Research Article

Does differentiated instruction affect learning outcome? Systematic review and meta-analysis

Muh. Asriadi AM1, Samsul Hadi2, Edi Istiyono3 and Heri Retnawati4

1Yogyakarta State University & Universitas Pendidikan Indonesia, Indonesia (ORCID: 0000-0003-4370-3550)
2Yogyakarta State University, Indonesia (ORCID: 0000-0003-3437-2542)
3Yogyakarta State University, Indonesia (ORCID: 0000-0001-6034-142X)
4Yogyakarta State University, Indonesia (ORCID: 0000-0002-1792-5873)

In an era of increasingly diverse education, the need to understand the effectiveness of differential instruction on learning outcomes is critical in creating inclusive and student-centered learning experiences. This study aims to comprehensively review the effectiveness of differentiated instruction on student learning outcomes and investigate what moderator variables contribute significantly to the effect of implementing differentiated instruction on student learning outcomes. The study was a systematic review and meta-analysis that utilized 49 primary studies, from which 63 effect sizes were extracted. It was found that there was a significant (gRE = 1.109, p < .01) of the application of differential instruction on student learning outcomes compared to learning in the control group. When measured from the aspect of the measured ability, subjects, grade level, sample size, and index type, there was no significant difference in effect size. However, there was a significant difference in effect size when viewed from the country's perspective. It was concluded that differential instruction is an appropriate learner strategy used to improve the quality of learning at all levels. However, it is prudent to acknowledge that the effectiveness of implementing differentiated instruction is intrinsically intertwined with the specific context and educational milieu of a given country. This study provides a nuanced understanding that while differentiated instruction holds universal promise, its optimal outcomes are sensitive to contextual influences.

Keywords: Differentiated instruction; Systematic review; Meta-analysis

Article History: Submitted 9 May 2023; Revised 26 August 2023; Published online 24 September 2023

1. Introduction

Education plays a vital role in the development of individuals and societies. In order to ensure that all students have a successful learning experience, it is essential to recognize and implement effective teaching methods within the educational setting (Candra & Retnawati, 2020). A one-size-fits-all strategy only sometimes works to fulfill each student's individual learning needs and preferences (Chung, 2023). Finding efficient teaching methods that respect students' differences and learning preferences is crucial for meeting the various learning needs of each student.

One approach that has become the focus of research and debate in education is Differentiated Instruction [DI]. DI is a learning approach designed to meet the individual needs of each student in the class (Tomlinson, 2017). This approach is based on the understanding that students have
different learning styles, abilities, and interests, which need to be recognized and accommodated in the classroom (Lavrijsen et al., 2021). This strategy acknowledges that every student has a unique learning style, level of expertise, and learning rate (Puzio et al., 2020). DI involves adjusting and adjusting content, processes, and learning assessments to meet the individual needs of students (Brigandi et al., 2019; Kohnke, 2022). It is an approach to teaching that recognizes the diverse needs of students and responds to them through various strategies, tactics, and instructional materials (Huang, 2022). The purpose of differentiated instruction is to give all students with equal opportunity to learn and succeed by tailoring instruction to each student's specific requirements (Paskevicius, 2021). This teaching approach is based on the belief that all students have unique learning styles and abilities and that there needs to be more practical than a one-size-fits-all approach to instruction.

In recent years, several studies have been conducted to explore the impact of DI on student learning (Peters et al., 2022; Zierwald et al., 2022). However, the results of these studies have been mixed and inconclusive, leading some researchers to question the validity of disparate teaching as an effective teaching strategy. Moreover, although DI has become a topic of concern in education, there is still debate about its influence in enhancing student learning. Several studies support the effectiveness of DI in improving student learning outcomes (Dalila et al., 2022), while other studies show inconsistent or even contradictory results (Peters et al., 2022). However, a significant concern arises elucidating the significance of this study. Filling the research gap not only adds depth to the existing body of knowledge but also contributes to refining pedagogical practices and advancing the field of education. By delving into the effectiveness of differentiated instruction and discerning the key moderator variables, this study is poised to shed light on a pivotal aspect of instructional methodologies. Such insights can guide educators, researchers, and policymakers in making informed decisions, ultimately elevating the quality of education provided to students.

