A review of the effects of active learning on high order thinking skills: a meta-analysis within Islamic education

Karwadi Karwadi¹, Abd Razak Zakaria², Ahmad Syafii³

¹Department of Islamic Studies, Faculty of Tarbiyah and a Teacher Training, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta, Indonesia
²Department of Educational Foundation and Humanities, Faculty of Education, University Malaya, Kuala Lumpur, Malaysia
³Department of Educational Research and Evaluation, Yogyakarta State University, Yogyakarta, Indonesia

ABSTRACT

This study aims to ascertain the impact of the Active learning method on high-order thinking skills (HOTS) within the realm of Islamic education. Our primary contribution is the comprehensive meta-analysis of various preliminary and university educational studies. Using RStudio, we analyzed the differences in effect sizes, variations, and groups among studies. The findings indicate that the Active learning method significantly enhances HOTS, with the mixed learning method showing the most pronounced effect. We adopted a systematic approach, examining diverse learning methods like genre-based learning and blended learning model. Our research design entailed a meticulous meta-analysis of pre-existing literature rather than primary experimental or survey-based methods. However, we identified high heterogeneity in effect sizes as a limitation, suggesting the need for more nuanced studies in varied contexts. This research underscores educators' need to adapt learning strategies suited for specific educational levels, pointing towards a brighter horizon for tailored pedagogical interventions.

1. INTRODUCTION

High-order thinking skills (HOTS) are pivotal in cultivating students' analytical and critical thinking capabilities and fostering a nuanced understanding of religious matters. These skills empower students to examine, judge, and integrate information seamlessly, thus anchoring a profound comprehension of multiple academic disciplines [1]. However, several challenges remain, such as the continued use of rote learning, inadequate teacher preparation, and restricted technological access, which obstructs effective HOTS implementation in educational settings [2]. Given this backdrop, our research endeavors to illuminate innovative approaches and techniques, supplemented by new empirical insights, to negotiate these challenges and fortify the prominence of HOTS in Islamic education.

Several studies have emphasized the importance of HOTS in various academic subjects. In mathematics, it was found that students with high conceptual understanding abilities demonstrated HOTS in analyzing, evaluating, and creating aspects. In contrast, those with moderate understanding abilities showed HOTS in analyzing and assessing, and those with low understanding abilities did not show HOTS [3]. Similarly, developing a HOTS instrument on the subject of Fiqh using a 4D development model resulted in
appropriate items with a good level of difficulty and discrimination index, which showed the need for understanding, analysis, and synthesis in responding to Fiqh-related issues [4]. Additionally, a study found that the integration of HOTS in teaching Islamic Education was positively and significantly related to a professional learning community and self-efficacy, where self-efficacy played a mediating role in the relationship, with teachers’ self-efficacy having the most significant influence on the teaching of Islamic Education integrated with HOTS [5]. Implementing 21st-century competency-based learning improved HOTS and numerical literacy among public Islamic university students in Indonesia. Gender was found to partially influence and relate to these skills, with male students demonstrating higher HOTS but lower numerical literacy than female students [6].

In the study of aqidah at madrasah alyiah, the significance of developing HOTS to connect moderate values in daily life has strengthened students’ analytical abilities regarding religious moderation [7]. The research found that problem-based learning (PBL) and case-based learning (CBL) can stimulate various cognitive skills among medical students, with PBL being more effective in promoting exploratory questions and CBL being more effective in fostering procedural interactions [8]. To enhance the quality of HOTS-based Islamic religious education (PAI) in elementary schools, it is necessary to improve teachers’ competencies in comprehensively integrating HOTS concepts into their planning, implementation, and evaluation and to receive support from various parties [9]. Wisdom pedagogy empowers students in Islamic education by enhancing HOTS, inquiry abilities, and independent thinking, thereby deepening the understanding and internalization of Islamic principles and values [10]. Using crossword puzzles to learn maharah qira’ah during the COVID-19 pandemic at Islam Sabilillah Malang high school has met HOTS standards and enhance students’ critical and in-depth thinking skills [11]. The development of religious chemistry teaching materials based on a 4D model that encompasses vocational contexts and character education produces validated teaching materials that promote HOTS and are suitable for contextual vocational chemistry [12].

