MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(3), em2592
ISSN:1305-8223 (online)
OPEN ACCESS Review Article https://doi.org/10.29333/ejmste/15988

The effect of inquiry-based learning on students' critical thinking skills in science education: A systematic review and meta-analysis

Zainal Arifin [1], Sukarmin [2*], Sulistyo Saputro [3], Azlan Kamari [4]

[1] Doctoral Program of Science Education, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Surakarta, INDONESIA
[2] Department of Physics Education, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Surakarta, INDONESIA
[3] Department of Chemistry Education, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Surakarta, INDONESIA
[4] Department of Chemistry, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Perak, MALAYSIA

Received 02 October 2024 ▪ Accepted 20 January 2025

Abstract

Inquiry-based learning (IBL) is regarded as an effective model for cultivating higher-order thinking skills but the precise components that enhance critical thinking skills remain inadequately comprehended. This study aims to investigate the impact of IBL on critical thinking skills in further detail. A literature search was performed utilizing the Scopus and ERIC databases for publications from 2000 to 2024. Specifically, quantitative empirical studies were reviewed with pre-experimental or quasi-experimental designs. 25 articles (divided into 36 studies) were selected and showed a substantial mean effect size of 1.27 [95% confidence interval: 0.78; 1.76], indicating a significant impact of IBL. Substantial heterogeneity (I^2 = 92.0%) indicated contextual variation, which necessitated subgroup analysis based on educational level, teaching strategies, information and communications technology utilization, duration, and evaluation. The findings highlight nuances in applying IBL and propose evidence-based recommendations to optimize its application in education, thus significantly contributing to the discussion on effective pedagogies for developing critical thinking skills.

Keywords: inquiry-based learning, critical thinking skills, science education, systematic review, meta-analysis

INTRODUCTION

The role of inquiry-based learning (IBL) has garnered significant interest in the field of science education in recent years (Jegstad, 2023; Strat et al., 2024). IBL is acknowledged as an effective method for enhancing students' higher-order thinking skills (Antonio & Prudente, 2023; Öztürk et al., 2022), particularly in fostering critical thinking skills (Gómez & Suárez, 2020; Kaczkó & Ostendorf, 2023). The cultivation of critical thinking skills is an essential objective in science education, as numerous studies highlight its importance in equipping students to address complex real-world challenges (Kabataş Memiş & Çakan Akkaş, 2020; Paige et al., 2024; Rizki & Suprapto, 2024). Critical thinking has been accommodated by inquiry learning models to enhance critical thinking (Kilbane & Clayton, 2023; Wale & Bishaw, 2020; Zheng et al., 2018).

However, the implementation of IBL faces various challenges (Bansal, 2021; Inel-Ekici & Ekici, 2022; Nawanidbumrung et al., 2022). For instance, a study by Arsal (2017) has found that pre-service teachers who practiced IBL teaching did not show statistically significant progress in critical thinking dispositions among students as compared to those taught with conventional teaching. Most teachers need a more transparent and adequate understanding of critical thinking or the requirements to develop students' critical thinking skills and competencies. Research by Essalih et al. (2024) has shown that little initiative is given to students who are expected to produce a sequence of steps or actions in a particular order without prior reflection and without necessarily understanding it. Teachers do not give the impression that they understand the requirements for cultivating critical

✉ zainarif@student.uns.ac.id ✉ sukarmin67@staff.uns.ac.id **(*Correspondence)** ✉ sulistyo_s@staff.uns.ac.id
✉ azlan.kamari@fsmt.upsi.edu.my

Contribution to the literature

- This systematic meta-analysis provides robust empirical evidence on the specific effects of IBL on students' critical thinking skills within the context of science education.
- Unlike prior studies that broadly examine higher-order thinking skills, this research specifically targets how various aspects of IBL impact critical thinking outcomes.
- The study applied a rigorous quantitative methodology, analyzing a large number of empirical research articles to guarantee thorough insights. Key moderating factors, including education level, teaching strategy, information and communications technology (ICT) media, duration, and type of assessment, were identified as elements that could improve the effectiveness of IBL.

thinking in students. Enugu and Hokayem (2017) studied the challenges pre-service teachers faced when implementing the 5E model. They found that they did not have sufficient content knowledge and thus had difficulties explaining and elaborating phases. Teachers should practice guiding, assisting, and facilitating the integration of critical thinking into their teaching style (Ramnarain, 2023; Soysal, 2021). Meanwhile, research by Yun and Crippen (2024) found a gap in utilizing new technologies. More attention needs to be paid to the use of technology to support the modelling process by students and the utilization of simulations. These findings indicate that there is an urgent need for effective educational interventions in promoting critical thinking in science education (Essalih et al., 2024).

Although previous studies have reviewed the evidence of the positive impact of IBL on critical thinking skills (Boonsathirakul & Kerdsomboon, 2023; Praminingsih et al., 2022; Sari et al., 2023), the weakness of previous findings is the need for a systematic and comprehensive approach in evaluating the impact of IBL on critical thinking skills. Many studies focus only on one aspect of IBL or only on a specific group without considering the various contextual factors that can influence learning outcomes (Harefa et al., 2016). Gaps in the literature regarding the systematic synthesis of existing research on the impact of IBL on students' critical thinking skills. Most studies focus on individual or isolated outcomes, and there is a need for comprehensive meta-analytic research that combines findings from multiple studies to provide a more robust and generalizable understanding of this relationship (Ananda & Usmeldi, 2023).

The novelty of this study lies in its more comprehensive approach to evaluating the impact of IBL on critical thinking skills through meta-analysis. The study examines the overall impact of IBL and explores how factors such as teaching strategies, ICT media, duration, education level, and assessment methods. By thoroughly considering the empirical evidence, this study offers critical new insights for developing more effective teaching strategies, media use, and assessment in IBL. This synthesis evaluates the effectiveness of IBL and identifies the conditions that enhance its benefits, thereby contributing to improved educational practices and policies.

This study aimed to assess the effectiveness of IBL on the critical thinking skills of students in science education. The specific objectives are to identify the characteristics of effective IBL interventions, examine the influence of contextual factors, and provide recommendations for future practice and research. This study aims to enhance science education and develop critical thinking skills essential for success in contemporary society. This study aims to elucidate the specific impact of IBL on critical thinking skills and to identify potential moderating factors. The following research questions (RQs) will be addressed:

RQ1. What is the overall effect of IBL on the critical thinking skills of students?

RQ2. What is the effect of IBL on students' critical thinking skills considering moderator variables?

METHOD

Meta-Analysis Framework

Meta-analysis serves as an essential method for integrating knowledge across diverse scientific disciplines, including education, to synthesize prior research and empirically validate theoretical frameworks (Chaudhary & Singh, 2022; Tikito & Souissi, 2019). The meta-analysis procedure in this study followed the steps proposed by Hansen et al. (2022). This study used the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (Barta et al., 2022; Wei et al., 2023) to guarantee a systematic and transparent review process. The 4-phase flow chart serves to visualize and monitor the various stages of this analysis (**Figure 1**).

The process encompasses distinct phases ranging from planning to the presentation of results (Geng & Su, 2024). Formulating RQs and establishing inclusion and exclusion criteria for selected studies (Ting et al., 2023). Gathering and choosing studies that meet the established criteria. Data were extracted from each selected study based on the relevant variables. The data extracted were organized and prepared for statistical

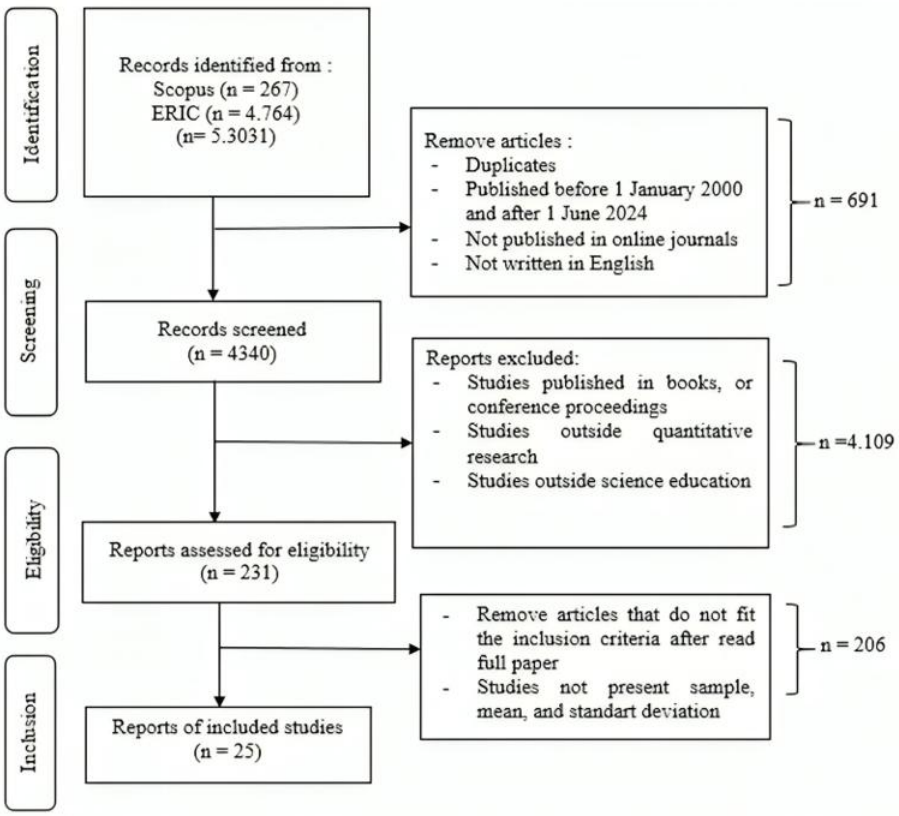

Figure 1. PRISMA flow diagram (Source: Authors' own elaboration)

Table 1. The specific search in the Scopus and ERIC databases

Database	Specific search
Scopus	TITLE-ABS-KEY (effect AND inquiry AND on AND critical AND thinking)
ERIC	Effect inquiry based learning on critical thinking

analysis. This analysis's results were interpreted by examining the practical and theoretical implications (Yu & Xu, 2022).

Search Strategy and Sources

Table 1 presents the search terms employed in the Scopus and ERIC databases to locate literature regarding the impact of IBL on critical thinking skills. The search aimed to identify empirical studies that examined the impact of IBL on critical thinking skills, facilitating a thorough meta-analysis of the findings.

Inclusion and Exclusion Criteria

Journal articles were chosen due to their rigorous peer review process, which ensures high academic reliability. The analyzed publication period spans from January 2000 to June 2024. This timeframe was selected to encompass pertinent data from the past to the present, facilitating a review that reflects the evolution of research in this domain over an extended duration (Smela et al., 2023). The review included only studies published in English. This decision was made to streamline the review process and facilitate accurate interpretation and analysis of the included studies by researchers. The review encompassed studies from various countries, providing a comprehensive perspective on the research topic (Kolaski et al., 2023). This review employed a quantitative research method, facilitating the analysis of statistically significant results. This review excluded qualitative and mixed-methods research to concentrate on studies with clearly measurable outcomes that allow for aggregation and comparison (Nakagawa et al., 2023). **Table 2** presents the inclusion and exclusion criteria employed for the selection of studies in this systematic review.

Statistical Analysis

This research employed R Studio, a widely utilized statistical analysis platform in meta-analysis (Le Thi Tuyet et al., 2024). The collected data typically encompasses effect size, standard deviation (SD), sample size (N), and moderator variables including education level, assessment, ICT media, and study location. After data collection, it is imported into R Studio for structural analysis. The pooled effect sizes of the analyzed studies were derived using a random-effects model. The analysis subsequently produces summary statistics of the aggregated effect sizes (Harrer et al., 2021).

The effect size measures the difference between the control and experimental groups according to the

Table 2. Inclusion and exclusion criteria (Strat et al., 2024; Winje & Løndal, 2020)

Criteria	Inclusion	Exclusion
Type of publication		
Journal articles	x	
Conference papers		x
Reports		x
Dissertations		x
Books and book chapters		x
Publication period		
January 2000-June 2024	x	
Language		
English	x	
Other		x
Place of study		
Worldwide	x	
Type of study		
Empirical investigations	x	
Literature reviews		x
Theoretical reviews		x
Research methods		
Qualitative		x
Quantitative	x	
Mixed methods		x
Participants in the study		
Primary student	x	
Secondary student	x	
Undergraduate	x	
Master student		x
PhD student		x
Focus on the subject		
Science	x	
Biology	x	
Chemistry	x	
Physics	x	
Others		x

criteria set by Cohen (1988). The effect size is calculated using Hedges' g, which adjusts for bias in the standardized mean difference (SMD) between the two groups (Hedges et al., 1989). Using the right formula, Hedges' g provides a more accurate estimate of how much the IBL approach affects students' critical thinking skills compared to the conventional teaching method, as presented in **Table 3**.

