A Meta-Analysis on the Effectiveness of Problem-Based Learning on Students' Cognitive Outcomes in Chemistry Education

Jessa Mae N. Jimenez[1] , Karen M. Napay[2]

Bicol University Graduate School, Bicol University, Legazpi City Philippines

DOI: https://doi.org/10.47772/IJRISS.2026.101000050

Received: 03 October 2026; Accepted: 08 October 2026; Published: 29 October 2026

ABSTRACT

Problem-based learning has been widely recognized as a learner-centered instructional approach for developing students' cognitive outcomes in chemistry education, but empirical findings on its effectiveness remain inconsistent, specifically in different regions, educational levels, and specific chemistry topics. In this meta-analysis, 14 experimental and quasi-experimental peer-reviewed studies from 2017-2026 were synthesized from six databases (ERIC, Cross-Reference (CrossRef), Google Scholar, ProQuest, OpenAlex, and ResearchGate) to determine the overall impact of PBL on students' cognitive outcomes in chemistry education compared to conventional methods and to examine how this effect differs based on the type of cognitive outcome level: academic achievement and higher-order thinking skills; continental location: Africa versus Asia & Middle East; and class size. The effect sizes were calculated grounded on the posttest means, standard deviations, and sample sizes from the combined total of 1,299 participants using Meta-Essentials v. 1.5. Findings showed that PBL had a large, statistically significant pooled effect compared to the conventional method (Hedges' $g = 1.18$, 95% CI [0.69, 1.68], $p < .001$), although high heterogeneity was observed across the studies ($I^2 = 90.98\%$), which led to further subgroup and moderator analyses. Subgroup analyses showed that cognitive outcome level and continental location did not significantly moderate this effect, indicating that PBL produced comparable benefits across both outcome types and both regions. Likewise, class size was not a significant moderator ($p = .572$). Assessment of publication bias using Trim-and-Fill and Egger's regression test found no significant evidence of bias. These findings suggest that PBL is broadly and consistently effective in improving students' cognitive outcomes in chemistry education, regardless of outcome type, regional context, or class size, although the high residual heterogeneity indicates that implementation should still be tailored to specific instructional contexts.

Keywords: Problem-Based Learning, Cognitive Outcomes, Conventional Method, Meta-Analysis, Chemistry Education

INTRODUCTION

In the 21st century, there have been significant changes in every aspect of human life. These changes resulted from faster technological growth, which led to increasingly complex challenges for human life across disciplines (Muyambo-Goto et al., 2023). Because of these changes in education, one of the most significant goals is to prepare students for the real challenges in life. An education system that can give the skills needed that are required in the 21st century for the students (Varas et al., 2023). It aims to develop learners who are scientifically literate and capable of applying scientific knowledge to real-life situations. Despite this goal, chemistry is often considered difficult because students' disengagement and limited appreciation of its relevance stem from traditional instruction that fails to connect theoretical concepts into real-life situations (Reyes, 2025). This difficulty is rooted in students' need to constantly connect three different levels of representation, which are the macroscopic, the microscopic, and the symbolic area in chemistry (Ahmar, D., et al., 2020). So students' disengagement and limited appreciation of its relevance started with traditional instruction that fails to connect these abstract and complex concepts to real-world situations.

According to Dewey (1923), traditional education created the wrong kind of learning experience because the world was constantly changing, and students needed to be active learners to be involved in the learning process. Also, among the factors linked to poor performance and attitude in chemistry, Babale et al. (2018) also mentioned poor teaching strategies, which are primarily the traditional teaching methods. Additionally, lecture-based learning (LBL), which is one of the traditional lecture based instructional methods that emphasizes rote memorization and passive listening, often leaves students with persistent different ideas or misconceptions because chemistry topics are so abstract, particularly in complex areas like bond energies, chemical equilibrium, and thermodynamic system enthalpy changes.

To meet the demands of the 21st century, students must learn how to learn and think. Constructivism and Vygotsky's Zone of Proximal Development, which essentially maintain that learners can construct meaning most successfully if knowledge is built and practiced through direct engagement and the use of authentic, contextualized problems rather than through passive teaching of information. Therefore, science teachers must utilize learner-centered instructional approaches that promote critical thinking, active engagement, and meaningful learning experiences (Funa & Prudente, 2021). To improve learners' performance and promote a positive attitude, an engaging approach is required, such as problem-based learning.

Problem-based learning, on the other hand, is more on teamwork since it focuses on student-centered learning and problem-solving; it emphasizes real-world problems and develops active learning and problem-solving skills in students (Kurt, 2020). Students are the navigators of their own learning journey, using their prior knowledge, collaborating with peers, and critically evaluating information to create a solution inside a PBL classroom (Sappaile et al., 2025). Because of this, PBL is widely recognized as a way of nurturing transferable skills and competencies such as collaboration, critical evaluation, self-directed learning, and effective communication as part of content mastery (Siregar, 2025).

In the context of chemistry, the teacher's role in a PBL classroom is to support students' learning by keeping an eye on their progress and asking scaffolding questions to help them advance their knowledge. The teacher serves as a guide while students look for more relevant resources rather than as the only source of content or process information. PBL implementation in the classroom has developed into unique structural modalities in recent years. It has direct applications of problem-solving modules to specific chemistry curricular units, such as using specialized inquiry worksheets to target and correct deep-seated misconceptions in thermodynamics (Rosyidah K., et al., 2024). And an interdisciplinary approach where chemistry concepts are blended with STEM/STEAM to structurally align with and to enhance multi-dimensional critical thinking skills (Monika, P., et al., 2023). In this review, "cognitive outcomes" refers to quantifiable improvements in students' academic achievement, conceptual understanding, critical thinking, and problem-solving skills, which are the fundamental knowledge and competencies commonly evaluated in PBL intervention studies. In the past ten years, studies have consistently shown that students' cognitive outcomes have improved with the use of problem-based learning in place of traditional lecture-based instruction. For instance, Ayyildiz et al. (2018) and Baran and Sozbilir (2018) found that PBL enhanced students' conceptual grasp of thermodynamics and enthalpy changes, respectively. Furthermore, studies indicate that problem-based learning enhances high school students' problem-solving skills and is strongly linked to the positive development of learners' cognitive styles (Arita, M., et al., 2026).

However, even with these various achievements of using PBL, the larger body of educational literature is still very dispersed. Different regional contexts, educational levels, and particular chemistry areas all have very different measured effects of PBL. Although some studies demonstrate different degrees of effectiveness when addressing traditional, formula-heavy chemistry topics, others report exceptionally large effect sizes in highly integrated PBL environments. It has also been concluded that the effect of PBL on students' outcomes depends on the regional and cultural context. A recent meta-analysis on PBL and critical thinking in medical education since PBL was first used in medical education, which is now widely adopted in other education. It concluded, based on the substantial heterogeneity across studies from different countries, that PBL may represent a more innovative and significant instructional shift in some regions than in others. Part of this variation resulted from differences in educational culture and student experience with discussion-based learning methods (Su et al., 2025). Another systematic review, but in science education,

also concluded that it improves students' critical thinking and collaboration skills but was also impacted by changes in instructional design, contextualization of learning materials, and teachers' readiness, since cultural and social diversity also affects how students interact with learning (Affandy et al., 2026). These results show the importance of creating more flexible, context-based curricula that encourage students' active participation in a variety of learning environments. This means that PBL's effectiveness might not be consistent across various educational systems and geographical locations, therefore requiring further study of how continental or regional context changes its effect on cognitive outcomes.