Many meta-analytical studies have scrutinized the evolution of HOTS. In contemporary times, the educational system is in constant flux as it explores novel and innovative methodologies to enhance the pedagogical process. One approach that has demonstrated favorable outcomes beyond cognition is the philosophy for children; however, its implementation can pose challenges for educators and learners [13]. Hence, the significance of acquiring effective learning strategies cannot be overstated. Other studies have ascertained that judicious employment of media, methods, and assessments can elevate students’ HOTS [14]. To further bolster HOTS and augment the learning environment, the integration of science, technology, engineering, and mathematics (STEM) has been proposed [15].

To effectively teach HOTS, metacognitive skills should also be utilized. A systematic literature review on metacognition-based HOTS teaching modules can guide teachers in applying these skills [16]. PBL has also been found to have a very high positive effect in enhancing Indonesian students’ HOTS in mathematics learning [17]. On the other hand, innovative education has been shown to affect HOTS student achievement positively and can be used by teachers or practitioners to improve HOTS [18]. Our paper contributes novel technical findings from an extensive meta-analysis, shedding light on underexplored dimensions of HOTS in Islamic education. This research not only elaborates on the importance of diversified learning strategies but also introduces fresh technical solutions founded on our comprehensive analysis to surmount associated challenges.

In addition, studies have shown that STEM enactment in Asia effectively improves students’ learning outcomes, with STEM integrated with project-based learning being preferred [19]. However, there is a lack of scholarly attention to HOTS in gifted education despite moderately large effect sizes found in HOTS interventions with talented students in Taiwan [20]. Furthermore, HOTS pedagogical practices in Malaysian classrooms include both techniques that promote brainstorming and inquiry teaching and practices that inhibit, such as teacher-centered learning and lower-order thinking skills [21]. These findings highlight the importance of incorporating various teaching strategies and methodologies to enhance HOTS in students effectively.

This study examines the effect of learning methods on HOTS in PAI subjects, seeking to fill the knowledge gap and provide recommendations for education practitioners, teachers, and policymakers. The significance of this research lies in its contribution to developing an effective curriculum and pedagogical approach to enhance the quality of PAI learning and students’ critical thinking skills through a comprehensive analysis of data from various previous studies. Moreover, we present novel pedagogical methodologies underpinned by unique empirical evidence to instill HOTS in students efficaciously.

2. METHOD

This study adopts the form of a meta-analysis, which examines empirical research about a singular subject or related area while grouping studies based on specific parameters. The present meta-analysis
integrates quantitative findings relevant to a particular study [22]. Cohen [23] describe meta-analysis as an “analysis of analyses.” Meta-analysis can enhance the validity of a specific research field by comparing the results of studies with those of related findings. In the current report, the meta-analysis procedure evaluates the effect size of multiple intelligence concepts on non-scientific successes. The databases include the Google Scholar database, Academic Microsoft, directory of open access journals (DOAJ), Emerald, Garuda Ristekdikti, Jstor, open access theses and dissertations (OATD), Sage, Science Direct, Springer, Taylor Francis, and Wiley employed in this meta-analysis. The researcher’s first step is to create relevant keywords for the study. The following keywords are used: “(ALL (high AND order AND thinking AND skill OR hots) AND ALL (method) AND ALL (learning OR teaching OR education))”.

The investigation yielded 2415 titles of scholarly articles, theses, and dissertations, subsequently subjected to the scrutiny of inclusion and exclusion criteria. The inclusion and exclusion criteria applied in this analysis were as: i) the research methodology employed must be quantitative, ii) the study must have been conducted within the time frame of 2016 to 2020, iii) the publication must have been a thesis, dissertation, or article published in a scholarly journal, and iv) the research must contain numerical data, specifically about the population size, means, and standard deviations of both experimental and control groups. Microsoft Excel was employed as the coding format to ensure the investigation's integrity, requiring the input of name, material, and analysis data.

Figure 1 illustrates the preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram of the study selection process, which commenced with a total of 2,415 article titles. After the application of inclusion and exclusion criteria, a subset of studies using an experimental design was identified, alongside those that did not employ this methodology. Further refinement was made by excluding studies that did not conform to formal coding standards, culminating in a final sample of studies utilized for research.