Visualizing the analysis results using a forest plot (Gillette et al., 2018). Forest plot displays the effect size of each study along with its confidence interval (CI), providing a visual representation of the distribution and homogeneity of the study results (Li et al., 2020; Malapane et al., 2022). A heterogeneity test (Q-test) is performed using the I^2 statistic to measure the extent to which variation among study results is due to natural differences between studies (Çalik & Wiyarsi, 2024). The I^2 statistic is calculated by comparing the Q value and the degrees of freedom (df), and the I^2 percentage indicates the degree of heterogeneity present (Kim et al., 2020). Moderator analysis was conducted to explore variables

Table 3. Interval effect size (Cohen, 1988)

Interval effect size	Interpretation
Effect size ≤ 0.50	Small effect
0.50 < Effect size ≤ 0.80	Medium effect
0.80 < Effect size	Large effect

Table 4. Summary of selected studies

Characteristics	Value
Number of studies (k)	36
Number of observations (O)	3,932
Observations (experimental [OE])	2,093
Observations (control [OC])	1,839
SMD	1.2691
95% CI	[0.7813; 1.7568]
z	5.1
p-value	< 0.0001

that might influence the results of the meta-analysis, such as education level or study location (Zhang et al., 2024). These analyses provide further insight into the factors contributing to outcome variation across different contexts (Cromley et al., 2023; Raposo-Rivas et al., 2024). To estimate the between-study variance (τ^2), to estimate the between-study variance (r^2), we used the restricted maximum likelihood estimator, which is known for providing more accurate and less biased estimates (Kooren et al., 2024; Wen et al., 2015). Publication bias assessment ensures that the results are not influenced by selective publication of studies that only report significant results (Nakagawa et al., 2023). Funnel plots (graphical analysis) and Egger's test (statistical test) were used to identify potential of publication bias (Egger et al., 1997).

RESULTS

The 25 empirical studies in **Table A1** in **Appendix A** were broken down into 36 different study entries (shown in the in **Table B1** in **Appendix B**) because some studies provided more than one set of statistical data, such as SD, mean, and N values. This meant that a single study could provide multiple relevant effect sizes to be analyzed separately. For example, the study by Lu et al. (2020) provides four different values (marked as Lu et al., 2020a, 2020b, 2020c, 2020d) indicating that the study had several different sample groups or experimental conditions.

A study by Ku et al. (2014) was divided into three sub-studies (Ku et al., 2014a, 2014b, 2014c). Other studies, such as Duran and Dökme (2016) and Kim et al. (2012), have two and three sub-entries indicating differences in the data presented. This process of splitting studies is essential to capture variation within the same study, allowing for a more comprehensive analysis of the conditions and variables tested. As such, the total number of studies analyzed increased to 36 entries despite coming from 25 studies to ensure all relevant data were included in this meta-analysis.

Table 5. Examination of the heterogeneity of studies based on I^2 and Cochran's Q-statistic

Criteria	Value
Heterogeneity (τ^2)	2.1363 [1.5552; 4.8359]
Heterogeneity (τ)	1.4616 [1.2471; 2.1991]
I^2	92.0% [89.9%; 93.7%]
H	3.54 [3.15; 3.97]
Q-test	437.47
df	35
Test of heterogeneity p-value	< 0.0001

Table 4 describes significant findings related to the impact of the interventions tested in the 36 studies. The main effect size measured using the SMD was 1.2691. We assume that the SMD value mentioned is the effect size using Hedges' g. The value suggests that, on average, there is a sizable effect of the intervention using the Hedges' g model. The value explains that the intervention using the inquiry model has a considerable effect on average. The difference between the experimental and control groups is significant in the context studied. The 95% CI ranged from 0.7813 to 1.7568. Since all these intervals are above zero, this result indicates that the effect is statistically significant. The z-value of 5.10 with a p-value of less than 0.0001 confirms that these results are highly statistically significant, indicating a high probability that these results are not due to chance.

The heterogeneity analysis in **Table 5** explains the significant variation among the study outcomes. The tau-squared (r^2) measures the actual variance among study effects, which is 2.1363, indicating considerable between-study variability in effect size. Tau (r), the square root of r^2, was 1.4616, giving a more intuitive idea of the level of variation. The sizable value confirms the presence of significant variability among the study results. The Q-test showed a Q value of 437.47 with 35 df and a minimal p-value ($p < 0.0001$), confirming that there is statistically significant heterogeneity among the results of these studies, meaning that the differences between studies are not only due to random error but also other factors that have not been identified.

Another measure of heterogeneity, I^2, had a value of 92.0%. The proportion of total variability in the results is due to heterogeneity between studies rather than random error. With a value of 92%, almost all of the variability in these results is due to natural differences between studies rather than random error. In addition, the H value of 3.54, which is greater than 1, also indicates considerable heterogeneity among the studies. Further Q-testing using the Q statistic revealed a value of 437.47 with df of 35 and a p-value of less than 0.0001. These results, suggesting significant heterogeneity among the studies, corroborate previous findings of substantial variation in results. **Figure 2** presents a forest plot graph displaying the results of a meta-analysis on the effect of

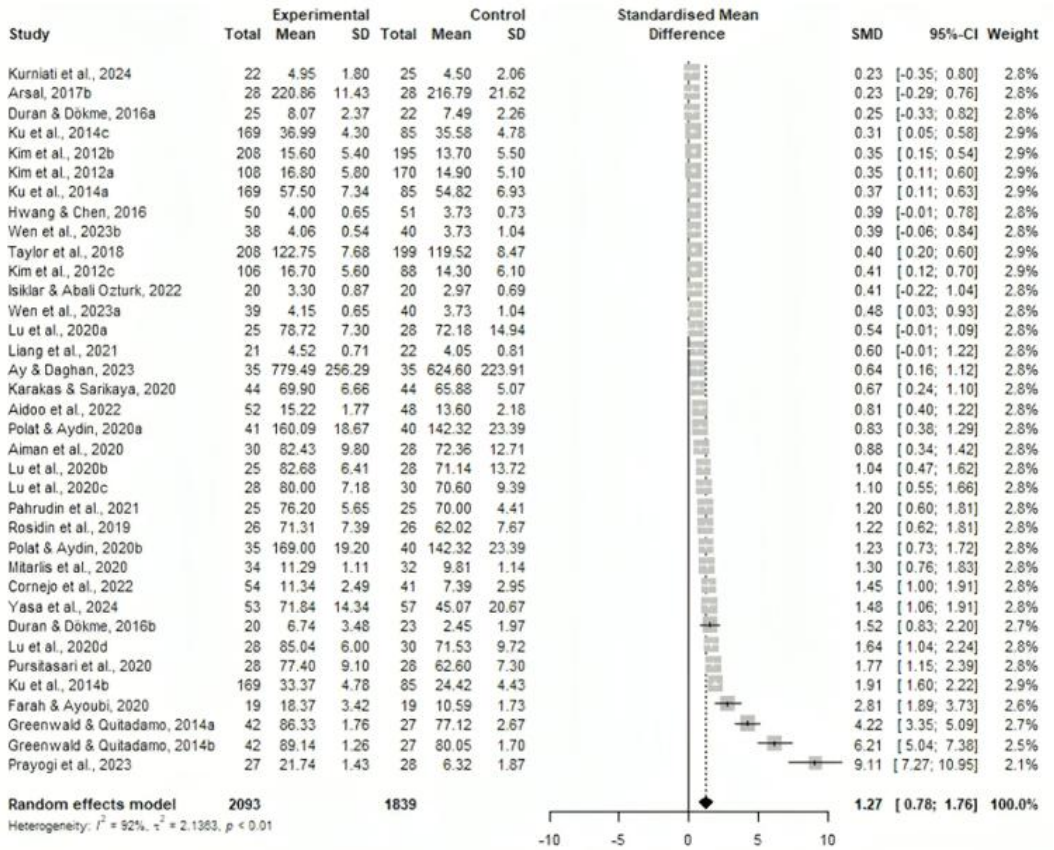

Figure 2. Effect sizes distribution and forest plot of studies about effect IBL on critical thinking skills (Source: Authors' own elaboration)

inquiry learning on students' critical thinking. The forest plot graph combines data from 36 studies involving various experimental and control groups to evaluate the effectiveness of the intervention.

The overall SMD value is 1.27 with a 95% CI between 0.78 and 1.76, indicating that, on average, there is a considerable positive effect of inquiry learning on students' critical thinking so that the effect is statistically significant as the CI does not include zero. The SMD value greater than one also indicates that the difference between the experimental and control groups is considerable in the context studied. The forest plot graph also showed high heterogeneity among the included studies, with an I^2 value of 92%, meaning that most of the variability in effect sizes between studies is due to fundamental differences rather than just random variation. High heterogeneity suggests significant variation in how the intervention was implemented or in study characteristics, which could affect the results (Öztürk et al., 2022; Yang et al., 2020).

The results of the forest plots indicated that the inquiry learning analyzed had a significant positive effect on students' critical thinking. However, there was substantial variability among the studies, and further analysis is needed to identify factors contributing to this variability and better understand how the inquiry learning model can be effectively implemented in various science teaching contexts.

Moderator Analysis

Moderation analysis was conducted to investigate the potential impact of various moderating variables on the effectiveness of IBL in improving critical thinking skills. Given the significant heterogeneity in the study results, as indicated by the heterogeneity value (τ^2) of 2.1363 and the I^2 value of 92.0%, it is important to examine how different factors may influence the relationship between IBL and critical thinking skills. The high I^2 value indicates that most of the variability in the results is due to natural differences between studies rather than random error. In addition, the H value of 3.54 and the Q-test result of 437.47 indicate highly significant heterogeneity among these studies. Therefore, variables such as education level, teaching strategy, country of study, duration of intervention, science discipline, ICT medium used, and specific assessment instruments were considered in this analysis to provide a deeper understanding of how contextual and methodological factors may affect the outcomes of IBL interventions. More detailed information can be seen in **Table C1** in **Appendix C**.

Education Levels

Students at the undergraduate level showed that IBL had a very strong influence on critical thinking skills with a SMD of 2.66 (N = 9). Students at the middle school level had an effect size of 1.47 (N = 7), which was greater than students at the high school level with a SMD of 0.81 (N = 5), explaining that the inquiry-based approach is also effective at a very early level of education. Meanwhile, students at the elementary school level had a moderate impact with a value of SMD = 0.57 (N = 12). When assessing the effect of students' education level on learning outcomes, the data shows that significant differences in effect sizes between education levels are different from the assumption that higher education levels will always show more significant effects.

Learning Strategy

The results of the analysis showed that the IBL strategy generally had a moderate impact on students' critical thinking skills with an SMD value of 0.65 ($p < 0.01$), indicating that the IBL approach as a whole was moderately effective in improving critical thinking skills, with effect sizes that could be considered moderate to large. However, when there is variation in the effectiveness of various models and approaches within the IBL framework. For example, the ethnoscience-based IBL learning strategy (ethno-inquiry) showed a massive effect with an SMD of 7.27 (N = 1) and the Inquiry-based clinical case with an effect size of 5.18 (N = 2). Some specific approaches in IBL, such as smart learning-based inquiry, IBL and reflection, and new IBL, also showed significant effect sizes with SMD of 1.41, 2.81, and 1.86, respectively. The strong effect can be attributed to students' more active approach and direct involvement in the learning process, which encourages critical and reflective thinking. The community of inquiry model and Inquiry-based ubiquitous gaming environment showed small effect sizes of 0.41 and 0.39, respectively, indicating more minor impact than the other IBL approaches. Nonetheless, these positive effects suggest that even approaches with more minor impacts can still be beneficial.

