

A Systematic Review and Meta-Analysis of Learning Stoichiometry: Challenges, Interventions, and Outcomes

Wilmer Orlando López González
Experimental Sciences Program, National University of Education, Chuquipata, Azogues, Ecuador
Marco Vinicio Vásquez Bernal
Experimental Sciences Program, National University of Education, Chuquipata, Azogues, Ecuador
José Enrique Martínez Serra
Experimental Sciences Program, National University of Education, Chuquipata, Azogues, Ecuador

ABSTRACT. Stoichiometry remains a major obstacle in chemistry education because students must integrate conceptual understanding, quantitative reasoning, representational competence, and problem-solving skills to learn it successfully. In this regard, the systematic review and meta-analysis highlighted empirical studies that provide insights into stoichiometry learning issues and the instructional strategies used to teach it. The selected studies met the eligibility criteria and produced correlational evidence indicating relationships between learning factors and stoichiometry performance, as well as intervention-based evidence showing the impact of teaching on learners. First, the correlational evidence yields inconsistent findings, ranging from no relationship to moderate positive associations, focusing on factors such as motivation and self-regulation. In contrast, the intervention meta-analysis showed consistently positive results, producing a large effect size for stoichiometry learning (d = 0.83) and even larger results for critical thinking abilities (d = 1.52). However, traditional stoichiometric

knowledge showed a small effect size (d = 0.23), and this measure drew on highly diverse studies.
General stoichiometry learning outcomes and achievements varied widely, whereas critical thinking effects were consistent across studies. These results suggest that intervention strategies emphasizing problem solving, supporting the representation of ideas, incorporating technology, and offering opportunities for higher-order thinking improve stoichiometry learning outcomes. Future investigations should use standardized outcome measures, adopt tighter experimental designs, and ensure uniform reporting across studies.

Keywords: Systematic Review; Stoichiometry Learning; Meta-Analysis; Chemistry Education Research; Student Learning Outcomes.

INTRODUCTION

Stoichiometry, the quantitative relationship between reactants and products in chemical reactions, is a cornerstone of chemistry education (Gulacar, et. al., 2022; Gulacar, et. al., 2019; Hamerská et. al., 2024). It serves as a critical bridge between the macroscopic world of observable substances and the microscopic world of atoms, molecules, and ions, enabling students to predict yields, determine limiting reagents, and understand reaction mechanisms (Rosa, et al., 2022). Mastery of stoichiometry is not only needed for success in advanced chemistry courses but also for developing broader scientific literacy, as it underpins concepts in fields ranging from biochemistry to environmental science (Innocenti, & Kamarulzaman, 2012). However, despite its fundamental importance, stoichiometry is widely recognized as one of the most challenging topics for students at both the secondary and university levels (Gulacar et al., 2013). Stoichiometric problem-solving requires multiple cognitive skills, such as proportional reasoning, symbolic manipulation, and understanding the mole concept, chemical equations, and the law of conservation of mass (Horvat et al., 2016). Many students have difficulty when going from one type of representation to another, which is crucial for solving stoichiometric problems (Kozma & Russell,2005), and the fact that the mole is conceptualized as a unit, an issue that students find difficult to grasp, makes the situation even worse (Siswaningsih, et. al., 2017).

The literature has extensively documented the challenges of learning stoichiometry. For instance, students often hold misconceptions about the mole, such as treating it as a mass or a number rather than a counting unit (Siswaningsih et al., 2017). They also struggle with the algorithmic aspects of problem-solving, often relying on rote memorization of formulas without grasping the underlying chemical principles (Schmidt, & Jignéus, 2003). In

addition, while students may show proficiency in proportional reasoning (a major component of stoichiometric calculations), not all students do; therefore, many make errors when scaling quantities or interpreting reaction ratios (Ramful & Narod, 2014). In addition to the academic ramifications associated with students performing poorly on stoichiometric calculations due to difficulties with proportional reasoning, if the educational system fails to address these issues, students' motivation to succeed will be adversely affected; consequently, they could develop negative feelings toward chemistry and ultimately lead to decreased performance in subsequent courses (Flaherty, 2020). Therefore, understanding how to teach students to reason proportionally while completing stoichiometric calculations is critical for teachers and curriculum developers.