In order to address this ongoing discussion and attain a more comprehensive comprehension, conducting thorough research using systematic reviews and meta-analyses is imperative. A systematic review is a systematic approach to identifying, evaluating, and synthesizing relevant literature on a research topic (Linnenluecke et al., 2020). The meta-analysis then uses statistics to combine the results of existing research and provide a more comprehensive summary of the effect of DI on learning (Kahmann et al., 2022). By conducting a systematic review and meta-analysis of the effect of DI on learning, it will be possible to gather and synthesize evidence from previous research and produce a clearer understanding of the effectiveness of DI.

The novelty of this research is the potential to provide more robust guidance for designing inclusive learning approaches. By better understanding the effect of differential instruction and the factors that influence it, this research can provide educators with a better framework for delivering learning that meets the diverse needs of students. Moreover, this research enriches the educational discourse by adding new insights about the effectiveness of differential instruction. This encourages continuous development in teaching methods and helps address the challenges faced by educators in creating responsive and adaptive learning environments.

However, the importance of this study goes beyond addressing a research gap. It is crucial to underline how filling this gap holds significance within the broader education research community. Through a nuanced understanding of how differentiated instruction influences student learning, this study offers novel perspectives that can potentially reshape teaching practices, enrich pedagogical discourse, and ultimately elevate the quality of education. The results of this research can also provide valuable insights for educators, researchers, and policymakers in designing more inclusive and effective learning approaches for all students. Based on the description above, this study aims to comprehensively review the effectiveness of differentiated instruction on student learning outcomes and investigate what moderator variables contribute significantly to the effect of implementing differentiated instruction on student learning outcomes.
2. Method

2.1. Research Method

This study employs a quantitative research methodology, focusing on collecting and analyzing numerical data to distinguish it from other research approaches. This study employs systematic review and meta-analysis methodologies to investigate the impact of differentiated instruction on student learning outcomes. These techniques involve the synthesis and statistical analysis of diverse research findings to assess the effectiveness of this instructional approach. The initial phase of the study involves conducting a systematic review, wherein relevant studies about differentiated instruction are gathered from reputable databases. Following this, the selection process focuses on identifying studies with a quasi-experimental design comprising two independent groups.

The second phase entails performing a Meta-analysis to thoroughly assess the quantitative data gathered (sample size, mean score, and standard deviation) from independent studies on a specific subject. Effect size is utilized as a standardized measure to evaluate the outcomes of the independent studies within the meta-analysis (Seidler et al., 2019). It allows for consistent and standardized criteria to assess the findings of each study. The data from the studies are coded according to specific criteria, involving extracting relevant data from each study for further analysis. Statistical analysis is conducted to interpret the research findings within the context of the meta-analysis (Chen, 2017).

2.2. Sources of Data and Search Methodologies

We comprehensively searched relevant studies published between January 2010 and April 2023 using various public databases, websites, search engines (PoP), journals, and specific search terms. The databases we searched included Scopus, ScienceDirect, and ERIC. Additionally, we utilized Google Scholar as a search engine. The search terms used for empirical studies included Differentiated Instruction, Differentiated Learning, Differentiated Strategy, Differentiated Reading, and Differentiated Practices.

2.3. Criteria for Inclusion and Exclusion

Inclusion criteria for studies in the meta-analysis were conducted between January 2010 and April 2023, focusing on Differentiated Instruction (DI), measuring learning outcomes, and using experimental research designs to test competition or learning outcomes comparisons. Conversely, studies were excluded from the meta-analysis if they were written in English or needed more data for effect size calculations.

2.4. Selection of a Study

After conducting an extensive search, 1,774 articles matching the keywords were identified. After applying the specified criteria, 94 articles remained. After assessing the abstracts and data availability, 49 articles were selected for further examination and analysis, as depicted in Figure 1.

The process of systematic review, six moderator variables are investigated and operationalized as follows:

- Measured Ability: The research encompasses various cognitive abilities, including academic achievement, critical thinking skills, letter word identification, reading comprehension, writing competency, picture vocabulary, student retention, decoding skills, fluency, interactive reading, and mathematical problem-solving. Additionally, it explores psychomotor abilities such as strength in mentally disabled pupils, long jump effectiveness, teacher’s differentiation activities, learning strategies, and flipped instruction. Furthermore, affective abilities such as motivation, attitude toward science, autonomy, and self-actualization are also considered in the study.
Subjects: The study includes five subject categories, which are mathematics, English, physical education, natural science (biology, chemistry, physics), and childhood education.