Level of Inquiry

Guided inquiry (N = 24) showed a significant positive effect with a mean value of 1.46, indicating that the result is statistically significant, signaling that the method may effectively achieve the desired goal. Open inquiry (N = 2) showed a moderately significant result with a mean of 1.41. Structured inquiry (N = 9) showed less significant results with a mean of 0.93 and a CI of 0.67 to 1.19. Since the CI includes a value of 1, the effect of this method is similar to zero, indicating the method may not have a clear impact. Confirmation inquiry (N = 1) explains the insignificant results with a mean of 0.41 and a CI from -0.22 to 1.04.

Research Across Countries

The analysis results by country showed variations in the effectiveness of the IBL strategy on students' critical

thinking skills. Results from Canada showed a significant effect size with the highest SMD of 5.18 (N = 2), indicating that the strategy significantly impacts the country's educational context. Similar results were also seen in Lebanon with an SMD of 2.81 (N = 1), consistently impacting students' critical thinking skills. Singapore and the United States showed large effect sizes with SMD of 1.45 (N = 1) and 1.43 (N = 4). Turkey, China, and Indonesia recorded moderate effect sizes with SMD of 0.71 (N = 8), 0.75 (4) and 0.59 (N = 8) . In contrast, in Chile, the effect is small, with a SMD of 0.49, indicating that although there is a positive impact, it is not as large as in other countries.

Science Disciplines

Analysis of the research results by discipline showed variations in the impact of IBL strategies on students' critical thinking skills. Overall, the studies that examined various science disciplines showed an SMD of 1.10 (N = 21). Although SMD indicates a moderately positive impact, the p-value of 0.51 indicates that the result is not statistically significant and could be due to heterogeneity in the approaches used or variations in the implementation of IBL strategies across different science sub-disciplines. The impact of IBL strategies on students' critical thinking skills shows considerable differences. Studies focusing on physics learning show a very strong effect size with a SMD of 2.27, indicating that IBL improves students' critical thinking skills in physics. In contrast, the physics discipline showed a lower effect size with a SMD of 0.95, indicating a more moderate impact of IBL strategies on students' critical thinking skills. Chemistry showed a higher effect size than biology, with an SMD of 1.58.

ICT Used

Studies that did not mention specific media showed a high effect size with an SMD of 1.28 and a 95% CI between 0.70 and 1.86. The p-value of 0.02 indicates that these results are statistically significant, indicating that the use of various media in IBL has a moderately strong positive impact on students' critical thinking skills. When looking more closely at the specific media, one can see significant differences in their effectiveness. For example, digital media such as digital games showed a significant impact with an SMD of 3.98 and 95% CI between 3.03 and 4.92, indicating that digital games have a very strong influence in improving students' critical thinking skills. Traditional media, such as books and printed materials, showed a more moderate effect size with a SMD of 1.02 and 95% CI between 0.53 and 1.51. While it still has a positive impact, more interactive digital media is needed. Other media, such as e-learning systems and mind mapping, also show moderate effect sizes with SMD of 0.89 and 1.02, respectively. These media are effective in certain situations, mainly when used to organize information or provide flexible access to learning materials. Technology-based media, such as computer simulation, augmented reality, and real-time media, show moderately high effect sizes.

Duration of Intervention

Analysis of the IBL duration showed that longer duration affects students' critical thinking skills substantially. The data showed that programs lasting 1-4 weeks had a reasonably low effect with a SMD of 0.23 and a 95% CI between 0.09 and 0.76, although these results were statistically significant ($p < 0.01$). However, once the duration reached four weeks, there was a significant increase in effectiveness, with an SMD of 1.29 and 95% CI between 0.75 and 1.83. The increase in effect size became more pronounced as the duration increased to 8 weeks, where the SMD reached 2.09 with a 95% CI between 0.66 and 3.53. Individual studies involving 6-week and 16-week durations showed SMD of 2.29 and 2.77, respectively, confirming that increasing the duration of learning significantly impacts improving critical thinking skills. The 10-week duration also showed a significant positive effect with an SMD of 1.48, suggesting an optimum point where longer duration starts to yield better results. However, the results vary when the duration reaches more than one year. For example, a 1-year duration only showed an SMD of 0.35 with a 95% CI between 0.10 to 0.60, while a 3-year duration had a much higher SMD of 1.97 with a 95% CI between 1.04 to 2.91.

Assessment Used

The instruments with the most significant effect sizes were the Ennis-Weir critical thinking essay test with a SMD of 5.39 with a 95% CI between 1.81 and 12.58 and the California critical thinking skills test with a SMD of 5.18 with a 95% CI between 2.32 and 7.13. These high effect sizes indicate that these instruments are highly effective in measuring improvements in critical thinking skills due to IBL interventions. Other instruments that showed high effect sizes were the James Madison critical thinking test with a SMD of 2.81 and the critical thinking skill essay test with a SMD of 1.24. Instruments with the lowest effect sizes include the critical thinking test (Miranda, 2003), with a SMD of 0.35 and 95% CI between 0.12 and 0.64.

Type of Assessment

The most frequently used assessment type, essay (N = 15), explained a strong effect size with a SMD of 1.95, and this result was statistically significant ($p < 0.01$). In contrast, rating scale-based assessments showed a lower effect size with a SMD of 0.7. Other assessments, such as closed-ended and multiple choices, showed exciting variations. Closed-ended questions had a sizable effect size with a SMD of 3.04. On the other hand, multiple-choice assessments have a SMD of 1.11 with a 95% CI

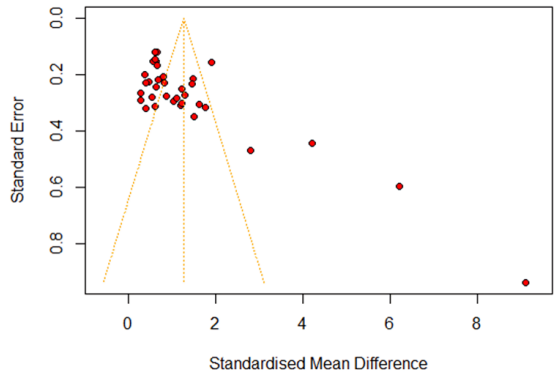

Figure 3. Funnel plot diagram of publication bias (Source: Authors' own elaboration)

between 0.65 and 1.72, which also shows a significant effect. The rating scale-based assessment showed an SMD of 0.70 with a 95% CI between 0.40 to 1.00, indicating that the rating scale was also effective in measuring the improvement of critical thinking skills with moderate effects. At the same time, the assessment type with the lowest effect is in the form of a questionnaire with a SMD of 0.39.

Publication Bias

Figure 3 shows no strong indication of publication bias in the meta-analysis (Egger et al., 1997). Publication bias occurs when studies with statistically significant results or larger effect sizes are more likely to be published than non-significant studies, thus creating biased results in the available literature (Demena, 2024). In this plot, the distribution of red dots, representing individual study results, is symmetrical around the central vertical line representing the standardized mean effect size. The symmetry in the plot shows that studies with both significant and non-significant results are evenly distributed, indicating that publications are not skewed to one side of the effect (Linden et al., 2024). Most points lie within the triangular boundary of symmetry formed by the dashed line, reflecting the normal distribution of study results, especially at the top of the graph where standard errors are lower (studies with larger N) (Sterne et al., 2011). Although there are some outliers with large effect sizes on the right side of the plot, their number and distribution pattern are not strong enough to suggest publication bias. Some of these outliers can be explained by methodological variations, different study characteristics, or the actual effects of the phenomenon under study rather than as a result of selection bias (Neethirajan et al., 2005). Based on the observations, there is no significant visual evidence of publication bias in the analyzed dataset, and the symmetric distribution of the study supports this result on the funnel plot.

DISCUSSION

General Features

Based on the collected research results, it is evident that the IBL approach, in its various variations, consistently improves students' critical thinking skills and academic achievement across different levels of education and science subject areas. This strategy is applied in various educational contexts worldwide, including countries in Europe, America, Africa and Asia (Ramnarain, 2023; Strat et al., 2024). The results show that the implementation of IBL is effective in promoting critical thinking, especially when combined with other learning elements such as technology (Alabidi et al., 2023; Pahrudin et al., 2021; Wen et al., 2023), argumentation (Kabataş Memiş & Çakan Akkaş, 2020), and local cultural context (Prayogi et al., 2022).

Previous research has shown that an IBL model can improve students' critical thinking skills. A research conducted by Lu et al. (2020) and Pursitasari et al. (2020) corroborate these findings by showing that learning models such as Critique-Driven Inquiry and Science Context-Based Inquiry Learning are not only effective in improving critical thinking skills but also enrich the learning process by involving students more actively. They introduced Smart Learning-Based Inquiry, which improves critical thinking skills and digital literacy, expanding the scope and impact of this inquiry-based approach.

Although numerous studies indicate substantial enhancements in critical thinking abilities, certain research, including Arsal (2017), revealed that not all IBL models produced significant alterations, implying that the efficacy of these methods is contingent upon the context of implementation and supplementary learning strategies employed. Learning methodologies that amalgamate inquiry-based approaches with argumentation or philosophy, as demonstrated in Karakas and Sarikaya (2020), yielded varied outcomes, suggesting potential for optimizing this strategy through the adjustment and integration of more diverse methods.

The Overall Effect of Inquiry-based Learning on Critical Thinking

The research findings show that implementing IBL in an educational setting significantly positively impacts the development of critical thinking skills, with a SMD 1.27. The conclusions of Cornejo (2022) highlighted that the inquiry-based approach effectively promotes critical thinking skills in students. Using IBL techniques associated with improving critical thinking skills compared to traditional teaching methods. The interactive nature of IBL encourages students' active participation in the problem-solving and investigation process, enhances cognitive engagement and promotes deeper learning as highlighted by Prayogi et al. (2023

The hands-on learning experience inherent in IBL allows students to simulate real-world scientific inquiry to practically develop critical thinking skills (Karakas & Sarikaya, 2020). Further analysis showed a statistically significant positive effect on critical thinking skills (SMD = 1.45), in line with previous research findings by Lu et al. (2020), who found that the IBL approach significantly improved students' analytical and scientific reasoning skills. Similar conclusions were also drawn by Duran and Dökme (2016), who noted that students exposed to inquiry-based science activities showed a marked improvement in critical thinking performance compared to those in a traditional learning environment.

The IBL approach, with its student-centered learning environment, encourages deeper engagement, motivation, and understanding of the subject matter, ultimately leading to more positive perceptions and attitudes toward critical thinking skills (Ku et al., 2014; Ay & Daghan, 2023). Despite some challenges, such as the need for structured guidance to prevent cognitive overload, overall, using IBL in education yields satisfactory results. Therefore, educators should consider integrating IBL into various educational contexts to enhance critical thinking skills and support the development of independent and analytical learners who are prepared for complex problem-solving in the real world (Wan et al., 2024).

Studies by Kurniati et al. (2024) and Arsal (2017) demonstrate lower or negligible effects, with SMD values approaching zero, suggesting that IBL does not consistently yield considerable enhancement. The brief duration of the intervention, deficiencies in execution, or discrepancies in student participation levels. These findings suggest that although IBL is typically beneficial, its efficacy is mostly contingent upon the manner of implementation, including the intervention duration, teacher training, integration with other educational approaches, and adaptation to the specific situation. Consequently, forthcoming research must ascertain the ideal conditions for executing IBL to optimize its advantages.

Moderator Analysis

Based on the results of moderator analysis, the effectiveness of IBL on critical thinking skills is influenced by various factors, including education level, learning strategy, inquiry level, country, intervention duration, science field, ICT media, and assessment instruments used.

Education level plays an essential role as a moderator in determining the effectiveness of IBL intervention. At the undergraduate level, the effect of the IBL intervention is powerful (SMD = 2.66) because students have benefitted more from complex learning methods (Cornejo et al., 2022). The ability of students to engage in a learning process that demands critical and analytical thinking supports the effectiveness of this intervention, thus improving their learning outcomes (Jeon et al., 2021; Prayogi et al., 2022).

Although the intervention still provided benefits at the elementary school level, the effects were more moderate compared to the undergraduate group. This result could be attributed to the more exploratory and play-focused nature of learning in primary schools that support student engagement in the inquiry process but with different levels of depth (Duran & Dökme, 2016; Yasa et al., 2024). Meanwhile, at the secondary school level, intervention effects were also cheerful and moderately strong, but there was more significant variability in the results. They indicate variations in the execution of the intervention or the distinct attributes of pupils at this level. Although the intervention had a positive effect at the senior high school level, the strength was smaller compared to the lower level of education (preschool). It could be due to variations in students' cognitive development and motivation at this level, which affect how they respond to the intervention (Polat & Aydın, 2020). The intervention effect is large at the preschool level, indicating that a more interactive and experiential learning approach is appropriate for this age group. A research study by Iskilarak and Abali Öztürk (2022) explains that children in this program exhibit various characteristics, including multi-dimensional thinking, the ability to view situations from multiple perspectives, and the establishment of cause-and-effect relationships. These features are believed to enhance the development of their problem-solving skills.