In addition to regional context, the specific type of cognitive outcome being measured is also an important aspect to be considered. Academic achievement and higher-order thinking skills, such as critical thinking and problem-solving, are represented by different levels of cognitive demand, which require more complex mental processing compared to checking only basic content comprehension according to the revised Bloom's taxonomy (Anderson & Krathwohl, 2001). These two types of outcomes are distinct from each other; therefore, the effectiveness of PBL may not be the same for each, making it necessary to examine them separately instead of treating them as one construct.

Although several reviews have examined generally the impact of PBL on chemistry learning outcomes, fewer studies have quantitatively synthesized PBL's specific impact on students' cognitive outcomes such as conceptual understanding, critical thinking, and problem-solving skills. Therefore, a meta-analysis is essential given the different effect sizes found in individual studies, the ongoing application of PBL as a learner-centered approach for 21st-century science instruction, and the determination of how specific subgroup variables, such as cognitive outcome level and continental location, as well as moderator variables such as sample size, can affect the attainment of cognitive learning outcomes of the students. To determine the overall scope, effectiveness, and consistency of PBL's impact on cognitive outcomes in chemistry education, this study attempts to statistically synthesize these inconsistent empirical findings.

General and Specific Objectives

This meta-analysis was conducted given these inconsistencies and gaps in the literature, along with the need to statistically synthesize the different effects in chemistry education of problem-based learning on students' cognitive outcomes. It specifically aims to answer the following research questions:

RQ1: How effective is problem-based learning (PBL) in improving students' cognitive outcomes in chemistry education compared to conventional methods?

RQ2: How does the effect of problem-based learning on students' cognitive outcomes in chemistry education vary according to the following subgroups:

A. Cognitive Outcome Level

B. Continental Location

RQ3: Is sample size a significant moderator of the relationship between problem-based learning and students' cognitive outcomes in chemistry education?

Conceptual framework

This meta-analysis investigated the effect of problem-based learning used in chemistry education on students' cognitive outcomes, measured by academic achievement and higher-order thinking skills, namely critical thinking and problem-solving skills. This focused on the concept of cognitive outcomes, described as the students' quantifiable improvements in academic achievement, theoretical understanding, critical thinking, and problem-solving skills because of exposure to problem-based learning. To further investigate the effectiveness of PBL, sample size was used as a moderator variable in the study, while the cognitive outcome level and the continent on which the study was conducted were used as subgroups.

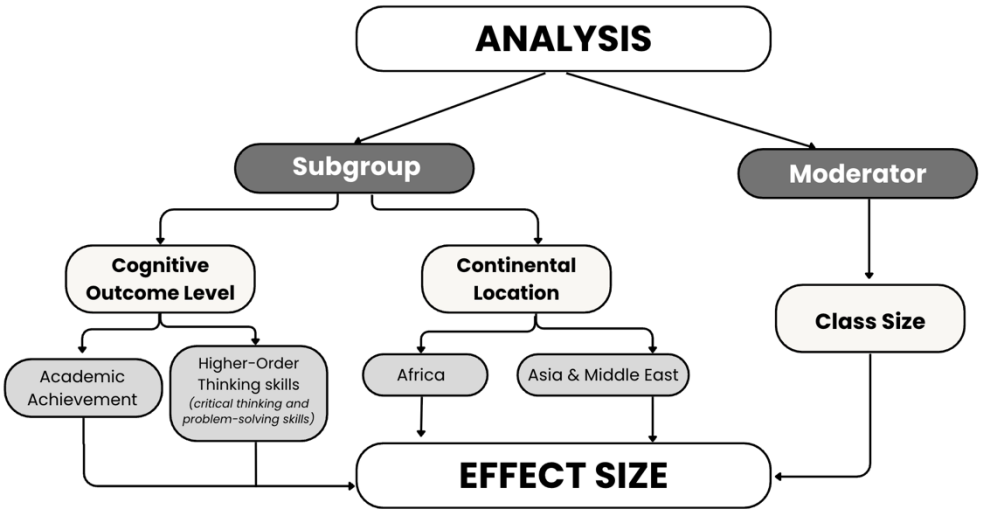


Fig. 1. Conceptual Paradigm

Figure 1 shows the conceptual paradigm of the study. Two analyses were conducted based on the effect sizes of the articles included: a subgroup analysis that was conducted on cognitive outcome level (academic achievement and higher-order thinking skills) and continental location (Africa and Asia & the Middle East), and a moderator analysis was conducted on sample size, which was considered continuous data.

METHODOLOGY

Research Design

A quantitative research method was used in this review, specifically meta-analysis, which statistically combines the results of different independent studies that address the same research question. This allows researchers to compute a pooled effect size that arrives at conclusions that are more precise and generalizable compared to those of a single study (Borenstein et al., 2021). This review relies on numerical data extracted directly from each included study, specifically the sample sizes, means, standard deviations, and reported effect sizes, which are then statistically weighted and combined. This research used software tools to compute effect sizes, assess heterogeneity, and synthesize the quantitative studies that are included in this review:

1. Harzing's Publish or Perish 8 was used to access and extract citation data from Cross-Reference (CrossRef), Education Resources Information Center (ERIC), Google Scholar, ProQuest, OpenAlex, and ResearchGate.
2. EndNote 21 (Free trial version) was utilized as the reference management tool for this meta-analysis to import, organize, and evaluate all the studies that were identified (EndNote, 2025); and
3. SPSS Statistics version 31 was used to combine the results of all included quantitative studies to have a unified conclusion out of the combined data. It was used for the graphical interface, which allows the researchers to easily analyze the data, test for differences, and create visual graphs.

Data Sources and Search Strategies

This meta-analysis uses the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for comprehensively classifying qualified research. To identify studies related to the meta-analysis, the search process used six (6) electronic databases in Harzing's Publish or Perish 8, namely Cross-Reference (CrossRef), Education Resources Information Center (ERIC), Google Scholar, ProQuest, OpenAlex, and ResearchGate, where search strings were developed, and the same keywords were used in manually searching for the articles. The results were saved as an EndNote file to be analyzed in the EndNote 21 (free trial version) app. The search specifically targeted studies that investigated the effects of problem-based learning on cognitive outcomes in chemistry education. The following keywords were used in the said

databases using various Boolean combinations: "effectiveness," "problem-based learning," “learning outcomes," “interest to learn," “problem-solving skills," “critical thinking," “academic achievement," "chemistry," and “chemistry education." To ensure full search reproducibility, the overall master Boolean search strings used were formulated as: Primary Master Search String ("problem-based learning" OR "PBL") AND ("chemistry" OR "chemistry education") AND ("academic achievement" OR "learning outcomes" OR "critical thinking" OR "problem-solving skills" OR "creative thinking" OR "mole concept" OR "thermochemistry"); and Secondary Master Search String: ("effectiveness" OR "impact" OR "comparative study") AND ("problem-based learning" OR "PBL") AND ("lecture-based learning" OR "conventional teaching") AND ("chemistry"). To select the most eligible and data-extractable articles, the current study also employs a set of eligibility criteria.

Eligibility Criteria

The meta-analysis was performed following the guidelines on PRISMA 2020 for identification, screening, eligibility, and inclusion procedures to ensure transparency and reproducibility (Page et al., 2021). The inclusion and exclusion criteria of this analysis are:

1. Studies must include students enrolled in formal chemistry education.
2. Studies should have used problem-based learning as the main teaching intervention in chemistry education, including blended or integrated PBL methods.
3. The results of the students' cognitive outcomes in the PBL classroom, such as problem-solving skills, learning outcomes, and/or critical thinking skills, should be presented.
4. Studies must be quantitative and must have a quasi-experimental or experimental research design that has data on the calculation of effect size.
5. Studies must be published in English in peer-reviewed journals with no restrictions on location.