Grade Level: The study encompasses five educational levels, which are Students in kindergarten, elementary school, junior high school, senior high school, and college.

Country: Due to the diversity of countries included in each study, this research categorizes the studies into four continents: Asia (including Bahrain, Egypt, Jordan, Iran, Indonesia, Iraq, Lebanon, Philippines, Saudi Arabia, Singapore, and Turkey), Africa (including Cameroon, Ethiopia, Kenya, and Nigeria), America (specifically the USA), and Europe (including Central Europe, Germany, Greece, the Netherlands, and Romania).

Sample Size: The study includes small sample sizes (N<30) as well as big sample sizes (N\geq30).

Index Type: The study includes articles published in Scopus-indexed and non-Scopus-indexed journals.

After applying the necessary filters, 49 suitable primary studies were identified and expanded to 63 studies for further analysis in the meta-analysis stage.

2.5. Data Analysis

The meta-analysis is carried out at the study level, with the effect sizes from numerous research gathered. Following the acquisition of the mean as well as the standard deviation for each study, the means are standardized on a standard scale, yielding the effect sizes for each study. After computing the effect sizes, the meta-analysis begins, which includes calculating the total effect size, also known as the summary effect. For combining effect sizes, researchers have two models: the model with fixed effects and the random-effect model (Lipsey & Wilson, 2001). Following the recommendation of Pigott and Polanin (2020), we used the random effects model in our meta-analysis because it considers the assumption that actual effects may vary across studies. Additionally, in these selected studies, the standardized mean effect size was used in comparing...
independent group means that were deemed equivalent for each of the two variables in each research.

We also performed a moderator analysis using robust variance estimates, treating the moderator as a predictor. The heterogeneity test assesses the variability among effect sizes and determines the statistical decision-making criteria. The Q-test (heterogeneity test) was used to assess variation between the results of different studies within the study. The significant Q statistic indicates that the individual studies did not draw from the same whole population (Suurmond et al., 2017). As a result, descriptive statistical analysis is used in this study to calculate the effect size and summary effect. Given the relatively small number of studies (63), we chose a critical alpha value of 0.01. Furthermore, the Funnel plot approach and fail safe-N were employed with criterion (N > 5K + 10) to detect publication bias in this analysis, where K = the number of studies (Ahn & Kang, 2018). In particular, we use the R Studio 023.3.1.446 version (Posit Team, 2023) with the R library (meta) (Balduzzi et al., 2019) and Library (metafor) (Viechtbauer, 2010) packages. It seeks to calculate the cumulative mean effect size of each study’s effect size, and the confidence interval.

3. Results

3.1. Overview of Primary Studies

The researcher got 63 independent samples from 49 primary studies (scientific articles) based on the inclusion criteria established by the researcher. Table 1 summarizes the studies that became the object of meta-analysis consisting of study name and year of publication, effect size (g), standard error [SE], measured ability, subjects, grade level, country, sample size, and index type.

According to the data in Table 1, it is clear that the size of the effect varies among research, as evidenced by both negative and positive values. It indicates the presence of variation in the results obtained from each study. Analyzing the distribution of studies based on measured abilities, it is evident that most studies focus on measuring cognitive ability (63.49%), while the fewest studies examine affective ability (14.29%). Furthermore, when considering the distribution of studies based on subjects, English subjects represent the highest proportion (50.79%), whereas childhood education subjects comprise the lowest percentage (3.18%). Regarding grade level, most studies were conducted among high school seniors (31.75%), with the least number of studies conducted at the Kindergarten level (6.39%). Regarding geographical distribution, most research was carried out in Asia (36.51%). The percentage of studies utilizing large sample sizes (82.54%) is more significant than those utilizing small (17.46%). Additionally, there is a relatively balanced representation of studies indexed in Scopus and non-Scopus journals.