The IBL strategy significantly impacted students' critical thinking skills, with effects varying between different approaches. Overall, IBL effectively improved critical thinking skills, with effect sizes ranging from moderate to large. Approaches incorporating advanced technologies such as smart learning-based learning (Yasa et al., 2024) and inquiry-based game environments (Liang et al., 2021; Wardani et al., 2017) showed an enormous effect. This high effectiveness can be attributed to the interactive and personalized nature of technology-based learning that enhances student engagement and encourages critical and reflective thinking (Spector & Ma, 2019).

However, not all IBL interventions have the same impact. Some interventions, such as the community of inquiry model and inquiry-based ubiquitous gaming environment, showed smaller effect sizes. Although the impact is positive, this difference suggests that the level of student engagement and the complexity of the task given can influence the effectiveness of learning in improving critical thinking skills (Kaczkó & Ostendorf, 2023). Approaches that require direct engagement and more challenging tasks appear more effective than passive or structured approaches (Ay & Dağhan, 2023; Martin et al., 2022).

The different levels of IBL shows that the level of student engagement and the structure applied in the inquiry method influence the effectiveness of this intervention towards improving critical thinking skills . Guided inquiry and open inquiry showed a significant and large impact on student's critical thinking skills, explaining that when students are given proper guidance or complete freedom to explore and investigate, their critical thinking skills can be better developed (van Uum et al., 2016; Wulandari et al., 2022). This intervention encourages students to actively engage in the learning process, ask questions, and seek answers through experimentation or investigation, which is essential to developing critical thinking skills (Butcher et al., 2023).

Conversely, structured inquiry also showed a significant but slightly lower effect compared to guided inquiry and open inquiry, explaining that although providing a framework and structure in the inquiry process helps students understand the steps that need to be taken, more limited freedom may reduce the opportunity for students to develop deeper critical thinking. Structured inquiry remains effective in contexts where students may need more support or direction to initiate the student learning process (Pahrudin et al., 2021).

Confirmation inquiry showed little effect on students' critical thinking skills. Research by Toma (2022) explained that there was no difference in students' expectations of success and the intrinsic value of school science when lecture, confirmation, or structured inquiry teaching strategies were used. It could be due to the nature of these methods, which focus more on confirming or verifying already known information rather than encouraging new exploration and discovery. Students tend to be more passive as they follow a predetermined procedure without thinking of new ways or solutions and suggests that a more active and participatory approach is needed to improve critical thinking skills.

In Canada, Lebanon, the United States, and Singapore, IBL has significantly impacted critical thinking skills. High effectiveness indicates that IBL can be very successful in contexts where students have access to adequate resources and strong support from their education system. However, in other countries such as Turkey, Indonesia, China, and Ghana, the effects of IBL were more moderate, suggesting that while IBL still provides positive benefits, other factors influence how effectively the method is implemented. Factors such as differences in pedagogical approaches, the quality of teacher training, or even cultural differences in learning can affect outcomes. Meanwhile, IBL showed a small effect in Chile due to implementation limitations or insufficient curriculum adaptation to support IBL effectively (Vergara-Díaz et al., 2020).

The length of time spent implementing this method affects its effectiveness in improving students' critical thinking skills. Interventions with shorter durations, such as a few weeks or months, often showed large effects, explaining that IBL can provide significant benefits even over a relatively short period of time, especially when the learning strategies are well-designed and intensively implemented. Interventions that last for years or those longer than one year show smaller effects, suggesting that while IBL can improve critical thinking skills in the long term, there is a risk of decreased effectiveness over time. Factors such as student boredom, decreased motivation, or lack of variety in teaching methods may influence lower results at longer durations (Marcos-Vílchez et al., 2024). Long-term interventions do not maintain the intensity and focus essential to maximize the effects of IBL (Shen et al., 2024).

IBL is generally effective in science disciplines, with results showing significant improvements in critical thinking skills. The effects of IBL are powerful in the field of biology. It could be due to the nature of biology, which often involves hands-on experiments and research, allowing students to more easily engage in the process of inquiry and development of critical thinking (Qu et al., 2024; Vančugovienė et al., 2024). Similarly, IBL has also shown significant effects in chemistry, reflecting how this method can enhance the understanding of concepts through in-depth laboratory activities and experiments (Strippel & Sommer, 2015). Although IBL showed a significant effect in physics, its slightly lower impact compared to biology and chemistry could be due to the complexity of the more abstract physics concepts that require additional learning strategies to ensure students can fully utilize the inquiry method (Agustini & Suyatna, 2018).

The use of ICT media in IBL has a diverse impact on improving students' critical thinking skills. Overall, ICT-based media are effective, with some media showing a tremendous impact (Chen & Chen, 2024). E-learning systems and digital games provided positive results, suggesting that digital technologies allowing in-depth interaction and personalized learning can significantly enhance critical thinking skills (Kaczkó & Ostendorf, 2023; Liang et al., 2021). The interactivity and flexibility offered by these digital platforms seem to encourage students to be more actively involved in the learning process and explore concepts in greater depth (Zheng et al., 2018). Wen et al. (2023) explained that integrating IBL and AR proved more beneficial for students with low critical thinking skills but had little impact on students with high critical thinking skills. Media such as books and printed materials also showed significant effects, although not as strong as digital media, suggesting that traditional and concrete media support IBL. These media do not have the same interactive features that can enhance engagement and critical thinking as in

technology-based media. Meanwhile, alternate and augmented reality games had less impact on critical thinking skills.

Various assessment instruments used to measure critical thinking skills show significant differences in the effectiveness of IBL. Such as the California critical thinking dispositions inventory (SMD = 0.9), California critical thinking skills test (SMD = 5.18), and Ennis-Weir critical thinking essay test (SMD = 5.39), which showed a considerable impact so indicates that these instruments are highly effective in measuring the improvement in critical thinking skills generated by the IBL intervention (Hix & Ramsey, 2023). These instruments cover various aspects of critical thinking, such as analysis, evaluation, and logical reasoning and can capture the various dimensions of critical thinking skills developed through IBL. On the other hand, some instruments, such as the Watson Glaser critical thinking appraisal (SMD = 0.47) and the Iowa test of basic skills (SMD = 0.4), showed smaller impacts on critical thinking skills. These results suggest that these instruments can still detect improvements in critical thinking skills less sensitive to the more subtle changes that occur due to IBL (Karakaş & Sarıkaya, 2020). Instruments with small impacts are often more focused on a particular aspect of critical thinking skills or use a more structured assessment format and may not fully capture the complexity and variety of critical thinking that can be developed through IBL methods (Butler, 2024).

Implications

The effectiveness of IBL depends not only on the method used but also on the educational level of the students. The unique characteristics of each educational level, such as cognitive readiness, motivation, and learning style, all play a role in determining how effective the intervention is. Therefore, educators and policymakers need to consider these factors when designing and implementing educational interventions, ensuring that learning strategies are tailored to the specific needs of students at each educational level. Yun and Crippen (2024) identified three important abilities to support students in science IBL. First, it assists students in capturing and collecting data; second, it provides relevant learning materials; and third, it connects learning with real-world contexts.

There have been significant successes in some countries; further research is needed to understand the barriers that may exist in countries where IBL has shown more moderate or modest results. One possible solution to address these gaps is to develop IBL implementation guidelines tailored to local needs and characteristics. Further research is also needed to explore how teacher training, infrastructure support, and curriculum adaptation can increase the effectiveness of IBL in different contexts, ensuring that all students, regardless of geographical location, can benefit the most from this learning approach.

Further research is required to comprehend the obstacles that may be present in countries where IBL has yielded more moderate or modest outcomes. A potential option to rectify these deficiencies is to establish IBL implementation guidelines customized to local requirements and attributes. Additional study is required to examine how teacher training, infrastructural support, and curriculum adaptation can enhance the effectiveness of IBL in all contexts, ensuring that all students, irrespective of geographical location, can optimally benefit from this educational method.

Certain research indicate substantial short-term advantages; nevertheless, modest long-term outcomes imply the necessity for alternative measures to sustain student engagement and learning efficacy (Pattipeilohy et al., 2022). Additional study is required to investigate how long-term treatments might be restructured to sustain their efficacy, including more adaptable methodologies or the integration of other learning mechanisms. Potential solutions involve creating curriculum that prioritize diverse learning methodologies and offering continuous support for students and educators during extended interventions to guarantee the sustainability of IBL advantages throughout the duration of the intervention. Educators must also contemplate two essential aspects for cultivating subject matter expertise: the allocation of time for science instruction and the establishment of a learning environment that prioritizes science (Nixon & Bennion, 2024).

Although some ICT media show great potential, there is a need for further research on how ICT can be used more effectively in various educational contexts. Potential solutions to address this gap include further training for teachers in the use of new technologies, as well as more in-depth research on how to better integrate technologies such as augmented reality, virtual reality, alternate reality games and artificial intelligence into existing curricula (Wang et al., 2024). By understanding the best ways to utilize various learning media, educators can improve students' critical thinking skills more consistently and effectively across various learning environments (Yeoh et al., 2024).

Variations in the effectiveness of assessment suggest gaps in understanding how best to measure critical thinking skills in the context of IBL. While some instruments effectively measure improvements in these skills, others may require customization or further development to be more sensitive to the multiple dimensions of critical thinking that develop through IBL. Combining IBL with customized assessment practices requires significant changes in teaching practices (Grangeat et al., 2021). Recent research could focus on developing more holistic and diverse assessment

instruments that can more comprehensively measure improvements in critical thinking skills (Cheung et al., 2023; Walton & Martin, 2023). Collaboration between researchers and teachers is essential to support changes in professional knowledge, beliefs and classroom practices.

CONCLUSION

Overall, this meta-analysis shows that IBL significantly improves critical thinking skills across different educational contexts. It is indicated by a high SMD of 1.2691 and a CI of [0.7813; 1.7568]. This effect was found to be robust with a highly significant p-value ($p < 0.0001$), as well as high heterogeneity ($I^2 = 92\%$), indicating considerable variability in effect size among the studies analyzed. Heterogeneity analysis confirmed substantial variation among the reviewed studies ($\tau^2 = 2.1363$), prompting further exploration of moderating factors that influence the effectiveness of IBL. Moderator analysis revealed several key variables affecting the success of IBL, including education level, inquiry strategy, level of inquiry, geographical context, duration of intervention, discipline, ICT media, and assessment tools.

The most effective levels of IBL are undergraduate, junior high, and senior high school. Additionally, significant positive effects have been observed at the early childhood education level, suggesting that inquiry can be tailored to a variety of cognitive levels. Nonetheless, its influence at the elementary school level is generally diminished owing to variations in cognitive development and students' preparedness for intricate learning methodologies. Diverse inquiry methodologies produced disparate outcomes, with clinical case-based inquiry and ethno-inquiry demonstrating notably high effect sizes, whereas confirmatory inquiry, which emphasizes information verification, exhibited diminished influence. The efficacy of IBL differs by country, with the most significant effects noted in nations such as Canada, Singapore, and the United States, underscoring the role of cultural and educational contexts on its adoption and execution. Moreover, extended interventions typically yielded superior results, however short-term therapies also demonstrated favorable outcomes, reflecting the adaptability of IBL. IBL proved to be most effective in science and biology, and its integration with ICT media, such as e-learning systems and digital games, enhanced learning outcomes, though traditional media also had significant effects. The effectiveness of IBL in enhancing critical thinking skills was demonstrated through many assessment instruments, with the California critical thinking skills test and the Ennis-Weir critical thinking essay test exhibiting substantial effect sizes, so affirming its relevance across varied evaluative settings.