The search was conducted using the keywords "effectiveness," “problem-based learning," “learning outcomes," “interest to learn," "problem-solving skills," “critical thinking," “academic achievement," "chemistry," and “chemistry education." The search resulted in a total of n = 3,768 records from six databases (CrossRef, ERIC, Google Scholar, ProQuest, OpenAlex, and ResearchGate). After duplicate record removal, n=99 records were excluded, leaving n=3,669 records for screening. These records were screened using EndNote to check the title, abstract, and other fields for relevant keywords. A total of n = 2,862 records were excluded because they were not aligned with education, involved samples from unrelated fields, or were not published between 2017 and 2026, leaving n = 807 records for title and abstract screening. Excluded n = 777 for being unrelated to chemistry education, not involving problem-based learning, or not written in English, leaving n = 30 that were fully read the paper for the assessment of eligibility. With these, n = 16 were excluded because they were not aligned with the study's objectives, were not experimental, or lacked data for effect size calculation. Finally, a total of n = 14 studies were included in the review. Figure 1 below shows the complete procedure undertaken by the researchers in identifying appropriate studies for analysis, following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Risk of Bias Assessment

To evaluate the methodological quality and internal validity of the 14 included studies, a domain-based appraisal framework modified from Cochrane standards was used. This framework included ROBINS-I (Risk Of Bias In Non-randomized Studies of Interventions) for quasi-experimental classroom designs and Cochrane Risk of Bias 2 (RoB 2) for randomized designs. Each study was independently assessed by two reviewers in five important methodological domains: (1) confounding bias (control of baseline characteristics and pre-test equivalence), (2) selection bias (allocation mechanism and group comparability), (3) performance bias (fidelity and consistency of instructional delivery), (4) attrition bias (completeness of outcome data), and (5) measurement bias (reliability and validity of the cognitive assessment instruments). Each domain was categorized as presenting a low, moderate, or high risk of bias. Disagreements between reviewers were resolved through collaborative discussion and consensus.

Data Analysis

The data extracted from the 14 included articles were analyzed using Meta-Essentials v.1.5. Means, standard deviations (SD), and sample sizes (N) of the posttest scores of the experimental group (PBL) and the control group (conventional method) were used to determine the effect sizes for each study. A p-value of <0.05 was considered significant for all analyses.

The fixed-effects model and the random-effects model, two common and popular statistical models, were used in this meta-analysis. The fixed-effects model was applied if the pooled data were homogeneous; otherwise, the random-effects model was applied. To determine the heterogeneity, the I^2 test was used. If the Cochran's Q statistic p-value was < 0.10, then the data were considered heterogeneous.

Another area for analysis is publication bias, which occurs if the researchers selectively include only studies that have favorable or statistically significant results. Wherein funnel plot inspection was used to assess publication bias, and to statistically validate the funnel plot's asymmetry, Egger's regression test was used.

In this meta-analysis, the researchers examined two subgroup variables: the cognitive outcome level and continental location. Cognitive outcome level was divided into two groups: academic achievement and higher-order thinking skills, which covered critical thinking and problem-solving skills. While the continental location was categorized into two groups, the first group was Africa, which comprised studies from Nigeria, South Africa, Zambia, and Ghana; the second was Asia & the Middle East, which comprised studies from Turkey, Pakistan, the Philippines, Indonesia, and Lebanon.Finally, moderator analysis was conducted to examine if the class size influenced the problem-based learning on students' cognitive outcomes. Then the class size was considered as a continuous moderator variable to determine whether studies with larger or smaller samples produced significantly different effect sizes.

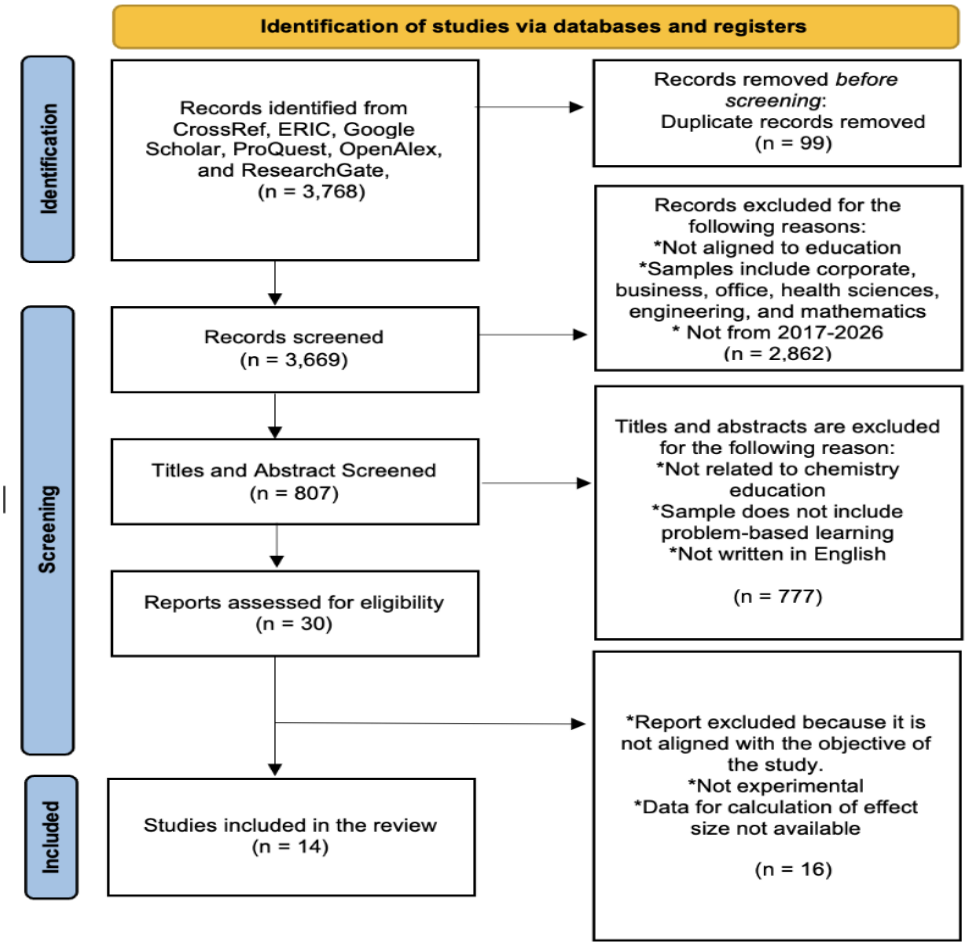


Figure 2. The Flow Diagram of Article Selection

RESULTS AND DISCUSSIONS

Characteristics of Studies Included

In this meta-analysis review, there are 14 articles included from 2017 to 2026. All of the studies used either experimental or quasi-experimental designs that compared Problem-Based Learning and the conventional or traditional method. The sample sizes across studies ranged from 36 to 215 total participants, and the total sample size was 1,299 (experimental group = 638 and control group = 661).