There is an abundance of research on the development of stoichiometric reasoning skills; however, there remain significant gaps in our understanding of how best to support students' development in this area. While several studies have identified various interventions (e.g., visual representations, analogies, computer simulations, scaffolded problem-solving) and evaluated their effectiveness, the evidence remains fragmented and context-dependent (García, 2025). Many studies have been conducted in a single classroom or with small sample sizes, limiting the generalizability of their results to the larger population (Anderson, Boersma, & Raubenheimer, 2004). Additionally, outcomes varied across studies (e.g., conceptual understanding, problem-solving accuracy, attitudes, self-efficacy), making it difficult to synthesize the overall impact of these interventions (Behrman et al., 2012). Furthermore, systematic investigation into moderating factors that could influence the effectiveness of these interventions remains limited (e.g., students' prior knowledge, instructional level (Beskeni et al., 2011) (e.g., high school vs. University), assessment method used) (Reyes, 2025). As a result, educators are left with a fragmented evidence base and are uncertain about which interventions are more likely to produce positive results in specific contexts.

This systematic review/meta-analysis was conducted to close the gaps in the literature on how students learn stoichiometry. It aims to provide a quantitative synthesis of studies that evaluate interventions to help students learn stoichiometry. We intend to provide a more precise estimate of the effect of instructional interventions on student learning outcomes by pooling results from multiple studies to reach a more valid overall conclusion about their effectiveness. We will also examine sources of heterogeneity to identify the characteristics that differentiate more successful interventions from less successful ones. Therefore, we aim not only to clarify what is currently known about learning stoichiometry, but also to provide an opportunity for additional research on differences between successful and unsuccessful

interventions. This systematic review/meta-analysis aims to provide evidence-based practices to guide teachers in selecting and implementing instructional interventions to help their students learn stoichiometry. By cataloging the types of challenges students have faced and the outcomes measured in these studies, we provide an opportunity for researchers to develop standardized instruments for assessing student learning in stoichiometry and to conduct future research more rigorously and consistently.

The remainder of this paper is organized as follows. Section 2 describes the methodological framework for the systematic review and meta-analysis, including the review protocol, inclusion criteria, study selection process, data extraction methods, and statistical analysis. Section 3 presents the results, starting with a description of the selected studies, heterogeneity analysis, meta-analytic synthesis of the identified outcome domains, and evaluation of stoichiometry learning outcomes, critical thinking skills, and achievement in stoichiometry learning. Section 3 also raises the question of possible publication bias from both a graphical and quantitative point of view. Section 4 discusses the results in relation to existing knowledge in chemistry education, stressing their theoretical and methodological importance. Section 5 concludes the paper by presenting the final conclusions from the systematic review and meta-analysis and emphasizing their implications for future teaching, assessment, and research.

Accordingly, this systematic review and meta-analysis aimed to clarify existing knowledge on stoichiometry learning by compiling correlational data and intervention outcomes within a single analytical framework. In particular, the review answers four connected research questions: what are the most frequently mentioned learning difficulties and teaching barriers in the existing publications on stoichiometry; what is the direction and the size of the relation between students' performance in stoichiometry and various learner-based factors and instructional conditions; to what extent, and in what way, does the application of instructional intervention strategies improve stoichiometry learning outcomes and whether there is any difference in the results of these intervention strategies related to outcomes in different aspects of learning; and how uniform is the effect of the learning intervention in studies considering publication bias and differences among research results. These questions aim to distinguish associative information from evidence of instructional effectiveness, identify unique intervention effects, and uncover the methodological and measurement conditions that drive variation in reported results.

METHODOLOGY

Review Protocol

The systematic review and the meta-analysis were conducted in accordance with the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) guidelines (Page et al., 2021). A broad-based search strategy was created to find empirical research articles on how students learn stoichiometry published between 2010 and 2025. This search included four major academic databases and search engines, each selected for its strength in covering educational research and chemical education. The first and primary database was Web of Science, which offers the most extensive coverage of high-quality, peer-reviewed journals across the sciences and social sciences, including chemical education. The second database was Scopus because of its broad interdisciplinary perspectives and strong citation-tracking capabilities, which augmented Web of Science's coverage. The third database selected was ERIC (Education Resources Information Center), as it focuses on education-related literature, ensuring that all areas of pedagogy are included in the final pool of research articles. Google Scholar was also used as an additional search method to identify relevant articles that were not indexed in the other databases. We then manually assessed all identified articles to ensure they met the predefined quality criteria.