3.2. Main Analysis Results

The random effects model analysis found that the average impact size of the 63 studies was 1.11 ($p < .001$), with a 95% confidence range ranging from 0.74 to 1.47. These findings imply a significant impact of differentiated instruction on learning outcomes compared to traditional learning methods. Notably, effect sizes greater than 0.8 falls into the enormous effect sizes category, as defined by Cohen (1998). Therefore, the positive effect of implementing differentiated instruction on learning outcomes, compared to traditional learning, falls within the high effect size category. The analysis also indicated heterogeneous effect sizes among the 70 studies ($Q = 1601.16$, $df = 62$, $p < .001$). These results indicate a substantial variation in effect sizes used in this study, highlighting the need for further analysis of moderator variables.

3.3. Moderator Analysis

After investigating the application of varied instruction to learning outcomes in many aspects, the results of the moderator variable analysis, including measured ability, subjects, grade level, country, sample size, and index type, are presented in Figure 2.
# Studies included and their properties

<table>
<thead>
<tr>
<th>No</th>
<th>Study</th>
<th>g</th>
<th>SE</th>
<th>Measured Ability</th>
<th>Subjects</th>
<th>Grade Level</th>
<th>Country</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Odicta (2017) [Study 1]</td>
<td>1.94</td>
<td>0.33</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>2</td>
<td>Odicta (2017) [Study 2]</td>
<td>1.16</td>
<td>0.3</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>3</td>
<td>Aliakbari and Haghighi (2014)</td>
<td>1.99</td>
<td>0.36</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>4</td>
<td>Marinescu et al. (2014)</td>
<td>0.81</td>
<td>0.76</td>
<td>Psychomotor</td>
<td>Physical Education</td>
<td>Elementary School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>5</td>
<td>Blessing et al. (2021)</td>
<td>1.56</td>
<td>0.15</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>6</td>
<td>Eysink et al. (2017) [Study 1]</td>
<td>0.85</td>
<td>0.56</td>
<td>Psychomotor</td>
<td>Natural Science</td>
<td>Elementary School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>7</td>
<td>Eysink et al. (2017) [Study 2]</td>
<td>0.16</td>
<td>0.53</td>
<td>Affective</td>
<td>Natural Science</td>
<td>Elementary School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>8</td>
<td>Eysink et al. (2017) [Study 3]</td>
<td>-0.50</td>
<td>0.14</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Elementary School</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>9</td>
<td>Salar and Turgut (2021)</td>
<td>-0.20</td>
<td>0.16</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Senior High School</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>10</td>
<td>Al Otaiba et al. (2015) [Study 1]</td>
<td>0.67</td>
<td>0.13</td>
<td>Psychomotor</td>
<td>English</td>
<td>Kindergarten</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>11</td>
<td>Al Otaiba et al. (2015) [Study 2]</td>
<td>0.29</td>
<td>0.12</td>
<td>Psychomotor</td>
<td>English</td>
<td>Kindergarten</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>12</td>
<td>Melka and Jatta (2022)</td>
<td>0.45</td>
<td>0.22</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>13</td>
<td>Ogunkunle and Henrietta (2014)</td>
<td>2.25</td>
<td>0.26</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>14</td>
<td>Endal et al. (2013)</td>
<td>0.33</td>
<td>0.18</td>
<td>Psychomotor</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>15</td>
<td>Hussein (2021)</td>
<td>2.8</td>
<td>0.46</td>
<td>Psychomotor</td>
<td>Physical Education</td>
<td>College Students</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>16</td>
<td>Awofala and Lawani (2020)</td>
<td>0.01</td>
<td>0.14</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>17</td>
<td>Dalila et al. (2022)</td>
<td>1.61</td>
<td>0.28</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>18</td>
<td>Saleh (2021)</td>
<td>-1.0</td>
<td>0.36</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>19</td>
<td>Kotob and Abadi (2019)</td>
<td>1.96</td>
<td>0.56</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>20</td>
<td>Al-Shehri (2020) [Study 1]</td>
<td>1.39</td>
<td>0.32</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>21</td>
<td>Al-Shehri (2020) [Study 2]</td>
<td>0.50</td>
<td>0.29</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>22</td>
<td>Njagi (2015)</td>
<td>1.86</td>
<td>0.12</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>23</td>
<td>Smadi and Al Masri (2018) [Study 1]</td>
<td>1.72</td>
<td>0.2</td>
<td>Psychomotor</td>
<td>English</td>
<td>Junior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>24</td>
<td>Smadi and Al Masri (2018) [Study 2]</td>
<td>6.84</td>
<td>0.45</td>
<td>Affective</td>
<td>English</td>
<td>Junior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>25</td>
<td>Haelermans (2022) [Study 1]</td>
<td>0.18</td>
<td>0.08</td>
<td>Affective</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>26</td>
<td>Haelermans (2022) [Study 2]</td>
<td>0.11</td>
<td>0.08</td>
<td>Psychomotor</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>27</td>
<td>Grain et al. (2022)</td>
<td>2.08</td>
<td>0.12</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>28</td>
<td>Wong et al. (2023) [Study 1]</td>
<td>1.17</td>
<td>0.34</td>
<td>Affective</td>
<td>English</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>29</td>
<td>Wong et al. (2023) [Study 2]</td>
<td>1.17</td>
<td>0.34</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>30</td>
<td>Magableh and Abdullah (2019)</td>