Table 1. Characteristics of Included Studies

Included Studies	Characteristics										
	Study Design	Country	Intervention	Comparator	Outcome Measure	Group Sizes		Posttest Means		Standard Deviation	
						Experimental Group	Control Group	Experimental Group	Control Group	Experimental Group	Control Group
Abubakar et al. (2026)	Pure Experimental	Ghana	Problem-Based Learning	Lecture-Based Learning	Academic Achievement	18	18	32	23.11	7.88	5.15
Adah (2022)	Quasi-experimental	Nigeria	Problem-Based Learning	Conventional Lecture	Higher Order Thinking Skills	59	66	48.11	36.76	11.13	13.36
Aidoo et al. (2017)	Quasi-experimental	South Africa	Problem-Based Learning	Traditional Lecture	Academic Achievement	51	51	31.76	19.98	4.394	7.279
Awan et al. (2017)	Quasi-experimental	Pakistan	Problem-Based Learning	Traditional Lecture	Academic Achievement	35	28	25.29	22.75	5.061	6.227
Ayyildiz & Tarhan (2018)	Quasi-experimental	Turkey	Problem-Based Learning	Traditional Teaching	Academic Achievement	21	20	87.23	72.14	12.75	9.35
Charif (2017)	Quasi-experimental	Lebanon	Problem-Based Learning	Conventional Teaching	Academic Achievement	26	27	88.27	79.89	14.21	14.74
Condong et al. (2026)	Quasi-experimental	Indonesia	Problem-Based Learning STEM	Conventional Teaching	Higher Order Thinking Skills	35	29	80.8	65.38	9.695	14.843
Dakabesi & Luoise (2019)	Quasi-experimental	Indonesia	Problem-Based Learning	Conventional Learning	Higher Order Thinking Skills	60	64	37.55	31.84	14.13	13.09
Jack (2022)	Quasi-experimental	Nigeria	Problem-Based Learning	Lecture-Based Learning	Academic Achievement	45	50	33.96	28.27	7.22	6.63
Jimoh & Fatokun (2020)	Quasi-experimental	Nigeria	Problem-Based Learning	Conventional Method	Academic Achievement	58	52	32.53	24.85	1.71	2.43

Onyi et al. (2022)	Quasi-experimental	Nigeria	Problem-Based Learning	Traditional Lecture	Academic Achievement	110	105	60.75	58.38	7.35	6.56
Simunza et al. (2020)	Quasi-experimental	Zambia	Problem-Based Learning	Traditional Method	Academic Achievement	45	50	61.58	53.74	7.838	11.517
Valdez & Bungihan (2019)	Experimental	Philippines	Problem-Based Learning	non-PBL	Higher Order Thinking Skills	46	50	0.79	0.4	0.31	0.16
Zudonu et al. (2020)	Quasi-experimental	Nigeria	Problem-Based Learning	Conventional Lecture	Academic Achievement	39	41	20.44	13.34	4.746	4.633

Risk of Bias Assessment for the Included Studies

The overall methodological quality across the 14 included experimental and quasi-experimental studies was evaluated as acceptable for classroom-based educational research. Ten studies (71.4%) had an overall Low Risk of bias, as shown in Table 2, while four studies (28.6%) had a Moderate Risk of bias, mainly because of non-randomized cluster allocation and possible baseline confounding present in intact classroom settings. There was no study that was deemed to have a high risk of bias. Because it is logistically impossible to double-blind teachers and high school students during instructional delivery in real-world school settings, performance bias was consistently rated as moderate across multiple quasi-experimental studies.

Table 2. Risk of Bias Assessment for the Included Studies (N = 14)

Included Study (Authors & Title)	Confounding Bias	Selection Bias	Performance Bias	Attrition Bias	Measurement Bias	Overall Risk of Bias
Abubakar et al. (2026) Comparative study of problem-based learning (PBL) and lecture-based learning (LBL) approaches in balancing redox equations among senior secondary school chemistry students	Low	Low	Moderate	Low	Low	Low Risk
Adah (2022) Effect of problem-based learning on acquisition of creative thinking skills among chemistry students in senior secondary schools	Moderate	Moderate	Moderate	Low	Low	Moderate Risk
Aidoo et al. (2017) Effect of problem-based learning on students' achievement in chemistry	Low	Moderate	Moderate	Low	Low	Low Risk
Awan et al. (2017) Effects of problem based learning on students' critical	Moderate	Moderate	Moderate	Low	Low	Moderate Risk

thinking skills, attitudes towards learning and achievement in chemistry						
Ayyildiz & Tarhan (2018) Problem-based learning in teaching chemistry: enthalpy changes in systems	Low	Low	Moderate	Low	Low	Low Risk
Charif (2017) Problem-based versus conventional teaching in chemistry education	Low	Moderate	Moderate	Low	Low	Low Risk
Condong et al. (2026) STEM-integrated problem-based learning as a cognitively aligned framework for enhancing critical thinking in thermochemistry	Low	Low	Low	Low	Low	Low Risk
Dakabesi & Luoise (2019) The effectiveness of problem-based learning model to increase the students' critical thinking skills	Moderate	Moderate	Moderate	Low	Low	Moderate Risk
Jack (2022) Problem-based versus lecture-based learning environments: Effect on students' attitude and performance in chemistry	Low	Moderate	Moderate	Low	Low	Low Risk
Jimoh & Fatokun (2020) Effect of problem-based learning strategy on chemistry students' achievement and interest in mole concept	Low	Moderate	Moderate	Low	Low	Low Risk
Onyi et al. (2022) Efficacy of problem based learning in promoting high achievement of students in chemistry	Low	Low	Moderate	Low	Low	Low Risk
Simunza et al. (2020) Impact of problem-based learning on	Low	Moderate	Moderate	Low	Low	Low Risk

learner achievement and problem-solving skills on environmental non-metal pollution						
Valdez & Bungihan (2019) Problem-based learning approach enhances the problem solving skills in chemistry of high school students	Low	Low	Low	Low	Low	Low Risk
Zudonu et al. (2020) Appraisal of effect of problem-based learning strategy on chemistry students' academic achievement in mole concept	Moderate	Moderate	Moderate	Low	Low	Moderate Risk

Note: To evaluate the risk of bias, a modified domain-based framework based on Cochrane RoB 2 (for randomized designs) and ROBINS-I (for quasi-experimental designs) was used. In classroom quasi-experiments where teaching interventions cannot be double-blinded, performance bias is inherently moderate.

Effect size

Table 3. Pooled Effect Size

Combined Effect Size	
Hedges' g	1.18
Standard Error	0.23
Cl Lower Limit	0.69
Cl Upper Limit	1.68
PI Lower Limit	-0.45
PI Upper Limit	2.82
Z-Value	5.22
One-tailed p-value	0.000
Two-tailed p-value	0.000