We created custom search strings for each database, following a consistent methodology to maximize results while ensuring consistency across databases. For Web of Science, we used the following search string: TS=("stoichiometry*" AND ("learning" OR "teach*" OR "instruction*" OR "pedagogy*" OR "education" OR "student*" OR "misconception*" OR "conceptual change")) NOT TS=("review" OR "survey" OR "META-ANALYSIS" OR "SYSTEMATIC LITERATURE") AND PY=(2010-2025), allowing for search refinement by document type, allowing only articles and excluding reviews from searches. The inquiry utilized for Scopus contained the next search term: TITLE-ABS-KEY("stoichiometry*" AND ("learning" OR "teach*" OR "instruction*" OR "pedagogy*" OR "education" OR "student*" OR "misconception*")) AND NOT TITLE-ABS-KEY("review" OR "survey" OR "META-ANALYSIS" OR "SYSTEMATIC") AND PUBYEAR > 2009, and enabled methods of filtering results by article types only, with review articles and conference proceedings being removed as well as being published after 2009. The following search term was used in ERIC: ("stoichiometry"+ "chemical calculation") AND ("learning" OR "teaching methods" OR "instructional effectiveness" OR "student achievement") NOT ("LITERATURE REVIEW" OR "META-ANALYSIS" OR "SYSTEMATIC REVIEW"). This query returned articles published between 2010 and the current date; publication type was limited to journal articles and reports/research, with analyses and literature review

articles filtered out. The query used in ScienceDirect was: TITLE-ABS-KEY("stoichiometry*" AND ("learning" OR "teach*" OR "instruction*" OR "pedagogy*" OR "education")) AND NOT TITLE-ABS-KEY("review" OR "survey" OR "meta-analysis") limited by a date filter spanning 2010 -2025 and with only research article types. For Google Scholar, we used: "stoichiometry" AND ("learning" OR "teaching" OR "instruction" OR "pedagogy" OR "education" OR "student understanding" OR "misconception") AND ("study" OR "experiment" OR "intervention" OR "assessment") -"review" -"survey" -"meta-analysis" -"systematic", with the year range filter set to since 2010 and result types manually verified to exclude reviews.

Inclusion And Exclusion Criteria

To ensure the relevance and consistency of selected studies, we defined clear inclusion and exclusion criteria before the screening process. Studies were eligible for this review if they were empirical studies published between 2010 and 2025, used quantitative data, were published in a peer-reviewed journal, were about secondary or post-secondary learning or teaching of stoichiometry, and were written in English. Review papers and meta-analyses, non-empirical studies without quantitative measurements, empirical studies containing no description or reporting of effect sizes (e.g., means, standard deviations, sample sizes, and/or correlation coefficients), qualitative measure-only studies, and/or studies that were not published in a peer-reviewed journal or were not peer-reviewed were exclusion criteria for the review. Inclusion criteria aligned with the review objectives (i.e., evaluating the effectiveness of an instructional intervention on stoichiometry learning) and contributed to objectivity and limited selection bias.

Study Selection Process

The study selection process followed a systematic, multi-stage approach to ensure rigor and reproducibility. Initially, the team imported all records retrieved from the database searches into reference management software, which identified and removed duplicates. The process of identifying suitable literature included two steps: initially searching literature based upon the inclusion/exclusion pre-screening criteria defined by the research team, and then assessing the applicability of retrieved literature against the established inclusion/exclusion criteria by using the same pre-defined inclusion/exclusion criteria as the researchers had defined for study eligibility at the point of primary investigation of each of the studies. Initially, studies identified through these two independent literature reviews underwent two independent reviews (one on title/abstract content and, when necessary, an independent full-text review) of either positive or negative results to determine whether to include the study in the primary review. The results

from these two independent literature reviews were combined to form a composite analysis of the stage of review and assessment accomplished before establishing definitive results of eligible literature; thus, an overall summary of eligible literature will provide for systematic identification of all available evidence on a specific topic for further studies associated with this research area.