Upon careful examination of the collected data, this study comprises 17 articles written by Guntur and Mustafa [24], Channa [25], Wildan [26], Putri and Setiawan [27], Harmiyanti [28], Sinaga et al. [29], Masran and Esha [30], Bahrun et al. [31], Susanto [32], Rahman and Noh [33], Anwar [34], Ghazali et al. [35], Ardhiyat et al. [36], Masoud [37], Rini and Mansur [38], Kusaeri et al. [39], Adam and Latif [40]. The research team thoroughly analyzed each article, ensuring the results were reliable and relevant to the study’s objectives. Two moderating variables were observed in this study: the learning methods employed and the educational level of the schools involved, both of which played a significant role in determining the outcomes.

This study identified and analyzed five distinct learning methods: genre-based learning, blended learning model, group learning, contextual learning approach, and technology-based learning. Each technique was investigated for its ability to enhance students’ HOTS. The educational levels of the schools involved in the study varied, ranging from primary, secondary, high schools, and universities, providing a diverse sample for analysis (refer to Table 1 for more details).

RStudio was applied to compare the various learning methods’ effect sizes, variances, and subgroups to facilitate an in-depth analysis. This statistical method allowed the research team to examine each method's impact on the development of HOTS in a systematic manner. The experimental group

Figure 1. PRISMA
The results of this research indicate significant variation in the use of learning methods across different levels of education. Although not directly observable from the data, several important insights can be gleaned. First, technology-based learning methods are more dominant than other methods. It indicates that technology has become an essential aspect of education and supports innovation in teaching and learning. In addition, the use of group learning methods and PBL methods also stands out, indicating that collaborative and problem-solving approaches are becoming increasingly important in the current educational context.

Meanwhile, data analysis based on school levels shows that high school is the level of education with the most significant variation in learning methods. It may indicate that educators at the high school level are more open to trying out various learning methods to improve student learning outcomes. On the other hand, other education groups, such as universities, junior high schools, and elementary schools, also implement various learning methods, but in a lower proportion. It indicates that at these levels of education, there is still room for developing innovation in learning methods and adapting approaches appropriate to the needs and characteristics of students at each level of education.

3. RESULTS AND DISCUSSION
3.1. Overview of studies and effect sizes
In this meta-analysis, RStudio was the primary tool for comparing effect sizes, variations, and groups in the context of HOTS intervention research. Studies were categorized based on their experimental design to cater to an international audience, with the experimental group receiving an intervention grounded in the active learning method and the control group receiving conventional learning treatments. Consequently, positive effect sizes were interpreted as reflective of the efficacy of HOTS interventions, while negative effect sizes indicated the prevalence of traditional learning practices.

Various studies were compiled and analyzed, adopting a significance level 0.05 to maintain consistency across the included research. As such, the meta-analysis adhered to this 0.05 level of statistical significance. Moreover, the interpretation of effect size values was grounded in Cohen’s [23] criteria for evaluating the magnitude of the effect. According to Cohen [23], effect sizes ranging from 0.20 to 0.49 denote a small impact, those falling within the 0.50 to 0.79 range signify a moderate effect, and values of 0.80 or greater suggest a substantial effect. Using Cohen's criteria as a benchmark not only establishes the validity of our assessment but ties back to the more extensive debate on quantifying the impact of educational interventions.

3.2. Publication bias
Acknowledging the international concern about publication bias, this research followed global best practices to detect it. The inclination towards publishing only positive results is emphasized to ensure comprehensive understanding. It's crucial to highlight that publication bias can distort the interpretations drawn from meta-analytic research. The propensity for articles to be published with positive or statistically significant results is more substantial than those displaying negative or non-significant outcomes, thereby potentially skewing the studies included in meta-analyses [41]. This phenomenon, known as publication bias,
can lead to inflated effect size estimates when conducting meta-analytic research [42]. Consequently, assessing for publication bias before conducting a meta-analysis is crucial. The issue of publication bias is not merely a matter of skewed representation. Still, it can lead to misguided policy decisions and instructional practices, mainly when educators rely on these research findings to inform pedagogical strategies.