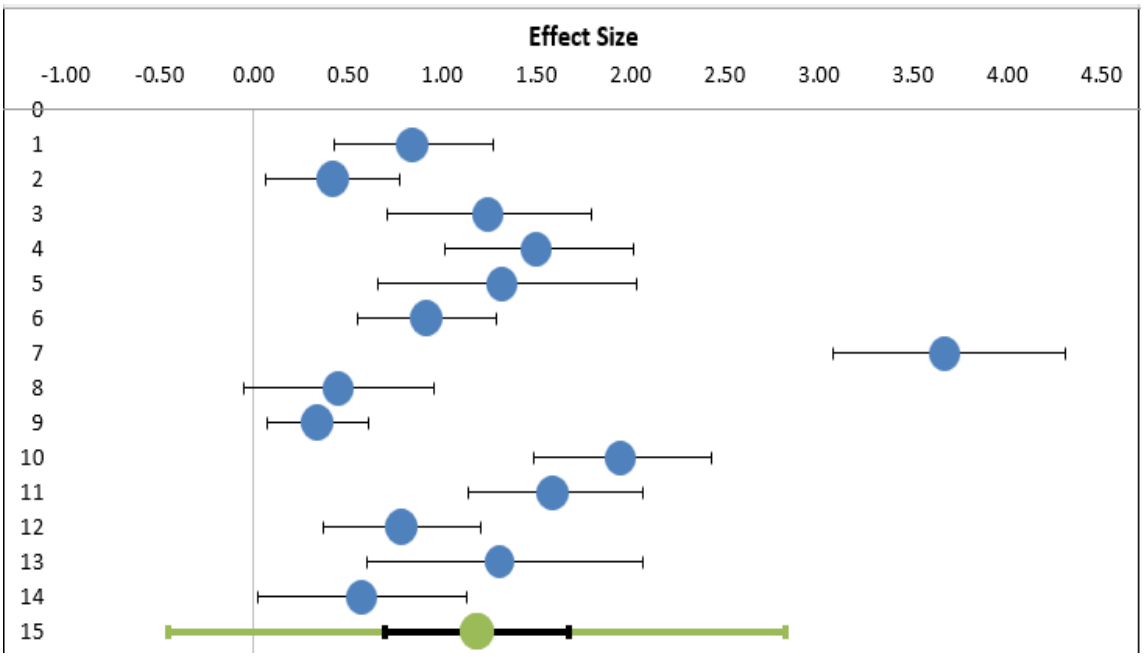


Figure 3. Forest Plot

The meta-analysis was conducted using Meta-Essentials v.1.5, specifically Workbook 3, which is designed to analyze effect sizes of two-group posttest studies (Van Rhee et al.,2015). A random-effects model meta-analysis was used to investigate the effect of problem-based learning on students' cognitive outcomes across fourteen included articles involving 1,299 participants. A 95% confidence interval of 0.69 to 1.68 indicates that the true average effect of problem-based learning is likely located within this range. This range is found entirely on the positive side of the plot, indicating strong evidence that problem-based learning positively improves students' cognitive outcome. As shown in Table 3, this statistically significant pooled effect (Hedges' $g = 1.18$, $Z = 5.22$, $p < .001$) is visually anchored in Figure 3 by the large green circle at the bottom of the forest plot. The position of this circle aligns precisely with the 1.18 value on the x-axis, while the intersecting thick black horizontal bar corresponds to the confidence interval limits of 0.69 to 1.68 reported in the table.

Based on a 95% prediction interval, the true effect of the intervention in a future study is expected to fall from -0.45 to 2.82. This interval ranges from a potential negative effect to a very large positive effect, which indicates that future studies may observe large variations in effect sizes. This wide dispersion is visually represented in Figure 3 by the extended light green bar that clearly crosses the vertical zero line, reflecting the substantial heterogeneity among the individual studies (blue circles), some of which show negligible or negative effects while others, like Study 7, demonstrate massive positive effects.

Heterogeneity test

Table 4. Heterogeneity Test

Heterogeneity	
Q	114.12
PQ	0.000
I^2	90.98%
T^2	0.52
T	0.72

The heterogeneity test results presented in Table 4 indicate substantial and statistically significant variability among the included studies. The Cochran's Q statistic is 114.12 with a corresponding p-value (PQ) of 0.000, which is well below the conventional significance threshold of .05. This rejects the null hypothesis of homogeneity, confirming that the true effect sizes are not identical across studies and that the observed dispersion is greater than what would be expected from sampling error alone. Because the Q statistic can be influenced by the number of studies and lacks a direct metric for the magnitude of inconsistency, the I^2 statistic is the preferred measure for quantifying the extent of heterogeneity. The I^2 value of 90.98% indicates that nearly 91% of the total variation in effect estimates is attributable to genuine between-study heterogeneity rather than random chance. According to the Cochrane Handbook's guidelines, an I^2 value in the range of 75% to 100% represents considerable heterogeneity, meaning the inconsistency among study results is substantial and the pooled estimate should be interpreted with significant caution (Deeks et al., 2024). This high I^2 value indicates that nearly 91% of the variance in effect sizes is due to genuine between-study differences rather than sampling error. The between-study variance ($T^2 = 0.52$, $T = 0.72$) and the wide prediction interval (-0.45 to 2.82) reinforce that the true effect varies considerably across contexts. Consequently, the pooled effect size should be interpreted with caution, and future research should conduct subgroup analyses or meta-regressions to identify potential moderators driving this variability.

Subgroup Analysis

Cognitive Outcome Level

A subgroup analysis was conducted to examine whether the effect of problem-based learning differs significantly according to the type of cognitive outcome measured. The overall meta-analytic model, encompassing 14 studies and 1,299 subjects, yielded a pooled effect size of Hedges' g = 1.15 (95% CI = -1.02 to 3.31). The result between subgroups shows that there was no statistically significant difference (Qbetween=0.09, df=1, p=0.763), indicating that the intervention produced comparable effects across the different types of cognitive outcomes measured. The residual heterogeneity that remained within the subgroups (Qwithin=13.39, df=12, p=0.341) suggests that other factors aside from cognitive outcome type have contributed to the variability in effect sizes. Additionally, the subgroup variable explained only 0.67% of the between-study variance (R^2=0.67%), indicating that cognitive outcome type had a negligible moderating effect on the effectiveness of problem-based learning.

Table 5. Subgroup Analysis for Type of Cognitive Outcome Measured

Analysis of Variance	Sum of Squares	df	p
Between / Model	0.09	1	0.763
Within / Residual	13.39	12	0.341
Total	13.48	13	0.411

Pseudo R^2	0.67%

Continental Location

Another subgroup analysis was conducted to investigate whether the continental location moderated the effect of problem-based learning on students' cognitive outcomes. For the studies conducted in Africa, the pooled effect size indicated a statistically significant and very large effect (Hedges' g=1.39, 95% CI=0.68 – 2.10). For the studies conducted in Asia and the Middle East, the pooled effect size also indicated a statistically significant and large effect (Hedges' g=0.91, 95% CI=0.50 – 1.33). Both groups also exhibited considerable heterogeneity.

Table 6. Subgroup Analysis for Continental Location

Analysis of Variance	Sum of Squares	df	p
Between / Model	1.49	1	0.223
Within / Residual	13.56	12	0.330
Total	15.04	13	0.305

Pseudo R^2	9.87%

However, the difference between the two continental locations was not statistically significant (Qbetween=1.49, df=1, p=0.223). The continental location explained 9.87% of the between-study variance (Pseudo R^2=9.87%), indicating that continental location had a negligible moderating effect on the effectiveness of problem-based learning. The residual heterogeneity that remained within the subgroups (Qwithin=13.56, df=12, p=0.330) suggests that other factors aside from continental location have contributed to the variability in effect sizes.

Moderator analysis

Class Size

A meta-regression analysis was conducted to examine whether class size significantly predicts the effect of the intervention on students' cognitive outcomes. The overall model was not statistically significant (Q=0.32, df=1, p=0.572; F=0.24, p=0.630). While the intercept was significant (B=1.43, p=0.003), indicating a significant baseline effect when the moderator is zero, the moderator itself was not a significant predictor (B=-0.00262, 95% CI=-0.01 to 0.01, p=0.572). The moderator explained only 2.00% of the between-study variance (R^2=2.00%), indicating a negligible moderating effect. Furthermore, the residual heterogeneity remained high (T^2=0.52), suggesting that other factors aside from sample size have contributed to the variability in effect sizes.

Table 7. Moderator Analysis on Class Size

Statistic	Value
Regression coefficient (β)	-0.14
F-value	0.24
p-value	0.630
R^2	2.00%

Publication Bias Analysis

A funnel plot was generated to visually assess publication bias and small-study effects. As shown in the figure 4, the individual study effect sizes (blue dots) exhibit considerable dispersion, which is expected given the substantial heterogeneity identified in the meta-analysis (I^2 = 90.98%).