The PRISMA Flow Chart (Figure 1) shows the selection process, including the number of records at each stage. In total, database searches found 1015 records. Following exclusions due to duplicates and for other reasons (such as whether they were withdrawn or did not have complete metadata) (65 duplicate records; 4 other excluded records), 946 records were screened based on their title and abstract. After applying the inclusion criteria to the 946 records, 807 were excluded, leaving 139 relevant reports for potential retrieval. Of the remaining 139 reports, 46 were unretrievable because they were unavailable or inaccessible; thus, only 93 reports were reviewed for full-text eligibility. Of these 93 reports, 83 were excluded for ineligible criteria (i.e., did not provide enough quantitative data, focused away from stoichiometric calculations, or did not have an empirical design). Consequently, we included 10 studies in the final systematic review and meta-analysis.

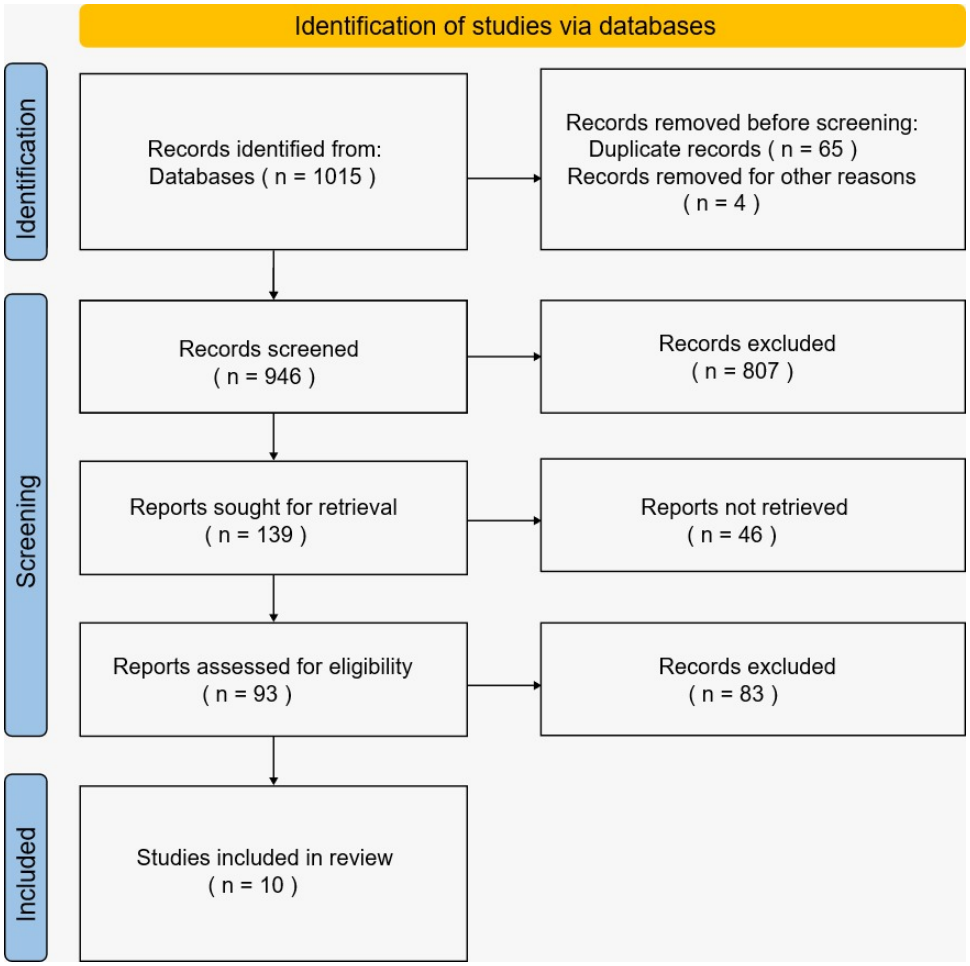


Figure 1. The PRISMA flow chart shows the procedures used to decide which studies to include in our systematic review and meta-analysis.

The literature selection process has several limitations and sources of potential bias. First, limiting the studies to English-language publications may have introduced language bias by excluding related studies published in other languages. Second, omitting grey literature, such as dissertations and conference proceedings, may have introduced publication bias because peer-reviewed journals typically publish studies with statistically significant results. Third, relying on database-specific search strings may have missed studies that used different terms or were indexed in lesser-known databases. Fourth, while necessary for quality purposes, manually verifying Google Scholar results may have involved subjective judgment during screening. Finally, the total number of studies included in the meta-analysis (n = 10) limits statistical power and the generalizability of the findings (Hedges & Vevea, 1998; Quintana, 2023). Despite these limitations, this review's systematic approach provides a transparent, replicable framework for synthesizing current evidence on learning stoichiometry.