<td>0.53</td>
<td>0.28</td>
<td>Cognitive</td>
<td>English</td>
<td>Junior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>31</td>
<td>Muthomi and Mbogua (2014)</td>
<td>1.81</td>
<td>0.12</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>No</td>
<td>Study</td>
<td>g</td>
<td>SE</td>
<td>Measured Ability</td>
<td>Subjects</td>
<td>Grade Level</td>
<td>Country</td>
<td>Sample Size</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>32</td>
<td>He et al. (2019)</td>
<td>-0.20</td>
<td>0.08</td>
<td>Psychomotor</td>
<td>Natural Science</td>
<td>College Students</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>33</td>
<td>Azah (2016) [Study 1]</td>
<td>0.76</td>
<td>0.49</td>
<td>Psychomotor</td>
<td>English</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Small</td>
</tr>
<tr>
<td>34</td>
<td>Azah (2016) [Study 2]</td>
<td>1.08</td>
<td>0.51</td>
<td>Psychomotor</td>
<td>English</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Small</td>
</tr>
<tr>
<td>35</td>
<td>Gomaa (2014) [Study 1]</td>
<td>5.20</td>
<td>0.55</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Junior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>36</td>
<td>Gomaa (2014) [Study 2]</td>
<td>2.24</td>
<td>0.33</td>
<td>Affective</td>
<td>Natural Science</td>
<td>Junior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>37</td>
<td>Little et al. (2014)</td>
<td>0.26</td>
<td>0.05</td>
<td>Cognitive</td>
<td>English</td>
<td>Junior High School</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>38</td>
<td>I. S. I. Magableh and Abdullah (2022)</td>
<td>1.81</td>
<td>0.27</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>39</td>
<td>Peters et al. (2022)</td>
<td>-0.30</td>
<td>0.11</td>
<td>Cognitive</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>40</td>
<td>Yavuz (2020)</td>
<td>1.07</td>
<td>0.48</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>41</td>
<td>Sapan and Mede (2022)</td>
<td>0.43</td>
<td>0.42</td>
<td>Cognitive</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>42</td>
<td>Sapan and Mede (2022)</td>
<td>0.17</td>
<td>0.41</td>
<td>Affective</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>43</td>
<td>Sapan and Mede (2022)</td>
<td>0.87</td>
<td>0.43</td>
<td>Affective</td>
<td>English</td>
<td>Junior High School</td>
<td>Europe</td>
<td>Small</td>
</tr>
<tr>
<td>44</td>
<td>I. S. I. Magableh and Abdullah (2020)</td>
<td>2.48</td>
<td>0.35</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>45</td>
<td>Aljaser (2019)</td>
<td>8.06</td>
<td>0.82</td>
<td>Affective</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>46</td>
<td>Bahgat et al. (2019)</td>
<td>3.12</td>
<td>0.38</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>47</td>
<td>Bont et al. (2009)</td>
<td>-0.60</td>
<td>0.37</td>
<td>Cognitive</td>
<td>Childhood Education</td>
<td>Kindergarten</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>48</td>
<td>Trninic et al. (2022)</td>
<td>0.26</td>
<td>0.23</td>
<td>Psychomotor</td>
<td>Physical Education</td>
<td>College Students</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>49</td>
<td>Tulbure (2011)</td>
<td>0.19</td>
<td>0.21</td>
<td>Psychomotor</td>
<td>Physical Education</td>
<td>College Students</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>50</td>
<td>Green et al. (2022)</td>
<td>0.44</td>
<td>0.23</td>
<td>Psychomotor</td>
<td>Physical Education</td>
<td>College Students</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>51</td>
<td>Mavidou and Kakana (2019)</td>
<td>0.07</td>
<td>0.16</td>
<td>Cognitive</td>
<td>Childhood Education</td>
<td>Kindergarten</td>
<td>Europe</td>
<td>Large</td>
</tr>
<tr>
<td>52</td>
<td>Osafor and Okigbo (2013)</td>
<td>0.52</td>
<td>0.25</td>
<td>Cognitive</td>
<td>Natural Science</td>
<td>Senior High School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>53</td>
<td>Eissa and Mostafa (2013) [Study 1]</td>
<td>-0.10</td>
<td>0.26</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>54</td>
<td>Eissa and Mostafa (2013) [Study 2]</td>
<td>0.03</td>
<td>0.26</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>55</td>
<td>Eissa and Mostafa (2013) [Study 3]</td>
<td>-0.60</td>
<td>0.26</td>
<td>Affective</td>
<td>Mathematics</td>
<td>Elementary School</td>
<td>Africa</td>
<td>Large</td>
</tr>
<tr>
<td>56</td>
<td>Matsumura et al. (2012)</td>
<td>0</td>
<td>0.4</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>America</td>
<td>Small</td>
</tr>
<tr>
<td>57</td>
<td>Nurasiah et al. (2020)</td>
<td>0.13</td>
<td>0.24</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Junior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>58</td>
<td>Aliakbari and Khales Haghighi (2014)</td>
<td>0.98</td>
<td>0.25</td>
<td>Cognitive</td>
<td>English</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>59</td>
<td>Suzzo et al. (2023)</td>
<td>1.47</td>
<td>0.4</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>60</td>
<td>Kado et al. (2021)</td>
<td>4.02</td>
<td>0.44</td>
<td>Cognitive</td>
<td>Mathematics</td>
<td>Senior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>61</td>
<td>I. Magableh and Abdullah (2020)</td>
<td>1.01</td>
<td>0.28</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>Asia</td>
<td>Large</td>
</tr>
<tr>
<td>62</td>
<td>Reis et al. (2011)</td>
<td>0.11</td>
<td>0.06</td>
<td>Cognitive</td>
<td>English</td>
<td>Elementary School</td>
<td>America</td>
<td>Large</td>
</tr>
<tr>
<td>63</td>
<td>I. S. I. Magableh and Abdullah (2021)</td>
<td>0.96</td>
<td>0.29</td>
<td>Cognitive</td>
<td>English</td>
<td>Junior High School</td>
<td>Asia</td>
<td>Large</td>
</tr>
</tbody>
</table>
3.3.1. Measured ability