Limitations and Future Studies

This study, while providing valuable insights into the impact of IBL on critical thinking skills, has several limitations. First, the literature search was restricted to articles from Scopus and ERIC, potentially excluding relevant studies from other databases, which could limit the comprehensiveness of the analysis. Second, the review only considered studies with pre-experimental or quasi-experimental designs, excluding other methodological approaches that may offer different perspectives. Third, despite the meta-analysis showing a high mean effect size, the substantial heterogeneity ($I^2 = 92.0\%$) among studies indicates variability in the effects of IBL across different contexts, which may limit the generalizability of the findings. The study's reliance on secondary data also constrains the control over variables, making it difficult to account for all moderating factors. Additionally, while subgroup analysis was conducted to explore potential moderators, other factors such as cultural influences or teacher characteristics may also play a significant role in the effectiveness of IBL but were not examined in detail. Lastly, the inclusion of only quantitative research may overlook the depth and insights that qualitative studies could provide on how IBL influences critical thinking skills in various educational settings.

Author contributions: **ZA:** conceptualization, data curation, resources, and writing–original draft; **S:** methodology, writing–review & editing, supervision, and validation; & **SS & AK:** writing–review & editing, supervision, and validation. All authors agreed with the results and conclusions.

Funding: This research was supported by Lembaga Pengelola Dana Pendidikan (LPDP) under the Indonesian Ministry of Finance with registration number 0001086/TRP/D/19/LPDP2023.

Ethical statement: The authors stated that the study does not require any ethical approval. It is a review of existing literature.

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

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Agustini, R., & Suyatna, A. (2018). Developing inquiry-based practice equipment of heat conductivity to foster the students' critical thinking ability. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni, 7*(1), 49-57. https://doi.org/10.24042/jipfalbiruni.v7i1.1848

Alabidi, S., Alarabi, K., Tairab, H., Alamassi, S., & Alsalhi, N. R. (2023). The effect of computer simulations on students' conceptual and procedural understanding of Newton's second law of motion. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(5), Article em2259. https://doi.org/10.29333/ejmste/13140

Ananda, P. N., & Usmeldi, U. (2023). Meta-analysis: effect of using inquiry based learning (IBL) model

on students' competence. *Jurnal Pendidikan Fisika, 11*(1), 86-102. https://doi.org/10.26618/jpf.v11i1.9802

Antonio, R. P., & Prudente, M. S. (2023). Effects of inquiry-based approaches on students' higher-order thinking skills in science: A Meta-Analysis. *International Journal of Education in Mathematics, Science and Technology, 12*(1), 251-281. https://doi.org/10.46328/ijemst.3216

Ay, K., & Dağhan, G. (2023). The effect of the flipped learning approach designed with community of inquiry model to the development of students' critical thinking strategies and social, teaching and cognitive presences. *Education and Information Technologies, 28*(11), 15267-15299. https://doi.org/10.1007/s10639-023-11809-2

Bansal, G. (2021). Indian pre-service teachers' conceptualisations and enactment of inquiry-based science education. *Education 3-13, 49*(3), 275-287. https://doi.org/10.1080/03004279.2020.1854957

Barta, A., Fodor, L. A., Tamas, B., & Szamoskozi, I. (2022). The development of students critical thinking abilities and dispositions through the concept mapping learning method–A meta-analysis. *Educational Research Review, 37*(17), Article 100481. https://doi.org/10.1016/j.edurev.2022.100481

Boonsathirakul, J., & Kerdsomboon, C. (2023). Synthesis of critical thinking research of basic education level students using meta-analysis in Thailand during 2010 to 2021. *Educational Research and Reviews, 18*(1), 1-8. https://doi.org/10.5897/ERR2022.4287

Butcher, K. R., Hudson, M., Lane, M., & Larson, M. (2023). Critical thinking during science investigations: What do practicing teachers value and observe? *Teachers and Teaching, 29*(6), 594-614. https://doi.org/10.1080/13540602.2023.2191186

Butler, H. A. (2024). Predicting everyday critical thinking: A review of critical thinking assessments. *Journal of Intelligence, 12*(2), Article 16. https://doi.org/10.3390/jintelligence12020016

Çalik, M., & Wiyarsi, A. (2024). The effect of socio-scientific issues-based intervention studies on scientific literacy: A meta-analysis study. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2024.2325382

Chaudhary, P., & Singh, R. K. (2022). A meta analysis of factors affecting teaching and student learning in higher education. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.824504

Chen, F., & Chen, G. (2024). Technology-enhanced collaborative inquiry in K-12 classrooms: A systematic review of empirical studies. *Science and Education*. https://doi.org/10.1007/s11191-024-00538-8

Cheung, K. K. C., Pun, J. K. H., & Fu, X. (2023). Development and validation of a reading in science holistic assessment (RISHA): A Rasch measurement study. *International Journal of Science and Mathematics Education, 22*, 299-320. https://doi.org/10.1007/s10763-023-10434-2

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates Inc. https://doi.org/10.4324/9780203771587

Cornejo, C. O., Luengo, M. P., Mujica, A. D., & Rovira, D. P. (2022). Promoting pedagogy students' scientific skills through critical thinking program. *Electronic Journal of Research in Educational Psychology, 20*(57), 267-290. https://doi.org/10.25115/ejrep.v20i57.5272

Cromley, J. G., Chen, R., & Lawrence, L. (2023). Meta-analysis of STEM learning using virtual reality: Benefits across the board. *Journal of Science Education and Technology, 32*(3), 355-364. https://doi.org/10.1007/s10956-023-10032-5

Demena, B. A. (2024). Publication bias in export promotion impact on export market entry: Evidence from a meta-regression analysis. *Applied Economics Letters*. https://doi.org/10.1080/13504851.2024.2306185

Duran, M., & Dökme, I. (2016). The effect of the inquiry-based learning approach on student's critical-thinking skills. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(12), 2887-2908. https://doi.org/10.12973/eurasia.2016.02311a

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Enugu, R., & Hokayem, H. (2017). Challenges pre-service teachers face when implementing a 5E inquiry model of instruction. *European Journal of Science and Mathematics Education, 5*(2), 178-209. https://doi.org/10.30935/scimath/9506

Essalih, S., Ourahay, M., & Khzami, S. (2024). Perceptions and practices of science teachers for the development of critical thinking through inquiry: A qualitative study of Moroccan primary schools. *Education 3-13, 52*(5), 662-677. https://doi.org/10.1080/03004279.2022.2116942

Geng, X., & Su, Y. S. (2024). Enhancing K-12 students' STEM learning through the integration of the metaverse into online and blended environments: A meta-analysis. In *International Journal of Science and Mathematics Education, 22*, 111-143. https://doi.org/10.1007/s10763-024-10484-0

Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., & Broedel-Zaugg, K. (2018). A meta-analysis of outcomes comparing flipped classroom and lecture. *American Journal of Pharmaceutical*

Education, 82(5), Article 6898. https://doi.org/10.5688/ajpe6898

Gómez, R. L., & Suárez, A. M. (2020). Do inquiry-based teaching and school climate influence science achievement and critical thinking? Evidence from PISA 2015. *International Journal of STEM Education, 7*, Article 43. https://doi.org/10.1186/s40594-020-00240-5

Grangeat, M., Harrison, C., & Dolin, J. (2021). Exploring assessment in STEM inquiry learning classrooms. *International Journal of Science Education, 43*(3), 345-361. https://doi.org/10.1080/09500693.2021.1903617

Hansen, C., Steinmetz, H., & Block, J. (2022). How to conduct a meta-analysis in eight steps: A practical guide. *Management Review Quarterly, 72*, 1-19. https://doi.org/10.1007/s11301-021-00247-4

Harefa, E., Sirait, M., & M. Ginting, E. (2016). Effect of inquiry learning model training and critical thinking skills on scientific knowledge class X. *Jurnal Pendidikan Fisika, 5*(2), Article 81. https://doi.org/10.22611/jpf.v5i2.4403

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R*. Chapman and Hall/CRC. https://doi.org/10.1201/9781003107347

Hedges, L., Shymansky, S., & Woodworth, G. (1989). *A practical guide to modern methods of meta-analysis*. National Science Teachers Association.

Hix, M., & Ramsey, S. J. (2023). Thinking again about critical thinking. *American Journal of Pharmaceutical Education, 87*(8), Article 100406. https://doi.org/10.1016/j.ajpe.2023.100406

Inel-Ekici, D., & Ekici, M. (2022). Mobile inquiry and inquiry-based science learning in higher education: Advantages, challenges, and attitudes. *Asia Pacific Education Review, 23*(3), 427-444. https://doi.org/10.1007/s12564-021-09729-2

Iskilarak, S., & Abali Öztürk, Y. (2022). The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills. *International Journal of Contemporary Educational Research, 9*(1), 130-142. https://doi.org/10.33200/ijcer.942575

Jegstad, K. M. (2023). Inquiry-based chemistry education: A systematic review. *Studies in Science Education, 60*(2), 251-314. https://doi.org/10.1080/03057267.2023.2248436

Jeon, A.-J., Kellogg, D., Khan, M. A., & Tucker-Kellogg, G. (2021). Developing critical thinking in STEM education through inquiry-based writing in the laboratory classroom. *Biochemistry and Molecular Biology Education, 49*(1), 140-150. https://doi.org/10.1002/bmb.21414

Kabataş Memiş, E., & Çakan Akkaş, B. N. (2020). Developing critical thinking skills in the thinking-discussion-writing cycle: The argumentation-based inquiry approach. *Asia Pacific Education Review, 21*(3), 441-453. https://doi.org/10.1007/s12564-020-09635-z

Kaczkó, É., & Ostendorf, A. (2023). Critical thinking in the community of inquiry framework: An analysis of the theoretical model and cognitive presence coding schemes. *Computers & Education, 193*, Article 104662. https://doi.org/10.1016/j.compedu.2022.104662

Karakaş, H., & Sarıkaya, R. (2020). The effect of argumentation-based teaching performed for environment-energy issues on critical thinking of prospective classroom teachers. *İlköğretim Online, 19*(1), 366-383. https://doi.org/10.17051/ilkonline.2020.661855

Kilbane, J. F., & Clayton, C. D. (2023). Conceptualising inquiry outcomes: A taxonomy to characterise the work of teacher inquirers. *Teachers and Teaching, 29*(3), 276-290. https://doi.org/10.1080/13540602.2023.2172395

Kim, K. H., VanTassel-Baska, J., Bracken, B. A., Feng, A., Stambaugh, T., & Bland, L. (2012). Project clarion: Three years of science instruction in title I schools among K-third grade students. *Research in Science Education, 42*(5), 813-829. https://doi.org/10.1007/s11165-011-9218-5

Kim, N. J., Belland, B. R., Lefler, M., Andreasen, L., Walker, A., & Axelrod, D. (2020). Computer-based scaffolding targeting individual versus groups in problem-centered instruction for STEM education: Meta-analysis. *Educational Psychology Review, 32*(2), 415-461. https://doi.org/10.1007/s10648-019-09502-3

Kolaski, K., Logan, L. R., & Ioannidis, J. P. A. (2023). Guidance to best tools and practices for systematic reviews. *Systematic Reviews, 12*, Article 96. https://doi.org/10.1186/s13643-023-02255-9

Kooren, N. S., Van Nooijen, C., & Paas, F. (2024). The influence of active and passive procrastination on academic performance: A meta-analysis. *Education Sciences, 14*(3), Article 323. https://doi.org/10.3390/educsci14030323

Ku, K. Y. L., Ho, I. T., Hau, K.-T., & Lai, E. C. M. (2014). Integrating direct and inquiry-based instruction in the teaching of critical thinking: An intervention study. *Instructional Science, 42*(2), 251-269. https://doi.org/10.1007/s11251-013-9279-0

Le Thi Tuyet, T., Nguyen Thi, K., Tran Duc, H., Tran Thi Giao, X., Vo Thi, N., Huynh Kim Tuong, V., Do Thi Nhu, U., & Le Thi Mai, A. (2024). STEM education in primary schools of southeast Asian countries: An analysis of scientific publications in the Scopus

database from 2000 to 2022. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(4), Article em2433. https://doi.org/10.29333/ejmste/14432

Li, B.-Z., Cao, N.-W., Ren, C.-X., Chu, X.-J., Zhou, H.-Y., & Guo, B. (2020). Flipped classroom improves nursing students' theoretical learning in China: A meta-analysis. *PLoS ONE, 15*(8), Article e0237926. https://doi.org/10.1371/journal.pone.0237926

Liang, H., Hsu, T., & Hwang, G. (2021). Promoting children's inquiry performances in alternate reality games: A mobile concept mapping-based questioning approach. *British Journal of Educational Technology, 52*(5), 2000-2019. https://doi.org/10.1111/bjet.13095