RESULTS

Overview Of Included Studies

The ten studies included in this systematic review and meta-analysis span diverse research contexts, participant populations, and instructional interventions focused on learning stoichiometry. The review organized the quantitative data into two major groups by statistical type and the interpretation of each outcome reported in the studies evaluated. The first outcome category is correlation; it identifies the direction and strength of the association between an instructional approach (or student characteristic) and learning outcome and is usually reported as a Pearson correlation (r). Using Pearson correlations as effect sizes for correlation outcomes helps evaluate strength and direction when a causal relationship cannot be established (e.g., students' problem-solving skills may correlate with stoichiometry-related learning outcomes) (Kusuma, Utami, & Mulyani, 2021). The second category of outcome is cognition and affective learning outcomes, which represent the amount of change/difference that results from an intervention, typically assessed with either a pre-test/post-test design or through a comparison between control and experimental groups (Jang & Kim, 2020; Martínez-Borreguero et. al., 2022; Post et al., 2019; Wei et al., 2021). To quantify effect sizes in this context, we calculated the standardized mean difference (SMD) using Cohen's d correction (Cohen, 1960). This applies to studies in which measurement methods differed across studies. The SMD can also compare

effect sizes across studies that used different measurement techniques or scales because it expresses effect size in standard deviation units.

Table 1 presents the coded outcomes for all included studies, along with the intervention examined, the sample's educational level, the sample size, and the effect size assigned to each outcome. The table reveals considerable variability in both the interventions and the reported effects. Several studies evaluated technology-enhanced approaches, including computer simulations and animation-integrated instruction (Alemneh et al., 2024), computer simulations for conceptualizing stoichiometric chemistry (Philip et al., 2023), and generative artificial intelligence integrated with digital curation (Gani, 2025). Others examined cooperative and collaborative arrangements, such as Numbered Heads Together supported by Google Sites (Nabila & Aznam, 2026), peer tutoring through a problem-solving approach (Hayat et al., 2018) , blended collaborative learning in a small private online course (Jong, 2016) , and structured collaborative learning grounded in Vygotskian theory (Aliyu & Talib, 2026) . A further group of studies focused on problem-solving and inquiry-oriented strategies, including the Selvaratnam and Fraser and Ashmore et al. strategies (Mandina & Ochonogor, 2018), Process Oriented Guided Inquiry Learning (Ahmad & Sari, 2026) , the EIGER strategy (Anggraeni et al., 2022), creative problem solving (Wijayanti et al., 2024), problem-based learning integrated with (Anwar et al., 2024), problem-based learning for numerical literacy (Yerimadesi et al., 2026), and the algorithm approach (Niaz, 1995). The remaining studies addressed representational and cognitive supports, such as multiple representation-based learning through cognitive dissonance (Widarti et al., 2021), concept mapping strategies for the mole concept (Bamidele et al., 2013), and animated presentations to support problem-solving maps (Su, 2017). Two studies examined artificial intelligence in stoichiometry learning with respect to problem-solving skills (Zeolita et al., 2026) and metacognitive dynamics under standard versus devil's advocate AI (Afriani, 2026), while one study investigated the influence of learning motivation and self-efficacy on achievement in the mole concept and stoichiometry (Marhadi et al., 2025). The remaining reports examined the TPACK model in stoichiometry teaching (López Cevallos et al., 2025) and comparative difficulties in stoichiometric problem solving across institutional types (Chonillo-Sislema, 2025).

Table 1. Coded Outcomes from Included Studies on Learning Stoichiometry.