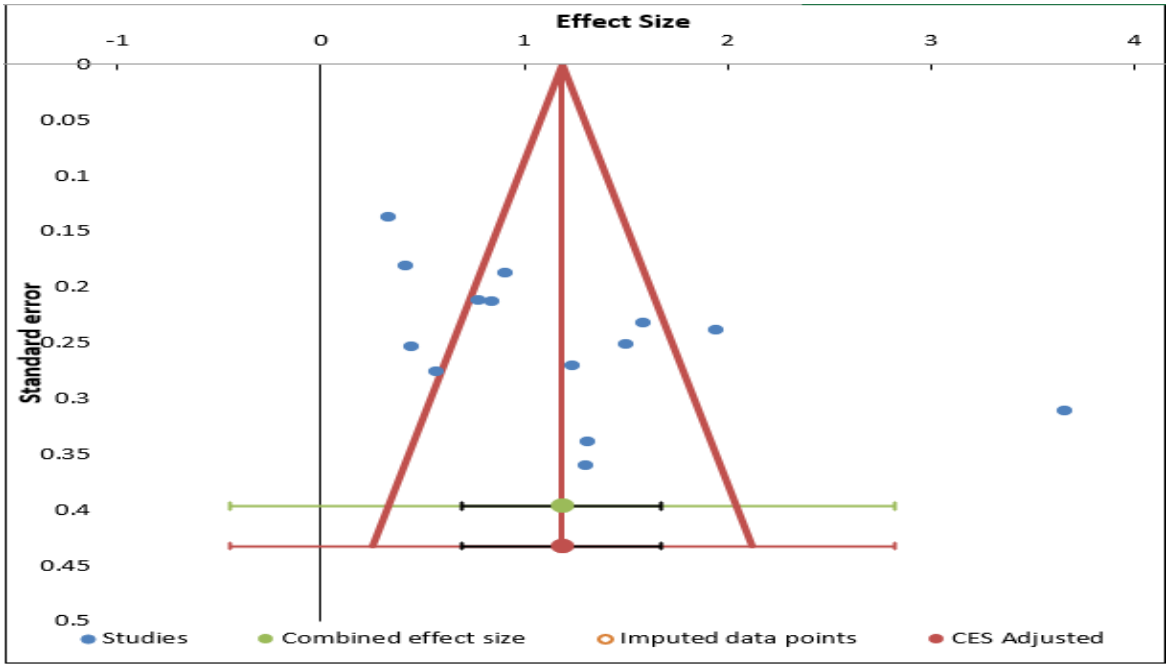


Figure 4. Funnel plot for the assessment of publication bias.

To statistically evaluate this asymmetry, a Trim-and-Fill analysis was conducted. As shown in Table 8, the estimator searched from the mean to the left and identified 0 missing studies. Consequently, the adjusted effect size (Hedges' g = 1.18, 95% CI = [0.69, 1.68]) is identical to the observed effect size. This indicates that no hypothetical missing studies were required to make the funnel plot symmetrical.

Table 8. Trim-and-Fill Analysis for Publication Bias

Analysis Type	Hedges' g	Standard Error	95% Cl		95% PI		Imputed Studies
			Upper	Lower	Upper	Lower	
Observe	1.18	0.23	1.68	0.69	2.82	-0.45	-
Adjusted	1.18	0.23	1.68	0.69	2.82	-0.45	0

Moreover, publication bias was also assessed using Egger's regression test. As shown in Table 9, the intercept was not statistically significant (estimate = 20.25, *SE* = 10.65, 95% CI = -2.75 to 43.26, t-test = 1.90, p-value = .081). Since the p-value exceeds the conventional threshold of .05, there is no significant evidence of funnel plot asymmetry. Consequently, the pooled effect size is unlikely to be substantially distorted by small-study effects or publication bias.

Table 9. Egger's Regression Test for Publication Bias

Variable	Estimate	Standard Error	95% Cl		t-test	p-value
			Upper	Lower		
Intercept	20.25	10.65	43.26	-2.75	1.90	0.081
Slope	-14.30	8.15	3.30	-31.90	-	-

SUMMARY OF FINDINGS

The meta-analysis, encompassing 14 experimental and quasi-experimental studies with a total of 1,299 participants, demonstrates that Problem-Based Learning (PBL) yields a large, statistically significant positive effect on students' cognitive outcomes in chemistry compared to conventional instructional methods (Hedges' g = 1.18, 95% CI [0.69, 1.68], Z = 5.22, p-value < .001). This robust pooled effect is visually anchored in the forest plot, where the confidence interval lies entirely on the positive side of the axis. However, the analysis revealed substantial and statistically significant heterogeneity across the included studies (Q = 114.12, p-value = .000; I^2 = 90.98%; T^2 = 0.52), indicating that nearly 91% of the variance in effect sizes stems from genuine between-study differences rather than sampling error. Consequently, the 95% prediction interval (-0.45 to 2.82) crosses zero, suggesting that while the average effect is strongly positive, future implementations in new contexts could vary considerably, ranging from negligible or negative to highly positive.

To explore potential sources of this variability, subgroup and moderator analyses were conducted. The subgroup analysis for cognitive outcome level (Academic Achievement versus Higher-Order Thinking Skills) revealed no statistically significant between-group difference (Q between = 0.09, df = 1, p-value = 0.763; Pseudo R^2 = 0.67%), indicating that PBL produces comparable effects across both types of cognitive outcomes. Similarly, the subgroup analysis for continental location (Africa versus Asia & the Middle East) showed no significant difference between regions (Q between = 1.49, df = 1, p-value = 0.223; Pseudo R^2 = 9.87%), with both regions demonstrating significant positive effects (Africa: Hedges' g = 1.39; Asia & Middle East: Hedges' g = 0.91). Furthermore, the meta-regression examining class size as a continuous moderator was not statistically significant (Q = 0.32, p-value = 0.572; F = 0.24, p-value = 0.630; R^2 = 2.00%), indicating that class size does not significantly predict the magnitude of PBL's effectiveness.

Assessment of publication bias yielded no significant concerns, as the Trim-and-Fill analysis identified zero missing studies (adjusted effect size identical to the observed Hedges' g = 1.18), and Egger's regression test was not statistically significant (intercept = 20.25, p-value = .081). Collectively, these findings suggest that PBL is a highly effective instructional approach for chemistry education across diverse geographical contexts and class sizes. However, the high residual heterogeneity and the wide prediction interval underscore that PBL is not a one-size-fits-all solution. To maximize higher-order thinking gains, curriculum developers and educators should pair PBL with structured scaffolds and ensure adequate teacher training, contextualization of materials, and adaptation to specific classroom goals to ensure consistent effectiveness across all learning environments.

CONCLUSION

Problem-based learning (PBL) has been one of the widely utilized types of instructional methods grounded in constructivist theory. Numbers of research has been used to compare the efficiency of problem-based learning to conventional methods with cognitive outcomes as a predictor variable. This meta-analysis analyzed 14 studies to provide a clear contrast of problem-based learning and conventional learning in improving students' cognitive outcomes. To conclude, problem-based learning (PBL) was concluded to be significantly effective in improving the cognitive outcomes of students in chemistry education better than conventional learning. This effect was found to be statistically consistent across both cognitive outcome level and continental location, with no significant difference in the magnitude of PBL's effectiveness between academic achievement and higher-order thinking skills (p = .763), or between studies conducted in Africa and in Asia & Middle East (p = .223). Similarly on the class size, it was found not to be a significant moderator of this relationship, which indicates that PBL's effectiveness does not depend on the number of learners involved in a study. On the other hand, class size was found not to be a significant moderator of the relationship, which indicates that problem-based learning's effectiveness doesn't depend on the number of learners involved in a study.