Linden, A. H., Pollet, T. V., & Hönekopp, J. (2024). Publication bias in psychology: A closer look at the correlation between sample size and effect size. *PLoS ONE, 19*(2), Article e0297075. https://doi.org/10.1371/journal.pone.0297075

Lu, Y., Lin, H., Smith, T. J., & Hong, Z. (2020). The effects of critique-driven inquiry intervention on students' critical thinking and scientific inquiry competence. *Journal of Baltic Science Education, 19*(1996), 954-971. https://doi.org/10.33225/jbse/20.19.954

Malapane, O. L., Musakwa, W., Chanza, N., & Radinger-Peer, V. (2022). Bibliometric analysis and systematic review of indigenous knowledge from a comparative African perspective: 1990-2020. *Land, 11*(8), Article 1167. https://doi.org/10.3390/land11081167

Marcos-Vílchez, J. M., Sánchez-Martín, M., & Muñiz-Velázquez, J. A. (2024). Effectiveness of training actions aimed at improving critical thinking in the face of disinformation: A systematic review protocol. *Thinking Skills and Creativity, 51*, Article 101474. https://doi.org/10.1016/j.tsc.2024.101474

Martin, F., Wu, T., Wan, L., & Xie, K. (2022). A meta-analysis on the community of inquiry presences and learning outcomes in online and blended learning environments. *Online Learning Journal, 26*(1), 325-359. https://doi.org/10.24059/olj.v26i1.2604

Nakagawa, S., Yang, Y., Macartney, E. L., Spake, R., & Lagisz, M. (2023). Quantitative evidence synthesis: A practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence, 12*, Article 8. https://doi.org/10.1186/s13750-023-00301-6

Nawanidbumrung, W., Samiphak, S., & Inoue, N. (2022). The impact of pre-service teachers' pedagogical beliefs on teaching science as inquiry: A silent antagonist for effective inquiry-based science lessons. *Science Education International, 33*(1), 112-121. https://doi.org/10.33828/sei.v33.i1.12

Neethirajan, S., Tuteja, S. K., Huang, S.-T., & Kelton, D. (2005). Publication bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Biosensors and bioelectronics* (Vol. 98). Wiley. https://doi.org/10.1002/0470870168

Nixon, R. S., & Bennion, A. (2024). Reflection time and valuing science: Elementary teachers' science subject matter knowledge development during teaching experience. *Science Education, 109*(1), 59-81. https://doi.org/10.1002/sce.21902

Öztürk, B., Kaya, M., & Demir, M. (2022). Does inquiry-based learning model improve learning outcomes A second-order meta-analysis. *Journal of Pedagogical Research, 6*(4), 201-216. https://doi.org/10.33902/JPR.202217481

Pahrudin, A., Misbah, Alisia, G., Saregar, A., Asyhari, A., Anugrah, A., & Susilowati, N. E. (2021). The effectiveness of science, technology, engineering, and mathematics-inquiry learning for 15-16 years old students based on K-13 Indonesian curriculum: The impact on the critical thinking skills. *European Journal of Educational Research, 10*(2), 681-692. https://doi.org/10.12973/eu-jer.10.2.681

Paige, D., Rupley, W. H., & Ziglari, L. (2024). Critical thinking in reading comprehension: Fine tuning the simple view of reading. *Education Sciences, 14*(3), Article 225. https://doi.org/10.3390/educsci14030225

Pattipeilohy, M., Rumahlatu, D., Salmanu, S. I. A., & Sangur, K. (2022). The inquiry investigation group learning model: Improving students' critical thinking skills, cognitive learning outcomes, and scientific attitudes. *Jurnal Pendidikan Biologi Indonesia, 8*(3), 205-215. https://doi.org/10.22219/jpbi.v8i3.22113

Polat, Ö., & Aydın, E. (2020). The effect of mind mapping on young children's critical thinking skills. *Thinking Skills and Creativity, 38*, Article 100743. https://doi.org/10.1016/j.tsc.2020.100743

Praminingsih, I., Miarsyah, M., & Rusdi, R. (2022). Meta-analysis: The effect of inquiry learning model on students critical thinking skill. *Journal of Biology Education, 5*(2), Article 168. https://doi.org/10.21043/jobe.v5i2.12940

Prayogi, S., Ahzan, S., & Rokhmat, J. (2022). Opportunities to stimulate the critical thinking performance of preservice science teachers through the ethno-inquiry model in an e-learning platform. *International Journal of Learning, Teaching and Educational Research, 21*(9), 134-153. https://doi.org/10.26803/ijlter.21.9.8

Pursitasari, I. D., Suhardi, E., Putra, A. P., & Rachman, I. (2020). Enhancement of student's critical thinking

skill through science context-based inquiry learning. *Jurnal Pendidikan IPA Indonesia, 9*(1), 97-105. https://doi.org/10.15294/jpii.v9i1.21884

Qu, M., Hou, Q., Yu, C., Li, X., Xia, J., & Dong, Z. (2024). Application and evaluation of the hybrid "problem-based learning" model based on "rain classroom" in experimental courses of medical molecular biology. *Frontiers in Medicine, 11*(2), 31-53. https://doi.org/10.3389/fmed.2024.1334919

Ramnarain, U. (2023). South African science teachers' experiences of inquiry-based teaching at disadvantaged schools. *Research in Science Education, 53*(4), 727-740. https://doi.org/10.1007/s11165-022-10095-1

Raposo-Rivas, M., Halabi-Echeverry, A. X., Sarmiento Campos, J. A., & García-Fuentes, O. (2024). Associations among variables in technology-enhanced phonological awareness programmes based on a meta-analysis. *Education Sciences, 14*(4), Article 343. https://doi.org/10.3390/educsci14040343

Rizki, I. A., & Suprapto, N. (2024). Project-oriented problem-based learning through SR-STEM to foster students' critical thinking skills in renewable energy material. *Journal of Science Education and Technology, 33*(4), 526-541. https://doi.org/10.1007/s10956-024-10102-2

Sari, N., Istiyono, E., Purbani, W., & Retnawati, H. (2023). The effect of the guided in meta-analysis study of the effect of guided inquiry model on critical thinking skills of students science concepts. *Jurnal Pendidikan Matematika Dan IPA, 14*(1), Article 82. https://doi.org/10.26418/jpmipa.v14i1.54483

Shen, Hu, X., Stein, D., & Liu, Q. (2024). Anatomizing online collaborative inquiry using directional epistemic network analysis and trajectory tracking. *British Journal of Educational Technology, 55*(5), 2173-2191. https://doi.org/10.1111/bjet.13441

Smela, B., Toumi, M., Świerk, K., Gawlik, K., Clay, E., & Boyer, L. (2023). Systematic literature reviews over the years. *Journal of Market Access and Health Policy, 11*(1), Article 2244305. https://doi.org/10.1080/20016689.2023.2244305

Soysal, Y. (2021). Talking science: Argument-based inquiry, teachers' talk moves, and students' critical thinking in the classroom. *Science and Education, 30*(1), 33-65. https://doi.org/10.1007/s11191-020-00163-1

Spector, J. M., & Ma, S. (2019). Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence. *Smart Learning Environments, 6*, Article 8. https://doi.org/10.1186/s40561-019-0088-z

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*, Article d4002. https://doi.org/10.1136/bmj.d4002

Strat, T. T. S., Henriksen, E. K., & Jegstad, K. M. (2024). Inquiry-based science education in science teacher education: A systematic review. *Studies in Science Education, 60*(2), 191-249. https://doi.org/10.1080/03057267.2023.2207148

Strippel, C. G., & Sommer, K. (2015). Teaching nature of scientific inquiry in chemistry: How do German chemistry teachers use labwork to teach NOSI? *International Journal of Science Education, 37*(18), 2965-2986. https://doi.org/10.1080/09500693.2015.1119330

Tikito, I., & Souissi, N. (2019). Meta-analysis of systematic literature review methods. *International Journal of Modern Education and Computer Science, 11*(2), 17-25. https://doi.org/10.5815/ijmecs.2019.02.03

Ting, F. S. T., Shroff, R. H., Lam, W. H., Garcia, R. C. C., Chan, C. L., Tsang, W. K., & Ezeamuzie, N. O. (2023). A meta-analysis of studies on the effects of active learning on Asian students' performance in science, technology, engineering and mathematics (STEM) subjects. *The Asia-Pacific Education Researcher, 32*(3), 379-400. https://doi.org/10.1007/s40299-022-00661-6

Toma, R. B. (2022). Confirmation and structured inquiry teaching: Does it improve students' achievement motivations in school science? *Canadian Journal of Science, Mathematics and Technology Education, 22*(1), 28-41. https://doi.org/10.1007/s42330-022-00197-3

van Uum, M. S. J., Verhoeff, R. P., & Peeters, M. (2016). Inquiry-based science education: Towards a pedagogical framework for primary school teachers. *International Journal of Science Education, 38*(3), 450-469. https://doi.org/10.1080/09500693.2016.1147660

Vančugovienė, V., Lehtinen, E., & Södervik, I. (2024). The impact of inquiry-based learning in a botanical garden on conceptual change in biology. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2024.2354542

Vergara-Díaz, C., Bustamante, K., Pinto, L., & Cofré, H. (2020). Exploring Chilean seventh grade students' conceptions of earth dynamics before and after model- and inquiry-based instruction. *Journal of Geoscience Education, 68*(4), 360-370. https://doi.org/10.1080/10899995.2020.1725406

Wale, B. D., & Bishaw, K. S. (2020). Effects of using inquiry-based learning on EFL students' critical thinking skills. *Asian-Pacific Journal of Second and Foreign Language Education, 5*(1), Article 9. https://doi.org/10.1186/s40862-020-00090-2

Walton, J., & Martin, J. L. (2023). Applying Sadler's principles in holistic assessment design: A retrospective account. *Teaching in Higher Education, 30*(1), 154-171. https://doi.org/10.1080/13562517.2023.2244439

Wan, Z. H., Zhan, Y., & Zhang, Y. (2024). Positive or negative? The effects of scientific inquiry on science achievement via attitudes toward science. *Science Education, 108*(1), 3-24. https://doi.org/10.1002/sce.21825

Wang, Y., Zhang, L., & Pang, M. (2024). Virtual experiments in physics education: A systematic literature review. *Research in Science & Technological Education*. https://doi.org/10.1080/02635143.2024.2327995

Wardani, S., Lindawati, L., & Kusuma, S. B. W. (2017). The development of inquiry by using android-system-based chemistry board game to improve learning outcome and critical thinking ability. *Jurnal Pendidikan IPA Indonesia, 6*(2), 196-205. https://doi.org/10.15294/jpii.v6i2.8360

Wei, B., Wang, H., Li, F., Long, Y., Zhang, Q., Liu, H., Tang, X., & Rao, M. (2023). Effectiveness of problem-based learning on development of nursing students' critical thinking skills. *Nurse Educator, 49*(3), 3-7. https://doi.org/10.1097/nne.0000000000001548

Wen, A. S., Zaid, N. M., & Harun, J. (2015). A meta-analysis on students' social collaborative knowledge construction using flipped classroom model. In *Proceedings of the 2015 IEEE Conference on E-Learning, e-Management and e-Services* (pp. 58-63). IEEE. https://doi.org/10.1109/IC3e.2015.7403487

Wen, Y., Wu, L., He, S., Ng, N. H. E., Teo, B. C., Looi, C. K., & Cai, Y. (2023). Integrating augmented reality into inquiry-based learning approach in primary science classrooms. *Educational Technology Research and Development, 71*(4), 1631-1651. https://doi.org/10.1007/s11423-023-10235-y

Winje, Ø., & Løndal, K. (2020). Bringing deep learning to the surface: A systematic mapping review of 48 years of research in primary and secondary education. *Nordic Journal of Comparative and International Education, 4*(2), 25-41. https://doi.org/10.7577/njcie.3798

Wulandari, D. S., Prayitno, B. A., & Maridi, M. (2022). Developing the guided inquiry-based module on the circulatory system to improve student's critical thinking skills. *Jurnal Pendidikan Biologi Indonesia, 8*(1), 77-85. https://doi.org/10.22219/jpbi.v8i1.16512