Study	**Outcome**	X_t	N_t	X_c	N_c
(Nabila & Aznam, 2026)	stoichiometry learning outcomes	80.27 (4.25)	64	76.31 (5.11)	62

(Afriani, 2026)	stoichiometry learning outcomes	86.33 (7.28)	30	68.67 (8.14)	30
(Alemneh et al., 2024)	stoichiometry learning outcomes	35.67 (6.02)	27	32.28 (7.23)	25
(Mandina & Ochonogor, 2018)	stoichiometry learning outcomes	56.72 (1.16)	117	40.62 (1.16)	275
(Hayat et al., 2018)	stoichiometry learning outcomes	76.26 (–)	34	56.85 (–)	34
(Chonillo-Sislema, 2025)	stoichiometry learning outcomes	8.06 (0.94)	47	7.25 (0.89)	43
(López Cevallos et al., 2025)	stoichiometry learning outcomes	5.88 (2.08)	28	3.70 (1.48)	24
(Ahmad & Sari, 2026)	stoichiometry learning outcomes	82.24 (3.94)	37	82.55 (3.34)	39
(Oyovwi & Otarigho , 2025)	stoichiometry learning outcomes	71.15 (10.56)	92	75.00 (10.66)	53
(Aliyu & Talib, 2026)	stoichiometry learning outcomes	14.23 (5.29)	120	9.95 (3.45)	120
(Zeolita et al., 2026)	stoichiometry learning outcomes	60.15 (–)	34	54.85 (–)	34
(Su, 2017)	stoichiometry learning outcomes	43.33 (10.90)	24	36.13 (11.16)	23
(Wijayanti et al., 2024)	Critical Thinking Skills	79.37 (3.92)	27	74.18 (4.29)	27
	stoichiometry learning outcomes	79.25 (7.51)	27	71.74 (7.51)	27
(Anwar et al., 2024)	stoichiometry learning outcomes	– (–)	49	– (–)	50
(Yerimadesi et al., 2026)	stoichiometry learning outcomes	– (–)	94	– (–)	94

(Jong, 2016)	stoichiometry learning outcomes	59.42 (27.15)	69	44.01 (27.21)	71
(Philip et al., 2023)	stoichiometry learning outcomes	52.94 (20.67)	32	41.20 (28.92)	30
(Gani, 2025)	Critical Thinking Skills	72.80 (8.20)	32	58.20 (9.10)	32
(Widarti et al., 2021)	stoichiometry learning outcomes	79.86 (11.07)	35	67.35 (13.72)	31
(Rudibyani, 2019)	Critical Thinking Skills	78.86 (12.19)	34	57.63 (13.15)	34
(Farwa et al., 2026)	Critical Thinking Skills	64.71 (–)	17	52.94 (–)	18
(Marhadi et al., 2025)	Stoichiometry Learning Achievement	0.67	30		
(Bamidele et al., 2013)	Stoichiometry Learning Achievement	0.11	131		

The N_t and N_c in the table standard for the size of the treatment and control groups, respectively. The X_t and X_c denote M (SD) for SMD. The X_t column denotes the correlation coefficient r for Stoichiometry Learning Achievement.

Heterogeneity Assessment

We assessed heterogeneity using the Q statistic, I^2, and τ^2 (Higgins & Thompson, 2002). As shown in Table 2, the three outcome domains differed markedly. Stoichiometry learning outcomes exhibited extreme heterogeneity ($Q = 780.76, df = 17, p < 1e^{-5}, I^2 = 97.82\%, \tau^2 = 2.27$), and stoichiometry learning achievement was also highly heterogeneous ($Q = 11.09, df = 1, p < 1e^{-3}$, $I^2 = 90.98\%$, $\tau^2 = 0.23$). Critical thinking skills, by contrast, were homogeneous ($Q = 1.49$, $df = 3$, $p = 0.68$, $I^2 = 0.00\%$, $\tau^2 = 0.00$). These results justify a random-effects model (Higgins & Thompson, 2002).

Table 2. Heterogeneity statistics by outcome domain.

Outcome	Q	df	p	I^2 **(%)**	τ^2
Stoichiometry learning outcomes	780.76	17	$p < 1e^{-5}$	97.82	2.27

Critical thinking skills	1.49	3	0.68	0.00	0.00
Stoichiometry learning achievement	11.09	1	$p < 1e^{-3}$	90.98	0.23

Meta-Analysis

We computed pooled estimates within a random-effects framework and organized the results by outcome domain to reflect the conceptual distinctions established in the coding scheme. We report the overall effect for each domain, followed by subgroup and moderator analyses that probe the sources of the substantial heterogeneity documented above. Forest plots accompany each domain-specific synthesis, and we conducted sensitivity analyses to verify that no single study disproportionately influenced the pooled estimates.