The moderator variable, measured ability, encompasses three categories: cognitive, psychomotor, and affective abilities. The analysis results, depicted in Figure 2, demonstrate that the average effect size among these three measured ability groups does not differ significantly ($Q_b = 1.26; p = .17$). It implies that, when compared to traditional learning, the effectiveness of applying differentiated instruction on learning outcomes is unaffected by the specific evaluated ability. Among the three measured ability groups, the application of differentiated instruction showed the highest effectiveness when used to measure affective abilities ($g = 2.08; p < .01$).

3.3.2. Subjects

Subject moderator variables consist of five groups: mathematics, English, physical education, natural science, and childhood education. According to the findings of the analysis (see Figure 2), the average impact size of the five subject groups is not substantially different ($Q_b = 0.82; p = .02$). It reveals that the subjects have no effect on the effectiveness of differentiated instruction implementation on learning outcomes when compared to traditional learning. Of the five subject groups measured, differentiated instruction was most effective when used to measure English subjects ($g = 1.25; p < .01$).

3.3.3. Grade level

Kindergarten, elementary, junior high, senior high, and college are the five grade level moderator groups. According to the findings of the analysis (see Figure 2), the average effect size of the five grade levels is not substantially different ($Q_b = 0.91; p = .02$). It demonstrates that, when compared to traditional learning, grade level variations have no effect on the effectiveness of applying differentiated instruction on learning outcomes. The application of differentiated instruction to
learning outcomes is most effective if applied at the senior high school level ($g = 1.50; p < .01$) and followed by junior high school ($g = 1.4; p < .01$).