Various methods can be employed to detect and account for publication bias in meta-analytic studies. This research implemented four widely utilized tests to evaluate potential publication bias: the Funnel Scatter Plot, the Trim and Fill method, and Rosenthal’s Fail-Safe N [43]. By incorporating these methods, the likelihood of publication bias affecting the meta-analytic results can be more accurately determined. Each test serves to quantify the potential presence of publication bias, offering insight into the overall reliability and validity of the meta-analysis findings. By conducting these tests, researchers can better ensure that the selective reporting of studies in the literature does not unduly influence their meta-analytic conclusions.

The funnel plot depicted as Figure 2 is a scatter plot that is used to assess the presence of publication bias in this research. In this graphical representation, the standard error (SE) of the estimated effect sizes from individual studies is plotted on the vertical axis against the effect sizes themselves on the horizontal axis. The expected distribution of data points should form a symmetrical, inverted funnel shape in the absence of bias. This is because larger studies, which appear at the top of the plot due to their smaller SE, are expected to cluster around the true effect size, while smaller studies with larger SE scatter more widely at the bottom of the plot. The presence of asymmetry in the funnel plot can suggest potential publication bias, with the plot in Figure 2 showing a degree of asymmetry indicating that smaller studies with less significant results may be underrepresented in the meta-analysis.

Based on the funnel plot graph, the position of each study is observed on the X and Y axis. The chart indicates that the studies are evenly distributed on both axis sides. The shape of the data distribution is symmetric, with a slight inclination towards the right. If consulted with various forms of funnel plots as described in academic books on a meta-analysis [44], [45], it can be understood that there is no publication bias in the observed study data on the meta-analysis of the effects of the active learning method on HOTS in Islamic education. Figure 3 presents the output from RStudio utilizing the Duval and Tweedie's Trim and Fill method, which is a statistical approach used to adjust for publication bias in meta-analyses. The Figure 3 summarizes the results of a meta-analysis that combines a total of 24 studies, with 6 additional studies included to correct for potential bias.

The findings indicate that four studies could potentially be added to enhance the precision and symmetry of the forest plot. A more precise judgment regarding publication bias may be achieved using the fail-safe N value. In the fail-safe N analysis of publication bias, it is posited that for any given study, there are several other unpublished studies whose findings are essentially null or, in other words, whose effect sizes are negligible. This assumption is based on the notion that publication bias arises from the selective reporting of statistically significant results, and thus, unpublished studies fail to reach significance.

Figure 4 shows the RStudio output for a publication bias assessment using the fail-safe N calculation based on the Rosenthal approach. This method is used to estimate the number of additional studies with null results (non-significant findings) that would be needed to bring the observed significance level of a meta-analysis above a specified threshold, here set at 0.05, which would indicate that the meta-analytic findings...
are robust to publication bias. Based on the output analysis of R Studio, the calculation results show that the value of Fail-Safe N is 4107, with a significant target level of 0.05. When the value of Fail Safe N > 5K + 10 (K = several studies), it can be concluded that there is no publication bias [46]. The calculation results indicate that the fail-safe N = 4107 exceeds 5 (18) + 10 = 100. Therefore, it can be concluded that there is no publication bias in the observed studies.

![Figure 3. Output RStudio duval and tweedie trim and fill method publication bias test](image)

<table>
<thead>
<tr>
<th>Number of studies combined: k = 24 (with 6 added studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMD 95% CI p-value</td>
</tr>
<tr>
<td>Random effects model (HR): 0.9338 [0.5640; 2.3415] 1.29 0.2100</td>
</tr>
<tr>
<td>Quantifying heterogeneity:</td>
</tr>
<tr>
<td>tau^2 = 12.2663 [7.3842; 24.6986]; tau = 3.5023 [2.7137; 4.9697]</td>
</tr>
<tr>
<td>I^2 = 98.6% [98.4%; 98.9%]; H = 8.50 [7.83; 9.23]</td>
</tr>
<tr>
<td>Test of heterogeneity:</td>
</tr>
<tr>
<td>Q d.f. p-value</td>
</tr>
<tr>
<td>1092.65 23 0</td>
</tr>
</tbody>
</table>