Yang, J. M., Sung, Y. T., & Chang, K. E. (2020). Use of meta-analysis to uncover the critical issues of mobile inquiry-based learning. *Journal of Educational Computing Research, 58*(4), 715-746. https://doi.org/10.1177/0735633119879366

Yasa, A. D., Rahayu, S., Handayanto, S. K., & Ekawati, R. (2024). Evaluating the impact of smart learning-based inquiry on enhancing digital literacy and critical thinking skills. *Ingenierie Des Systemes d'Information, 29*(1), 219-233. https://doi.org/10.18280/isi.290122

Yeoh, C.-P., Li, C.-T., & Hou, H.-T. (2024). Game-based collaborative scientific inquiry learning using realistic context and inquiry process-based multidimensional scaffolding. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2024.2354944

Yu, Z., & Xu, W. (2022). A meta-analysis and systematic review of the effect of virtual reality technology on users' learning outcomes. *Computer Applications in Engineering Education, 30*(5), 1470-1484. https://doi.org/10.1002/cae.22532

Yun, M., & Crippen, K. J. (2024). The use of educational technology in inquiry-based elementary science education: A systematic review. *International Journal of Science Education, 46*(12), 1256-1279. https://doi.org/10.1080/09500693.2023.2284689

Zhang, W., Guan, Y., & Hu, Z. (2024). The efficacy of project-based learning in enhancing computational thinking among students: A meta-analysis of 31 experiments and quasi-experiments. *Education and Information Technologies, 29*, 14513-14545. https://doi.org/10.1007/s10639-023-12392-2

Zheng, L., Li, X., Tian, L., & Cui, P. (2018). The effectiveness of integrating mobile devices with inquiry-based learning on students' learning achievements: A meta-analysis. *International Journal of Mobile Learning and Organisation, 12*(1), 77-95. https://doi.org/10.1504/IJMLO.2018.089238

APPENDIX A

Table A1. Description of the results of empirical studies about effect IBL on critical thinking skills

Reference	Journal	Strategy	Country	Tittle	Results
Arsal (2017)	International Journal of Science Education	IBL	Turkey	The impact of inquiry-based learning on the critical thinking dispositions of pre-service science teachers	The results of the study showed that prospective teachers in the experimental group did not show statistically significant progress in terms of critical thinking tendencies compared to those in the control group.
Lu et al. (2020)	Journal of Baltic science education	Critique-driven inquiry	Taiwan	The effects of critique-driven inquiry intervention on students' critical thinking and scientific inquiry competence	The results of the study showed that students taught with the CDI model were significantly more effective in terms of critical thinking and scientific inquiry skills during and after the CDI intervention.
Ku et al. (2014)	Instructional Science	IBL	China	Integrating direct and inquiry-based instruction in the teaching of critical thinking: An intervention study	Students who received training showed greater improvement on at least one critical thinking assessment compared to those who did not receive training.
Greenwald and Quitadamo (2014)	Journal of Undergraduate Neuroscience Education	Inquiry-based clinical case	Canada	A mind of their own: Using inquiry-based teaching to build critical thinking skills and intellectual engagement in an undergraduate neuroanatomy course	The results showed that students in conventional neuroanatomy courses earned less than 3rd percentile national rankings while IBCC students earned more than 7.5 in one academic semester using the valid and reliable California test of critical thinking skills.
Pursitasari et al. (2020)	Jurnal Pendidikan IPA Indonesia	Science context-based inquiry learning	Indonesia	Enhancement of student's critical thinking skill through science context-based inquiry learning	The results of the implementation of the SCOIL model showed an increase in activity with a high category and the N-gain of critical thinking skills was in the medium category. The significance test showed that students' critical thinking skills with the SCOIL model were greater than those with the guided inquiry learning model.
Duran and Dökme (2016)	Eurasia Journal of Mathematics, Science and Technology Education	IBL	Turkey	The effect of the inquiry-based learning approach on student's critical-thinking skills	The research findings reveal that science and technology learning supported by guided activities developed in line with the IBL approach has a significant influence on students' critical thinking skills in science and technology subjects.
Yasa et al. (2024)	Ingenierie des Systemes d'Information	Smart learning-based inquiry	Indonesia	Evaluating the impact of smart learning-based inquiry on enhancing digital literacy and critical thinking skills	The results of the study showed that the SLBI strategy, which includes six stages (orientation, conceptualization, investigation, digital report design, reflection, and publishing), significantly improved critical thinking skills (effect size of 1.504 is categorized as large) compared to the traditional inquiry learning method.
Hwang and Chen (2016)	British Journal of Educational Technology	Inquiry-based ubiquitous gaming environment	China	Influences of an inquiry-based ubiquitous gaming design on students' learning achievements, motivation, behavioral patterns, and tendency towards critical thinking and problem solving	The experimental results showed that the inquiry approach improved students' performance in learning achievement and intrinsic motivation; furthermore, students' perceptions of their problem solving and critical thinking also improved significantly.
Ay and Daghan (2023)	Education and Information Technologies	Flipped inquiry learning	Turkey	The effect of the flipped learning approach designed with community of inquiry model to the development of students' critical thinking strategies and social, teaching and cognitive presences	The findings suggest that it is possible to eliminate the lack of interaction and feedback processes, and to develop students' critical thinking strategies and their perceptions of teaching.
Pahrudin et al. (2021)	European Journal of Educational Research	STEM-inquiry	Indonesia	The effectiveness of science, technology, engineering, and mathematics-inquiry learning for 15-16 years old students based on K-13 Indonesian curriculum: The impact on the critical thinking skills	The results of the hypothesis test show that students' critical thinking skills before and after using the STEM-based inquiry model are different, so it can be concluded that the application of the STEM-based inquiry model is effective in improving students' critical thinking skills.

Table A1 (Continued). Description of the results of empirical studies about effect IBL on critical thinking skills

Reference	Journal	Strategy	Country	Tittle	Results
Isiklar and Abali Öztürk (2022)	International Journal of Contemporary Educational Research	Community of inquiry model	Turkey	The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills	The results of the analysis showed a significant difference between the pre-test and post-test scores of the two groups in terms of critical thinking skills of the experimental group through philosophical inquiry in the group comparison. Although the average post-test score of the critical thinking skills of the experimental group through philosophical inquiry was higher than the control group , no statistically significant difference was observed between the two.
Farah and Ayoubi (2020)	Journal of Education in Science, Environment and Health	IBL and reflection	Lebanon	Enhancing the critical thinking skills of grade 8 chemistry students using an inquiry and reflection teaching method	The findings showed that students in the experimental group improved their critical thinking skills significantly, while the critical thinking of students in the control group remained the same.
Karakas and Sarikaya (2020)	İlköğretim Online	Argument-driven inquiry	Turkey	The effect of argumentation-based teaching performed for environment-energy issues on critical thinking of prospective classroom teachers	significant differences were found in favor of the mean scores of the experimental group's post-test critical thinking skills ($p < 0.05$), and the effect of its implementation was a "middle effect" (R2: 0.11). The critical thinking level of the experimental group's prospective classroom teachers was found to be higher than that of the control group.
Kurniati et al. (2024)	KnE Social Sciences	IBL	Indonesia	The effect of inquiry learning on the critical thinking skills of students in the era of the Industrial Revolution 4.0	The results obtained after the research proved that students who were taught using the inquiry learning model had higher critical thinking skills compared to students who were taught using the conventional learning model.
Aidoo et al. (2022)	European Journal of Science and Mathematics Education	Flipped inquiry learning	Ghana	A mixed-method approach to investigate the effect of flipped inquiry-based learning on chemistry students learning	The results of the quantitative data analysis showed a significant improvement in students' academic performance and critical thinking skills in pre- and post-test scores in the experimental class.
Liang et al. (2021)	British Journal of Educational Technology	Concept mapping-based questioning approach	Taiwan	Promoting children's inquiry performances in alternate reality games: A mobile concept mapping-based questioning approach	Quantitative and qualitative analyses were examined to identify the effects of different learning strategies to support children's learning with ARG. ARG-CMQ improved children's learning achievement, critical thinking and learning attitudes better than conventional ARG.
Aiman et al. (2020)	European Journal of Educational Research	Process oriented guided inquiry learning	Indonesia	The influence of process oriented guided inquiry learning (POGIL) model assisted by realia media to improve scientific literacy and critical thinking skill of primary school students	There is a significant difference between the critical thinking of students who receive POGIL learning assisted by realia media and students who receive expository learning.
Rosidin et al. (2019)	Jurnal Cakrawala Pendidikan	Argument-driven inquiry	Indonesia	Can argument-driven inquiry models have impact on critical thinking skills for students with different personality types?	The results of the study showed that there was an influence of ADI learning on critical thinking skills in students with high and low academic abilities.
Mitarlis et al. (2020)	European Journal of Educational Research	New IBL	Indonesia	The effectiveness of new inquiry-based learning (NIBL) for improving multiple higher-order thinking skills (M-HOTS) of prospective chemistry teachers	The NIBL model effectively improves the M-HOTS of prospective chemistry teachers in terms of CACP thinking skills and contributes significantly to improving students' mastery of organic chemistry concepts.
Taylor et al. (2018)	Journal of Science Education for Students with Disabilities	Argument-driven inquiry	USA	Using argument-based science inquiry to improve science achievement for students with disabilities in inclusive classrooms	students in the SWH group scored significantly better than the comparison group on post-test science achievement scores. The authors also found stronger effect size results for the SWH group.

Table A1 (Continued). Description of the results of empirical studies about effect IBL on critical thinking skills

Reference	Journal	Strategy	Country	Tittle	Results
Cornejo et al. (2022)	Electronic Journal of Research in Educational Psychology	Inquiry probabilistic reasoning	Chile	Promoting pedagogy students' scientific skills through critical thinking program	The results showed statistically significant differences in both variables for the experimental group, with a large effect on the investigation variable, and a moderate effect on the probabilistic reasoning variable.
Wen et al. (2023)	Educational Technology Research and Development	QIMS inquiry	Singapore	Integrating augmented reality into inquiry-based learning approach in primary science classrooms	The use of AR and QIMS has a significant effect in improving students' critical thinking skills and knowledge creation skills.
Polat and Aydın (2020)	Thinking Skills and Creativity	IBL	Turkey	The effect of mind mapping on young children's critical thinking skills	In line with the results of the analysis, it is understood that the critical thinking skill scores of children in experimental group-1 and experimental group-2 were significantly higher than those of children in the control group. It was also found that experimental group-2 who conducted large group mind mapping made further progress than children in experimental group-1 who were subjected to individual mind mapping studies.
Kim et al. (2012)	Research in Science Education	IBL	USA	Project clarion: Three years of science instruction in title I schools among K-third grade students	The results indicate that all ability groups of students benefit from an inquiry-based science learning approach that emphasizes science concepts, and that there are positive achievement effects for young children of low socioeconomic status who engage in such a curriculum.
Prayogi et al. (2023)	Journal of Education and e-Learning Research	Ethno-inquiry	Indonesia	Dynamic blend of ethnoscience and inquiry in a digital learning platform (e-learning) for empowering future science educators' critical thinking	According to the study findings, there has been significant progress in CT due to the dynamic blend of Ethnoscience and inquiry on digital learning platforms. This study suggests that the dynamic blend of ethnoscience and inquiry on digital learning platforms can serve as a cutting-edge learning method to empower CT especially in pre-service science teachers.