3.3.4. Country

The nation moderator variable is divided into four categories: Asia, Africa, America, and Europe. The analysis results (shown in Figure 2) show that the average effect sizes of the four country groups varied significantly ($Q_b = 0.88; p < .01$). It suggests that when compared to traditional learning, nation variations have an impact on the effectiveness of applying differentiated instruction on learning outcomes. Of the four countries studied, the application of differentiated instructions was most effective in countries on the Asian continent ($g = 1.79; p < .01$).

3.3.4. Sample size

The moderator variable sample size is divided into two categories: large samples and small samples. The findings of the analysis (see Figure 2) show that the average effect size of the two sample size groups is comparable ($Q_b = 0.93; p = .06$). It shows that sample size differences have no effect on the effectiveness of adopting differentiated teaching on learning outcomes when compared to traditional learning. Of the two sample size groups measured, applying differentiated instructions was most effective for large samples ($g = 1.19; p < .01$).

3.3.5. Index type

The index-type moderating variable is divided into two groups: Scopus and non-Scopus. According to the findings of the analysis (see Figure 2), the mean effect size of the two index-type groups is not statistically different ($Q_b = 1.13; p = .22$). It shows that the change in index type has no effect on the effectiveness of differentiated instruction implementation on learning outcomes when compared to traditional learning. Of the two groups of index types measured, the application of differentiated instructions is most effective in non-Scopus-indexed journals ($g = 1.36; p < .01$).

3.4. Evaluation of Publication Bias

Evaluation of publication bias can be seen from the results of the Funnel plot and the fail-safe N values of 63 studies. The funnel plot results show that all effect size plots form a pattern that tends to be symmetrical (see Figure 3).

Figure 3
Funnel plot
There is no evidence of publication bias in the data used for this study. Borenstein et al. (2009), state that when the fail-safe N value reaches $5K + 10$ (where $K$ indicates the number of individual studies), there is no publication bias in the meta-analysis. $K$ equals 63 in this study, resulting in a computation of $5(63) + 10 = 325$. The failed-safe N value for this study was 19849, with a goal significance of .05 and $p < .001$. These findings further confirm the absence of publication bias in this meta-analytic study. Therefore, there are no issues of publication bias encountered within this research.

4. Discussion

The collective findings from all 63 studies demonstrate that implementing differentiated instruction positively affect learning outcomes compared to traditional learning. This outcome can be attributed to several factors associated with the application of differentiated instruction in the learning process, contributing to its significant impact on student learning outcomes (Yavuz, 2020). Differentiated instruction emphasizes acknowledging and accommodating individual student differences, offering relevant and challenging learning experiences, promoting active student engagement, and providing constructive feedback (Marks et al., 2021). These findings are supported by research (Cruz et al., 2019; Whitley et al., 2021) that By taking into account the unique characteristics and needs of each student, the implementation of differentiated instruction enhances student engagement, motivation, understanding, and learning skills, ultimately leading to improved learning outcomes.

In the context of moderator variable analysis, it is known that the variables measured ability, subjects, grade level, sample size, and index type have no significant difference in effect size. It can happen because a differentiated learning approach accommodates individual student differences in measured ability, subjects, grade level, and sample size. By using different strategies, such as providing material according to the level of understanding of students or providing interesting learning approaches for students with various approaches, differentiated instruction can provide benefits for aspects of measured ability, subjects, grade level, and sample size (Laari et al., 2021). In terms of index type, these findings indicate no difference in findings between articles published in Scopus and non-Scopus indexed journals. Overall findings from each study resulted in consistent findings.

However, when viewed from a country perspective, there are significant differences in effect sizes. Several factors can cause the existence of significant differences. First, differences in educational policies between countries can affect the approach and level of support for implementing differentiated instruction (Faber et al., 2018). Countries with more robust education policies and support their implementation tend to achieve a more significant effect. Second, differences in educational resources such as facilities, teacher support, and curricula can impact effective implementation (Thapliyal et al., 2022). Countries with better educational resources tend to produce more significant effects. Third, cultural and contextual factors can also influence students’ responses to the application of differentiated instruction. Learning culture, educational values, and social norms can differ from country to country, which may influence the impact of different learning approaches (Bondie et al., 2019). Therefore, these factors play a role in the effect of implementing differentiated instruction on student learning outcomes in different educational contexts.