Stoichiometry Learning Outcomes

The pooled effect size for stoichiometry learning outcomes was d=0.83 (95% CI: 0.72, 0.93, z=15.81, p<1e^(-55)), which indicates a large and statistically significant benefit of the interventions relative to comparison conditions. As shown in Figure 2, the majority of the 18 effect sizes fell on the positive side of the null, and several were substantial, including the problem-solving strategies examined by Mandina & Ochonogor, (2018) and the generative AI and digital curation intervention reported by Gani (2025). Nevertheless, the distribution was not uniformly favorable. Two studies yielded negative point estimates, namely the POGIL intervention of Ahmad & Sari (2026) and the 7E learning strategy of Oyovwi & Otarigho (2025), and the confidence intervals of several other studies, such as the computer simulation and animation instruction of Alemneh et al. (2024), the AI-supported learning of Zeolita et al. (2026) , and the simulation-based instruction of Philip et al. (2023) , included zero. The precision of these estimates varied considerably, with weights ranging from 3.86 to 53.82, so a few large, precise studies dominated the pooled value, most notably the collaborative learning study of Aliyu & Talib (2026) and the peer tutoring study of Hayat et al. (2018).

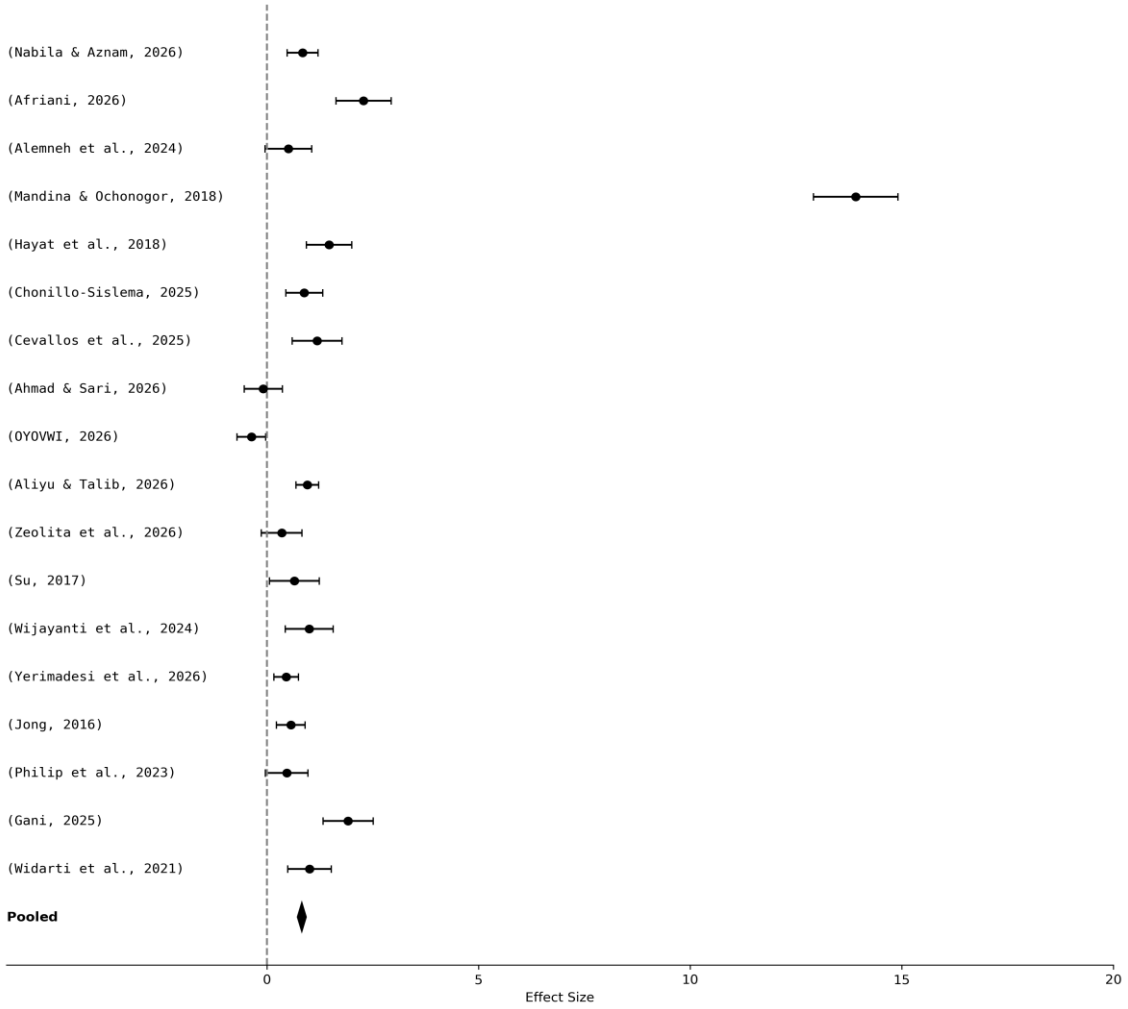


Figure 2. Forest plot for Stoichiometry learning outcomes

Critical thinking skills

The pooled effect size for critical thinking skills was d=1.52 (95% CI: 1.22, 1.82, z=9.94, p<1e^(-22)), which is the largest effect observed in any outcome domain and indicates that the interventions substantially strengthened students' higher-order reasoning. As shown in Figure 3, all four effect sizes were positive and statistically significant, ranging from d=1.26 in the creative problem-solving study of Wijayanti et al. (2024) to d=1.69 in the generative AI and digital curation study of Gani (2025). The problem-based learning intervention of Rudibyani (2019) and the algorithm approach of T (Pangilinan, 2026) occupied intermediate positions, with d=1.67 and