Figure 3. Output RStudio duval and tweedie trim and fill method publication bias test

![Figure 4. RStudio output bias test analysis publication fail-safe n rosenthal method](image)

<table>
<thead>
<tr>
<th>Fail-safe N Calculation Using the Rosenthal Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Significance Level: &lt; .0001</td>
</tr>
<tr>
<td>Target Significance Level: 0.05</td>
</tr>
<tr>
<td>Fail-safe N: 4107</td>
</tr>
</tbody>
</table>

3.3. Combined findings and sub-group analysis

In this section, the study’s core findings are integrated with relevant literature, placing them in the broader context of current international educational research. The study discovered that the effect size value on student achievement by implementing multiple intelligence-dependent learning practices was 1.640558 with the fixed-effect model and 2.4254 with the random effect model, as evidenced by meta-analysis test results conducted using the RStudio. Concerning the homogeneity test’s outcomes, the data appears heterogeneous concerning the Q-value of 749.62 and the P-value of 0.0001. This paper builds on Yildiz’s assertion that, due to its homogeneous effects, the random effect paradigm is employed. Moreover, the analysis encompasses the results of the Active learning method on HOTS, utilizing a random paradigm. Building upon Yildiz’s assertion, adopting the random effect paradigm was deemed necessary. This decision was rooted in the need for our study to be congruent with established research methods, ensuring the validity of our findings. The research estimates indicate that the random effect value observed in this meta-analysis is 1.64055. These findings correspond to the substantial effects classification, as delineated by Cohen [23]. Further details can be found in Figure 5.

Figure 5 is a forest plot, a graphical display designed to illustrate the relative strength of treatment effects in multiple quantitative scientific studies addressing the same question. This forest plot includes individual studies listed on the y-axis, with their corresponding SE, standardized mean differences (SMD), and 95% confidence intervals (CI). The size of the squares reflects the weight of each study in the meta-analysis, while the horizontal lines represent the 95% CI.

In addition, the I^2 value obtained surpasses 98%, suggesting a high level of heterogeneity in the effect size of studies. Moderator variables, such as the level of education and type of learning method, are employed to elucidate the distribution of this heterogeneity. Moderator variables influence the outcomes in meta-analysis and serve to ascertain the extent of the meta-analysis impact. Consequently, a subgroup analysis was conducted to assess the moderator variables within the scope of this review.

Figure 6 provides the RStudio output for a subgroup analysis within a meta-analysis, examining the effect sizes of different types of learning methods. The subgroup analysis allows for the comparison of effects across categorically different studies to understand if and how the impact of the interventions varies. The Figure 6 lists various learning methods as subgroups, along with their respective SMD, 95% CI, tau^2

J Edu & Learn, Vol. 18, No. 1, February 2024: 97-106
values (representing between-study variance), tau (the standard deviation of underlying effects across studies), and I² (the percentage of variation across studies due to heterogeneity rather than chance).

