., same teacher or not when teaching using OER and non-OER materials) which could moderate the effects of OER and OEP on learning achievement (Hilton, 2020). Therefore, future studies could focus on this line of research. Finally, this present meta-analysis did not consider OER quality, which has been shown to have a significant impact on students’ learning outcomes (Butcher, 2015). Future research could systematically assess and incorporate OER quality as a moderating variable, hence further enhancing the understanding of the intricate relationship between OER and learning achievement. However, despite these limitations, this present study provided quantitative evidence about the OER and OEP effects on students’ learning achievement.

Acknowledgements
Not applicable.

Author contributions
Each author contributed evenly to this manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The datasets generated and/or analyzed during the current study are presented within this study.

Declarations

Competing interests
The authors have no conflict of interest to declare.

Received: 19 July 2023   Accepted: 26 September 2023
Published online: 13 October 2023
References

References with an asterisk (*) indicate studies included in the analysis


**Publisher’s Note**

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