Educational implications

The findings of this meta-analysis carry several inferences for chemistry teaching, curriculum design, and future research. One of them is the suggestion for chemistry teachers to adopt problem-based learning as a primary instructional strategy instead of a supplementary activity. This is statistically supported by the significant pooled effect that was discussed in this review, where PBL continuously outperformed conventional, lecture-based instruction in improving students' cognitive outcomes, especially for subjects like thermodynamics, chemical equilibrium, and enthalpy changes that have been linked and recognized to common misconceptions.

For curriculum developers, this review's findings show that PBL's effectiveness is not just limited in a single type of cognitive outcome. The subgroup analysis showed that PBL produced comparable, statistically significant effects on both academic achievement and higher-order thinking skills such as critical thinking and problem-solving, with no significant difference between the two ($p = .763$). This suggests that PBL can be integrated into chemistry curricula to support both content mastery and higher-order skill development simultaneously, without the need of choosing one outcome over the other. Pairing PBL with structured scaffolds, such as reflection activities, can help maximize gains in both areas. PBL can be implemented across diverse educational systems; it just needs to have a specific context as an implementation support. Therefore, to strengthen its impact beyond content-based learning, pairing PBL with additional scaffolds, such as structured reflection activities or precise critical-thinking frameworks, may help. PBL can be implemented across diverse educational systems; it just needs to have a specific context as an implementation support.

Since PBL produced significant effects in both Africa and Asia & the Middle East, it appears to be effective across different countries. However, the higher heterogeneity found among African studies compared to Asia & Middle East studies shows that outcomes may vary more within some regional contexts than others, which is consistent with findings that PBL's effectiveness depends heavily on different factors such as teacher readiness, contextualization of learning materials, and instructional design (Affandy et al., 2026). With this, education stakeholders, especially in regions that show higher differences, should consider strengthening teacher training and localizing PBL materials.

Another implication is that class size should not be treated as a barrier to adopting PBL, since sample size did not significantly predict the magnitude of PBL's effect on cognitive outcomes, making it a possible instructional approach even in resource-constrained settings or with a large number of students. Given the substantial heterogeneity observed across studies overall, every teacher training program and school administrator should avoid having a one-size-fits-all approach to PBL and emphasize how to adapt PBL design and delivery to the specific goals and context of each classroom, instead of assuming the same outcomes across all applications.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

J.M. Jimenez and K. Napay conducted the research, analyzed the data, and wrote the paper.

Declaration of Generative AI and AI-assisted technologies in the writing process.

During the preparation of this work, the authors used Claude AI to guide the development of the research outline. Furthermore, the authors utilized Grammarly to enhance the quality of their writing. Following the use of these tools, the authors carefully examined and modified the content as needed and accepted full responsibility for the published material.

REFERENCES

1. Abubakar, M., Abukari, M. A., Yakubu, A., & Najah, A. N. (2026). Comparative study of problem-based learning (PBL) and lecture-based learning (LBL) approaches in balancing redox equations in senior high school. Education Journal, 15(1), 35–46. Retrieved from https://doi.org/10.11648/j.edu.20261501.15
2. Adah, S. A. (2022). Effect of problem-based learning on acquisition of creative thinking skills among chemistry students in Ogoja Education Zone, Cross River State, Nigeria. International Journal of Education and Evaluation, 8(7), 1–10. Retrieved from https://doi.org/10.56201/ijee.v8.no7.2022.pg1.10
3. Affandy, H., Sunarno, W., Suryana, R., & Harjana. (2026). Unveiling the global instructional impact of problem-based learning in science education: A comprehensive systematic review. Interdisciplinary Journal of Problem-Based Learning, 20(1). Retrieved from https://doi.org/10.14434/ijpbl.v20i1.37007
4. Ahmar, Dewi & Azzajjad, Muhammad & Syahrir, Muh. (2020). Students' Representation Ability in Chemistry. Journal of Applied Science, Engineering, Technology, and Education. 2. 181-187. 10.35877/454RI.asci22124.
5. Aidoo, B., Boateng, S. K., Kissi, P. S., & Ofori, I. (2017). Effect of problem-based learning on students' achievement in chemistry. Journal of Education and Practice, 7(33), 103–108.
6. Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.
7. Arita Marini, Uyu Muawanah, Asep Marfu, (2026). Enhancing critical thinking through problem-based learning: The role of student engagement and technology for education sustainability in Indonesia, Sustainable Futures, Volume 11,2026,101846, ISSN 2666-1888, Retrieved from https://doi.org/10.1016/j.sftr.2026.101846.
8. Awan, R. u. N., Hussain, H., & Anwar, N. (2017). Effects of problem based learning on students' critical thinking skills, attitudes towards learning and achievement. Journal of Educational Research, 20(2), 28–41.
9. Ayyildiz, Yildizay & Tarhan, Leman. (2018). Problem-based learning in teaching chemistry: enthalpy changes in systems. Research in Science and Technological Education. 36. 35-54. 10.1080/02635143.2017.1366898.
10. Babale Abdullahi, Lawal Aisha Yusuf, Ibrahim Aliyu Mohammed. (2019). A Study of the Effects of Computer-Assisted Packages on the Performance of Senior Secondary Students in Chemistry in Zaria Educational Zone, Kaduna State Nigeria. International Journal of Secondary Education, 6(4), 59-65. Retrieved from https://doi.org/10.11648/j.ijsedu.20180604.12
11. Baran, M., Sozbilir, M. An Application of Context- and Problem-Based Learning (C-PBL) into Teaching Thermodynamics. Res Sci Educ 48, 663–689 (2018). Retrieved from https://doi.org/10.1007/s11165-016-9583-1.
12. Bolarinwa, O. T., Abubakar, A., Akintokun, O. A., Adedoyin, S. A., & Adeoye, B. A. (2025). Effect of problem-based learning on secondary school chemistry students' achievement in the law of electrolysis, Oyo Town. International Journal of Innovative Science and Research Technology, 10(5), 3963–3969. Retrieved from https://doi.org/10.38124/ijisrt/25may2177
13. Condong, D. M., Sihaloho, M., Thayban, T., Laliyo, L. A. R., Musa, W. J. A., Kurniawati, E., Munandar, H., & Nggole, S. I. (2026). STEM-integrated problem-based learning as a cognitively aligned framework for enhancing critical thinking in thermochemistry. Jambura Journal of Educational Chemistry, 8(1), 120–128. Retrieved from https://doi.org/10.37905/jjec.v8i1.37514
14. Dakabesi, D., & Yosephine Luoise, I. S. (2019). The effectiveness of problem-based learning model to increase the students' critical thinking skills. Journal of Education and Learning (EduLearn), 13(4), 543–549. Retrieved from https://doi.org/10.11591/edulearn.v13i4.12940
15. Deeks, J. J., Higgins, J. P. T., Altman, D. G., McKenzie, J. E., & Veroniki, A. A. (2024). Chapter 10: Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), Cochrane handbook for systematic reviews of interventions (Version 6.5). Cochrane. Retrieved from https://training.cochrane.org/handbook/current/chapter-10
16. Dewey, J. (1923). Democracy and education: an introduction to the philosophy of education. New York: McMillan. Retrieved from

https://books.google.com.ph/books/about/Democracy_and_Education.html?id=WGEWAAAAIAAJ&redir_esc=y