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>seTE</th>
<th>Standardised Mean Difference</th>
<th>SMD</th>
<th>95%-CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muhammad Firdq Nur Iman Syukurullah (2021)</td>
<td>0.01</td>
<td>0.2236</td>
<td>-0.01</td>
<td>0.45</td>
<td>0.43</td>
<td>5.7%</td>
</tr>
<tr>
<td>Ayu Alim Fitri, et al. (2020)</td>
<td>0.00</td>
<td>0.0301</td>
<td>0.00</td>
<td>0.60</td>
<td>0.62</td>
<td>5.6%</td>
</tr>
<tr>
<td>Fazlin Bibi Rahman, et al. (2021)</td>
<td>0.02</td>
<td>0.2236</td>
<td>0.02</td>
<td>0.42</td>
<td>0.44</td>
<td>5.6%</td>
</tr>
<tr>
<td>Ahmad Zubaidi (2021)</td>
<td>0.68</td>
<td>0.2256</td>
<td>0.68</td>
<td>0.36</td>
<td>0.32</td>
<td>5.7%</td>
</tr>
<tr>
<td>Dina Azlina Sarihi (2019)</td>
<td>0.45</td>
<td>0.3203</td>
<td>0.45</td>
<td>0.17</td>
<td>0.19</td>
<td>5.6%</td>
</tr>
<tr>
<td>Wirda Seri Habibawati, et al. (2022)</td>
<td>0.58</td>
<td>0.3456</td>
<td>0.58</td>
<td>0.10</td>
<td>0.12</td>
<td>5.6%</td>
</tr>
<tr>
<td>Irwan Aziz Fathiyah (2022)</td>
<td>0.72</td>
<td>0.3284</td>
<td>0.72</td>
<td>0.08</td>
<td>0.10</td>
<td>5.6%</td>
</tr>
<tr>
<td>Ricky Hamyanto (2023)</td>
<td>0.90</td>
<td>0.3326</td>
<td>0.90</td>
<td>0.85</td>
<td>0.95</td>
<td>5.6%</td>
</tr>
<tr>
<td>Hamzani Salim, et al. (2022)</td>
<td>1.00</td>
<td>0.4340</td>
<td>1.00</td>
<td>0.39</td>
<td>0.39</td>
<td>5.6%</td>
</tr>
<tr>
<td>Muhammad Gunjur (2022)</td>
<td>2.96</td>
<td>0.3579</td>
<td>2.96</td>
<td>2.20</td>
<td>3.62</td>
<td>5.6%</td>
</tr>
<tr>
<td>Deyas Ane Limarao, et al. (2021)</td>
<td>2.96</td>
<td>0.4554</td>
<td>2.96</td>
<td>2.04</td>
<td>3.56</td>
<td>5.6%</td>
</tr>
<tr>
<td>Usanto, et al. (2020)</td>
<td>2.00</td>
<td>0.9531</td>
<td>2.00</td>
<td>2.30</td>
<td>3.48</td>
<td>5.6%</td>
</tr>
<tr>
<td>Siti Maisarah Bani, et al. (2019)</td>
<td>3.14</td>
<td>0.4511</td>
<td>3.14</td>
<td>2.25</td>
<td>4.02</td>
<td>5.6%</td>
</tr>
<tr>
<td>Lili Orma (2020)</td>
<td>3.86</td>
<td>0.4742</td>
<td>3.86</td>
<td>2.80</td>
<td>4.76</td>
<td>5.6%</td>
</tr>
<tr>
<td>Usanto, et al. (2020)</td>
<td>3.84</td>
<td>0.4160</td>
<td>3.84</td>
<td>3.12</td>
<td>4.75</td>
<td>5.6%</td>
</tr>
<tr>
<td>Effendi Aulia Marding (2018)</td>
<td>0.15</td>
<td>0.0807</td>
<td>0.15</td>
<td>0.02</td>
<td>0.29</td>
<td>5.4%</td>
</tr>
<tr>
<td>Tiera Kusum Jandiyet, et al. (2020)</td>
<td>5.57</td>
<td>0.2725</td>
<td>5.57</td>
<td>6.03</td>
<td>7.10</td>
<td>5.6%</td>
</tr>
<tr>
<td>Hoesva Mostafa (2023)</td>
<td>7.71</td>
<td>0.7621</td>
<td>7.71</td>
<td>6.21</td>
<td>8.12</td>
<td>5.2%</td>
</tr>
</tbody>
</table>

Figure 5. Forest plot

Figure 6. Output RStudio subgroup variable type of learning method

Considering international educational settings, moderator variables such as education level and learning method type were examined. For global understanding, the subgroups were identified as genre-based learning, blended learning model, group learning, contextual learning approach, and technology-based learning. The analysis results indicate the presence of heterogeneous variance estimates for the level of education variable [47]. The test for subgroup differences between groups reveals a Q-value of 4.03 and a p-value of 0.5458 < α (0.05). A distinction between the types of learning methods on the HOTS in religious education subjects is evident, with mixed learning methods exhibiting the highest effect [47], [48]. Combining learning methods, such as mixed learning methods, yields superior outcomes in enhancing high-level thinking skills [49]. The primary challenge in implementing STEM-based contextual learning methods is integrating subjects and providing contextual aspects of students’ real-life experiences [50]. It highlights the importance of considering the context and needs of students when implementing learning methods such as contextual learning and mixed learning methods.