APPENDIX B

Table B1. Quantitative studies relating the effect of IBL on students' critical thinking

Reference	Grade level	Strategy	Inquiry Level	Country	Duration	Subject	Media ICT	Assessments	Types
Arsal (2017)	Undergraduate	IBL	Guided inquiry	Turkey	14 weeks	Science	Not specified	California critical thinking dispositions inventory	Questioner
Lu et al. (2020a)	Elementary school	Critique-driven inquiry	Structured inquiry	Taiwan	20 weeks	Physics	Not specified	California critical thinking dispositions inventory	Essay
Lu et al. (2020b)	Elementary school	Critique-driven inquiry	Structured inquiry	Taiwan	20 weeks	Physics	Not specified	California critical thinking dispositions inventory	Essay
Lu et al. (2020c)	Middle school	Critique-driven inquiry	Structured inquiry	Taiwan	20 weeks	Physics	Not specified	California critical thinking dispositions inventory	Essay
Lu et al. (2020d)	Middle school	Critique-driven inquiry	Structured inquiry	Taiwan	20 weeks	Physics	Not specified	California critical thinking dispositions inventory	Essay
Ku et al. (2014a)	High school	IBL	Guided inquiry	China	4 weeks	Science	Not specified	Watson Glaser critical thinking appraisal	Close-ended
Ku et al. (2014b)	High school	IBL	Guided inquiry	China	4 weeks	Science	Not specified	Halpern critical thinking assessment	Open-ended
Ku et al. (2014c)	High school	IBL	Guided inquiry	China	4 weeks	Science	Not specified	Halpern critical thinking assessment	Close-ended
Greenwald and Quitadamo (2014a)	Undergraduate	Inquiry based clinical case	Guided inquiry	Canada	3 weeks	Biology	Not specified	California critical thinking skills test	Essay
Greenwald and Quitadamo (2014b)	Undergraduate	Inquiry-based clinical case	Guided inquiry	Canada	3 weeks	Biology	Not specified	California critical thinking skills test	Essay
Pursitasari et al. (2020)	Middle school	Science context-based inquiry learning	Guided inquiry	Indonesia	3 weeks	Science	Books and printed materials	Ennis-Weir critical thinking essay test	Multiple choices
Duran and Dökme (2016a)	Middle school	IBL	Guided inquiry	Turkey	4 weeks	Science	Books and printed materials	Critical thinking skills scale	Essay
Duran and Dökme (2016b)	Middle school	IBL	Guided inquiry	Turkey	4 weeks	Science	Books and printed materials	Critical thinking skills scale	Essay
Yasa et al. (2024)	Elementary school	Smart Learning Based Inquiry	Open inquiry	Indonesia	6 weeks	Science	E-learning system	CTST-MC	Multiple choices
Hwang and Chen (2016)	Elementary school	Inquiry-based ubiquitous gaming environment	Structured inquiry	China	3 weeks	Science	Digital game	CTST-RS	Rating scale
Ay and Daghan (2023)	Undergraduate	Flipped Inquiry learning	Guided inquiry	Turkey	15 weeks	Physics	E-learning system	Critical thinking strategies for learning questionnaire	Rating scale
Pahrudin et al. (2021)	High school	STEM-inquiry	Structured inquiry	Indonesia	4 weeks	Physics	Not specified	CTST-E	Essay
Isiklar and Abali Ozturk (2022)	Preschool	Community of inquiry model	Confirmation inquiry	Turkey	10 weeks	Science	Not specified	Critical thinking through philosophical inquiry	Rating scale

Table B1 (Continued). Quantitative studies relating the effect of IBL on students' critical thinking

Reference	Grade level	Strategy	Inquiry Level	Country	Duration	Subject	Media ICT	Assessments	Types
Farah and Ayoubi (2020)	Middle school	IBL and reflection	Guided inquiry	Lebanon	2 semesters	Chemistry	Not specified	James Madison critical thinking test	Multiple choices
Karakas and Sarikaya (2020)	Undergraduate	Argument-driven inquiry	Guided inquiry	Turkey	10 weeks	Physics	Not specified	Watson Glaser critical thinking appraisal	Multiple choices
Kurniati et al. (2024)	High school	IBL	Guided inquiry	Indonesia	Not specified	Biology	Not specified	CTST-MC	Multiple choices
Aidoo et al. (2022)	Undergraduate	Flipped inquiry learning	Guided inquiry	Ghana	4 weeks	Chemistry	E-learning system	CTST-MC	Multiple choices
Liang et al. (2021)	Elementary school	Concept mapping-based questioning approach	Structured inquiry	Taiwan	Not specified	Science	Alternate reality game mobile	Critical Thinking Strategies for Learning Questionnaire	Questioner
Aiman et al. (2020)	Elementary school	Process oriented guided inquiry learning	Guided inquiry	Indonesia	4 weeks	Science	Realia media	CTST-MC	Multiple choices
Rosidin et al. (2019)	Middle school	Argument-driven inquiry	Guided inquiry	Indonesia	Not specified	Science	Not specified	CTST-E	Essay
Mitarlis et al. (2020)	Undergraduate	New IBL	Open Inquiry	Indonesia	Not specified	Chemistry	Not specified	CTST-E	Essay
Taylor et al. (2018)	Elementary school	Argument-driven inquiry	Guided inquiry	USA	1 month	Science	Computer simulation	The Iowa test of basic skills	Multiple choices
Cornejo et al. (2022)	Undergraduate	Inquiry probabilistic reasoning	Guided inquiry	Chile	1 semester	Science	Not specified	Critical thinking task (Miranda, 2003)	Multiple choices
Wen et al. (2023a)	Elementary school	QIMS inquiry	Guided inquiry	Singapore	3 weeks	Biology	Augmented reality	Critical thinking test	Open ended questions
Wen et al. (2023b)	Elementary school	QIMS inquiry	Guided inquiry	Singapore	3 weeks	Biology	E-learning system	Critical thinking test	Open ended questions
Polat and Aydin (2020a)	Preschool	IBL	Structured inquiry	Turkey	8 weeks	Science	Not specified	Critical thinking through philosophical inquiry	Rating scale
Polat and Aydin (2020b)	Preschool	IBL	Structured inquiry	Turkey	8 weeks	Science	Not specified	Critical thinking through philosophical inquiry	Rating scale
Kim et al. (2012a)	Elementary school	IBL	Guided inquiry	USA	1 year	Science	Not specified	Paul-Elder critical thinking test	Essay
Kim et al. (2012b)	Elementary school	IBL	Guided inquiry	USA	2 years	Science	Not specified	Paul-Elder critical thinking test	Essay
Kim et al. (2012c)	Elementary school	IBL	Guided inquiry	USA	3 years	Science	Not specified	Paul-Elder critical thinking test	Essay
Prayogi et al. (2023)	Undergraduate	Ethno-inquiry	Guided inquiry	Indonesia	Not specified	Science	E-learning system	Ennis-Weir critical thinking essay test	Essay

APPENDIX C

Table C1. Summary of the moderator analysis

Moderate variable	Category	Number of studies	p	SMD	95% CI	Effect
Grade level	Undergraduate	9	0.01	2.66	[0.71; 4.62]	Big
	Elementary school	12		0.57	[0.38; 0.76]	Moderate
	Middle school	7		1.43	[0.89; 1.97]	Big
	High school	5		0.81	[0.15; 1.46]	Big
	Preschool	3		0.86	[0.43; 1.28]	Big
Strategy	IBL	12	< 0.01	0.65	[0.33; 0.98]	Moderate
	Critique-driven inquiry	4		1.07	[0.63; 1.51]	Big
	Inquiry based clinical case	2		5.18	[2.32; 7.13]	Big
	Science context-based inquiry learning	1		1.77	[1.15; 2.39]	Big
	Smart learning-based inquiry	1		1.41	[0.61; 1.91]	Big
	Inquiry-based ubiquitous gaming environment	1		0.39	[0.00; 0.78]	Small
	Flipped inquiry learning	2		0.74	[0.43; 1.05]	Moderate
	STEM-inquiry	1		1.20	[0.61; 1.80]	Big
	Community of inquiry model	1		0.41	[-0.22; 1.04]	Small
	IBL and reflection	1		2.81	[1.89; 3.73]	Big
	Argument-driven inquiry	1		1.30	[0.76; 1.83]	Big
	Concept mapping-based questioning approach	1		0.60	[-0.01; 1.22]	Moderate
	Process oriented guided inquiry learning	1		0.88	[0.34; 1.42]	Big
	New IBL	1		1.86	[0.44; 3.28]	Big
	Inquiry probabilistic reasoning	1		1.45	[1.00; 1.91]	Big
	QMIS-inquiry	2		0.43	[0.12; 0.75]	Small
	Ethno-inquiry	1		7.27	[2.27; 10.95]	Big
Level	Guided inquiry	24	0.02	1.46	[0.69; 2.23]	Big
	Structured inquiry	9		0.93	[0.67; 1.19]	Big
	Open inquiry	2		1.41	[1.08; 1.74]	Big
	Confirmation inquiry	1		0.41	[-0.22; 1.04]	Small
Country	Turkey	8	< 0.01	0.71	[0.43; 1.00]	Moderate
	Taiwan	5		0.96	[0.56; 1.36]	Big
	China	4		0.75	[0.02; 1.51]	Moderate
	Canada	2		5.18	[2.32; 7.13]	Big
	Indonesia	8		0.59	[0.22, 0.96]	Moderate
	Lebanon	1		2.81	[1.89; 3.73]	Big
	Ghana	1		0.87	[0.40; 1.22]	Big
	USA	4		1.43	[1.00; 1.91]	Big
	Chile	1		0.43	[0.12; 0.75]	Small
	Singapore	2		1.45	[1.00; 1.91]	Big
Duration	14 weeks	1	< 0.01	0.23	[0.29; 0.76]	Small
	20 weeks	4		1.07	[0.63; 1.51]	Big
	4 weeks	8		2.20	[0.48; 3.91]	Big
	3 weeks	1		1.48	[1.06; 1.91]	Big
	15 weeks	1		1.66	[1.22; 2.11]	Big
	10 weeks	2		2.81	[1.89; 3.73]	Big
	2 semester	1		2.47	[0.41; 5.56]	Big
	Not specified	11		2.40	[0.70; 6.01]	Big
	1 month	1		2.00	[0.20; 6.00]	Big
	1 semester	5		1.24	[0.64; 1.85]	Big
	8 weeks	1		1.02	[0.51; 1.40]	Big
	1 year	2		0.35	[0.15; 0.64]	Small
	2 years	1		0.41	[0.12; 0.70]	Small
	3 years	1		0.41	[0.12; 0.70]	Small
Discipline	Science	21	0.51	1.10	[0.49; 1.71]	Big
	Physics	7		0.95	[0.67; 1.23]	Big
	Biology	5		2.27	[0.42; 4.65]	Big
	Chemistry	3		1.58	[0.45; 2.70]	Big
ICT media	Not specified	21	0.02	1.28	[0.70; 1.87]	Big
	Books and printed materials	5		1.17	[0.23; 2.10]	Big
	E-learning system	3		2.4	[0.41; 5.56]	Big
	Digital game	1		3.5	[-0.01; 7.00]	Big
	Alternate reality game mobile	1		0.39	[0.00; 0.78]	Small
	Realia media	1		0.88	[0.34; 1.42]	Big
	Computer simulation	2		0.80	[0.20; 1.22]	Big
	Augmented reality	1		0.48	[0.03; 0.93]	Small
	Mind mapping	2		1.02	[0.63; 1.40]	Big

Table C1 (Continued). Summary of the moderator analysis

Moderate variable	Category	Number of studies	p	SMD	95% CI	Effect
Assessment	California critical thinking dispositions inventory	5	< 0.01	0.9	[0.42; 1.37]	Big
	Watson Glaser critical thinking appraisal	2		0.47	[0.19; 0.75]	Small
	Halpern critical thinking assessment	2		1.11	[0.45; 2.68]	Big
	California critical thinking skills test	2		5.18	[2.32; 7.13]	Big
	Ennis-Weir critical thinking essay test	2		5.39	[1.81; 12.58]	Big
	Critical thinking skills scale	2		0.87	[0.34; 1.35]	Big
	CTST-MC	1		0.39	[0.01; 0.78]	Small
	CTST-RS	1		1.11	[0.26; 1.97]	Big
	Critical thinking strategies for learning questionnaire	1		0.62	[0.09; 1.15]	Moderate
	CTST-E	3		1.24	[0.65; 1.83]	Big
	Critical thinking through philosophical inquiry	1		0.86	[0.43; 1.28]	Big
	James Madison critical thinking test	1		2.81	[1.89; 3.73]	Big
	The Iowa test of basic skills	1		0.4	[0.20; 0.60]	Small
	Critical thinking task (Miranda, 2003)	1		0.35	[0.15; 0.64]	Small
	Critical thinking test	1		0.43	[0.12; 0.75]	Small
	Paul-Elder critical thinking test	3		0.36	[0.23; 0.50]	Small
Type of assessment	Questioner	2	< 0.01	0.39	[0.01; 0.79]	Small
	Essay	15		1.95	[0.75; 3.15]	Big
	Close-ended	2		3.04	[1.06; 5.01]	Big
	Open-ended	3		1.91	[1.60; 2.22]	Big
	Multiple choices	9		1.11	[0.64; 1.58]	Big
	Rating scale	5		0.7	[0.40; 1.00]	Moderate