17. Funa, A. A., & Prudente, M. S. (2021). Effectiveness of problem-based learning on secondary students' achievement in science: A meta-analysis. International Journal of Instruction, 14(4), 69-84. Retrieved from https://doi.org/10.29333/iji.2021.1445a
18. Jack, G. U. (2022). Problem-based versus lecture-based learning environments: Effect on students' attitude and performance in chemistry at secondary school level. Unizik Journal of Educational Research and Policy Studies, 14(2).
19. Jimoh, S. B., & Fatokun, K. V. F. (2020). Effect of problem-based learning strategy on chemistry students' achievement and interest in mole concept in Federal Capital Territory, Abuja. Anchor University Journal of Science and Technology, 1(1), 136–143.
20. Monika, Putri & Suharno, Suharno & Rahmasari, Lita. (2023). Effectiveness of Science Technology Engineering Mathematics Problem Based Learning (STEM PBL) and Science Technology Engineering Mathematics Project Based Learning (STEM PjBL) to Improve Critical Thinking Ability. Jurnal Penelitian Pendidikan IPA. 9. 9593-9599. 10.29303/jppipa.v9i11.4910.
21. Muyambo - Goto, O., Naidoo, D., & Kennedy, K. J. (2023). Students' conceptions of 21st-century education in Zimbabwe. Interchange, 54(1), 49–80. Retrieved from https://doi.org/10.1007/s10780-022-09483-3
22. Onyi, I. E., Njoku, Z. C., & Nwafor, S. C. (2022). Efficacy of problem based learning in promoting high achievement of students in chemistry. Journal of Critical Reviews, 9(1), 49–58.
23. Reyes, R. L. (2025). Integrating real-world problems into chemistry curricula: enhancing relevance and student engagement. Forum for Education Studies, 3(2), 2177. Retrieved from https://doi.org/10.59400/fes2177
24. Rosyidah*, K., Lutfi, A., Sanjaya, I., & Astutik, J. (2024). Identification of Students' Misconceptions and Understanding on Thermochemistry Material with Four-Tier Multiple-Choice Tests. Jurnal Pendidikan Sains Indonesia, 12(1), 155-171. Retrieved from https://doi.org/10.24815/jpsi.v12i1.34899
25. Sappaile, Baso & Dhaksietsatsura, Virzha & Dewi, Eka & Pinanta, Raden & Yuliyanti, Yuliyanti. (2025). Problem-Based Learning (PBL) in Action: Fostering Critical Thinking Among Middle School Students. International Journal of Educational Research Excellence (IJERE). 4. 446-453. 10.55299/ijere.v4i2.1253.
26. Simunza, J., Banda, A., & Hayumbu, P. (2020). Impact of problem-based learning on learner achievement and problem-solving skills on environmental non-metal pollution. International Journal of Research and Innovation in Social Science, 4(8), 522–533.
27. Siregar, Torang. (2025). Article Research Problem-Based Learning (PBL) Strategy in Enhancing Students' Critical Thinking Skills. The New Educator. 2. 1-26.
28. Sternberg, R. J. (1998). Abilities are forms of developing expertise. Educational Researcher, 27(3), 11-20.
29. Su, T., Liu, J., Meng, L., Luo, Y., Ke, Q., & Xie, L. (2025). The effectiveness of problem-based learning (PBL) in enhancing critical thinking skills in medical education: A systematic review and meta-analysis. Frontiers in Education, 10, 1565556. Retrieved from https://doi.org/10.3389/feduc.2025.1565556
30. Valdez, J. E., & Bungihan, M. E. (2019). Problem-based learning approach enhances the problem solving skills in chemistry of high school students. Journal of Technology and Science Education, 9(3), 282–294. Retrieved from https://doi.org/10.3926/jotse.631
31. Van Rhee, H. J., Suurmond, R., & Hak, T. (2015). User manual for Meta-Essentials: Workbooks for meta-analysis (Version 1.5). Erasmus Research Institute of Management. Retrieved from www.erim.eur.nl/research-support/meta-essentials
32. Varas, D., Santana, M., Nussbaum, M., Claro, S., & Imbarack, P. (2023). Teachers' strategies and challenges in teaching 21st-century skills: Little common understanding. Thinking Skills and Creativity, 48, 101289. Retrieved from https://doi.org/10.1016/j.tsc.2023.101289
33. Zudonu, O. C., Ikwut, E. F., Ekpeno, J. N., & Onyije, C. (2020). Appraisal of effect of problem-based learning strategy on chemistry students' academic achievement in senior secondary two in mole concept in Ahoada West Local Government Area, Rivers State. Chemistry and Materials Research, 12(6), 27–34. Retrieved from https://doi.org/10.7176/CMR/12-6-05

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title Page (Page 1)
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Abstract (Page 1)
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction (Pages 1–3)
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction – A. General and Specific Objectives (Page 3)
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Methodology – B. Eligibility Criteria (Pages 4–5)
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methodology – Introduction & Subsection A (Pages 4–5)
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Methodology – A. Data sources and search strategies (Pages 4–5)
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Methodology – A, B, & Figure 2 Flow Diagram (Pages 4–6)
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Methodology – D. Data Analysis (Page 6)
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Introduction – B. Conceptual framework & Methodology (Pages 2–3 & 6)
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Introduction – B. Conceptual framework & Methodology (Pages 2–3 & 6)
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Methodology – C. Risk of Bias Assessment (Page 5)
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Methodology – D. Data Analysis & Results – Table 3 (Pages 6 & 13)
Synthesis methods	13a	Describe the processes used to decide which studies were	Results – A.

		eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Characteristics of Studies Included & Table 1 (Pages 7–10)
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Methodology – D. Data Analysis (Page 6)
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Results – Table 1, Table 3, & Figure 3 (Pages 7, 13, & 14)
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Methodology – D. Data Analysis & Results – Effect Size (Pages 6, 13–15)
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Methodology – D & Results – Subgroup Analysis (Pages 6 & 15–16)
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Results – Moderator Analysis (Class Size) (Page 16)
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Methodology – D & Results – F. Publication Bias Analysis (Pages 6 & 16–17)
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Results & Discussion – Heterogeneity Test & Summary of Findings (Pages 14–15 & 17)
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Results – Figure 2 Flow Diagram of Article Selection (Pages 5–6)
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Methodology – B. Eligibility criteria & Figure 2 (Pages 5–6)
Study characteristics	17	Cite each included study and present its characteristics.	Results – Table 1 Characteristics of Included Studies (Pages 7–10)
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Results – Table 2 Risk of Bias Assessment (Pages 10–13)
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Results – Table 1, Table 3, & Figure 3 (Pages 7–14)
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Results – A & B Text Summaries (Pages 7 & 10)
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Results – Effect Size, Heterogeneity Test, & Tables 3 & 4 (Pages 13–15)
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Results – Subgroup & Moderator Analyses

			(Pages 15–16)
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Results – F. Publication Bias Analysis / Trim-and-Fill (Pages 16–17)
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Results – F. Publication Bias Analysis / Egger's test (Page 17)
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Results & Discussion – Summary of Findings (Pages 17–18)
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Results & Discussion – Summary of Findings (Pages 17–19)
	23b	Discuss any limitations of the evidence included in the review.	Results & Discussion – Heterogeneity & Limitations (Pages 14–17)
	23c	Discuss any limitations of the review processes used.	Discussion – Educational Implications (Page 19)
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion – Educational Implications (Pages 18–19)
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Not registered (Standard independent review format)
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Not registered (Standard independent review format)
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not registered (Standard independent review format)
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Unfunded / Not applicable
Competing interests	26	Declare any competing interests of review authors.	Conflict of Interest (Page 19)
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Methodology – EndNote, Meta-Essentials, data tables (Pages 4–6)