The findings from this meta-analysis study provide a significant theoretical contribution to understanding the application of differentiated instruction in learning contexts. This study presents solid empirical evidence regarding the positive impact of implementing differentiated instruction on learning outcomes compared to traditional learning. These findings support theories and concepts about accommodating individual student differences in learning. In addition, through the analysis of moderator variables, this study contributes to a deepening understanding of the factors that can moderate the effectiveness of implementing differentiated instruction, which
can become the basis for the development of more comprehensive theories and models in the educational context.

This study also provides a significant practical contribution to education practitioners and policymakers. These findings suggest that the consistent application of differentiated instruction positively impacts student learning outcomes, reinforcing evidence that this approach has value in diverse learning settings. By understanding the effectiveness of implementing differentiated instruction, educational practitioners can consider its use in designing and implementing learning strategies that are more inclusive and responsive to students' individual needs. This study, while confirming many expected findings, unearths valuable insights that warrant further exploration and deliberation in the sections such as the nuanced impact of cultural and contextual factors, can contribute to a richer dialogue about the practical implications and future directions of differentiated instruction in educational.

5. Conclusions

In summary, the findings underscore a substantial and positive influence of differentiated instruction on student learning outcomes. The implementation of strategies tailored to individual student needs and learning preferences emerges as a robust means of enhancing overall learning achievements. This study has unearthed insights into several critical dimensions, each contributing to a richer understanding of the educational landscape. Notably, the analysis reveals a remarkable consistency in effect size across various factors, including measured ability, subjects, grade levels, sample sizes, and index types. These results emphasize the universality of differentiated instruction's efficacy, indicating its potential to be a versatile tool irrespective of these contextual variables. However, a significant discrepancy in effect size comes to light when contemplating the vantage point of the country. This disparity underscores the pivotal role of cultural, educational, and institutional contexts specific to each country. These elements significantly shape the effectiveness of differentiated instruction, highlighting the need for a nuanced approach that accounts for regional specificities. From a broader perspective, the implications of these findings radiate across multiple realms of the educational ecosystem. Educators gain valuable insights into how tailored instructional methods can be harnessed to foster optimal learning outcomes for diverse students. Policymakers are presented with evidence that underscores the importance of incorporating differentiated instruction strategies within curricular frameworks. This study echoes a resounding call for educational practitioners to acknowledge the influence of contextual factors and adapt instructional strategies accordingly. As this manuscript's story unfolds, it imparts crucial takeaways that researchers, educators, policymakers, and stakeholders must internalize. The narrative not only affirms the potency of differentiated instruction but also accentuates the need for a holistic and context-sensitive implementation. By encapsulating these insights, this research contributes to the ongoing dialogue about the evolution of educational practices, facilitating a more tailored, effective, and inclusive learning environment for students across the globe.

Despite the considerable data indicating that differential instruction has a favorable impact on student learning outcomes, there are certain caveats to consider. To begin, the study was limited to a small number of nations, which may not adequately represent the global diversity of educational systems and cultural situations. It limits the generalizability of the findings and emphasizes the need for additional research that includes a broader range of countries and educational settings.

Recommendations for future research are to conduct comparative studies involving more countries to explore the impact of cultural, contextual, and systemic factors on the effectiveness of differential instruction. Moreover, investigations into specific subgroups, such as students with learning disabilities or from disadvantaged backgrounds, could shed light on the differential effects of this instructional approach across diverse student populations. Additionally, longitudinal studies that assess the long-term effects of differential instruction on student learning outcomes provide valuable insights into the sustained benefits of this approach. Furthermore,
incorporating qualitative research methods to gain a deeper understanding of teachers' and students' experiences and perspectives regarding the implementation of differential instruction would enhance the overall understanding of its effectiveness and practical implications.

**Author contributions:** All authors are agreed with the results and conclusions.

**Declaration of interest:** No conflict of interest is declared by authors.

**Ethics declaration:** This study does not require ethic approval, because it only analyzes the data from published literature.

**Funding:** No funding source is reported for this study.

**References**

Note: The meta-analyses included in the study are starred (*) in the list.