The analysis revealed a significant influence of learning methods on students’ high-level thinking abilities (2.4592), with high heterogeneity among studies. High heterogeneity is also apparent in subgroups divided based on education levels (Q = 5.42, p-value = 0.1434). The highest effect was found among middle school students (3.5382), followed by university students (2.5800), high school students (2.3955), and elementary school students exhibiting the lowest effect (1.8051). These findings align with prior literature suggesting that school-based programs enabling students to actively take responsibility for their learning effectively promote awareness and healthy lifestyles [51]. Additionally, active learning strategies were similar to middle and high school instruction [52]. Other research identified alternative ways of encouraging critical thinking and active participation through library activities and smartphone applications [53]. However, these findings contrast with more specific studies, such as problem-solving in mathematics teacher
education and vocabulary instruction for low-income background students [54], [55], highlighting the diverse global academic landscape.

The interrelation of the findings with the broader body of literature underscores this meta-analysis's contributions to the Islamic education field. Nevertheless, both studies also emphasize the importance of active learning in enhancing students’ understanding and skills. Overall, the findings of this research reinforce the significance of practical learning methods in augmenting students’ high-level thinking abilities and highlight variations in the effects of learning methods based on education levels. Therefore, strategies and interventions designed for different education levels may be more effective in improving student learning outcomes. The implications of these findings suggest that educators and policymakers in the realm of Islamic education need to be keenly aware of the learning methods they employ, considering their students’ diverse needs and contexts. Traditional pedagogies might not yield the desired results in promoting HOTS, hence the importance of exploring alternative, evidence-based strategies.

4. CONCLUSION

In summary, this meta-analysis aimed to evaluate the effects of the Active learning method on HOTS in Islamic education. The findings demonstrate a noteworthy positive correlation between the intervention and enhanced student achievement, with the mixed learning method showing the most pronounced effect. This research underscores the significance of hands-on learning techniques while emphasizing the imperative nature of considering students’ unique contexts and individual needs when deploying these methodologies. Beyond their immediate academic context, these findings have broader implications for pedagogical approaches across various disciplines, suggesting that active learning can significantly bolster students' critical thinking and analytical capabilities. The paramount importance of these discoveries lies in their potential to guide the creation of tailored strategies and interventions for diverse educational tiers, thereby elevating student academic outcomes. Recognizing its constraints, this study points out the evident high variability in effect size across the studies considered. As such, it paves the way for potential avenues of research, advocating for a more granular investigation into the efficacy of the Active learning method on HOTS across varying educational landscapes.

REFERENCES

A review of the effects of active learning on high order thinking skills: a meta-analysis ... (Karwadi Karwadi)


**BIOGRAPHIES OF AUTHORS**

**Karwadi Karwadi** is a senior lecturer at Tarbiyah and a Teacher Training Faculty at Universitas Islam Negeri (UIN) Sunan Kalijaga. He received his doctorate in Islamic Studies Department from UIN Sunan Kalijaga in 2009. He obtained his bachelor’s degree from IAIN Sultan Thaha Salfuddin Jambi in 1995 and his master’s degree from the Islamic Studies Department of IAIN Ar-Raniry Banda Aceh in 1998. He has written several papers in the area of Islamic education. His research interests include evaluating learning, learning media, and curriculum improvement. He can be contacted at email: karwadi@uin-suka.ac.id.

**Abd Razak Zakaria** is a senior lecturer at the Department of Educational Foundation and Humanities, Universiti Malaya (UM), Malaysia. His academic qualifications include a Ph.D. from the University of Malaya in 2011, a Master’s degree from Universiti Kebangsaan Malaysia (UKM) in 2005, and a Bachelor’s degree from UKM in 2003. His expertise includes social and community psychology, cultural history, deviance, and delinquency. He can be contacted at email: abdzazak@um.edu.my.

**Ahmad Syafii** is a postgraduate student in the educational research and evaluation study program at Yogyakarta State University, Yogyakarta, Indonesia. He worked as an assistant editor at an Islamic education journal. His current research interests include Information Management systems, Islamic Studies, and big data in education. His publication topics include electronic resources, education management, and educational curricula. He can be contacted at email: azfy99@gmail.com.