d=1.39, respectively. Because the four estimates were homogeneous, as reported in Section 3.2, the pooled value can be interpreted as a common effect that is not attributable to systematic differences among the interventions or their measurement instruments. Weights were relatively balanced, ranging from 7.05 to 12.59, so no single study dominated the synthesis. The two largest effects were associated with interventions that combined problem-solving scaffolds with technological or representational supports, whereas the

smallest effect came from a purely collaborative creative problem-solving model, although the domain's homogeneity cautions against overinterpreting this pattern.

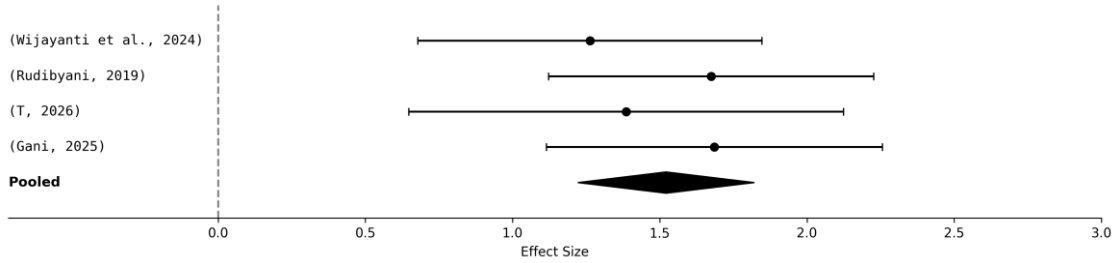


Figure 3. Forest plot for Critical thinking skills

Stoichiometry learning achievement

The pooled effect size for stoichiometry learning achievement was d=0.23 (95% CI: 0.07, 0.39, z=2.87, p=0.004), which is statistically significant but considerably smaller than the effects observed for the other two domains. As shown in Figure 4, only two studies contributed to this synthesis, and their estimates diverged sharply. The study by Marhadi et al. (2025) reported a moderate positive association between learning motivation, self-efficacy, and chemistry achievement in the mole concept and stoichiometry, with d=0.81 (95% CI: 0.44, 1.19), and it carried a weight of 27.00. By contrast, the concept mapping study of Bamidele et al. (2013) yielded a negligible effect, d=0.11 (95% CI: -0.07, 0.28), yet it received a weight of 128.00 and therefore dominated the pooled estimate. This imbalance in weights, combined with the high heterogeneity reported in Section 3.2 (I^2=90.98%), indicates that the pooled value should be interpreted cautiously, since it largely reflects the near-null finding of a single large study rather than a consistent pattern across the domain. The contrast between the two estimates also suggests that the association's strength depends heavily on the predictor variables examined, with affective and self-regulatory constructs predicting achievement more strongly than comparisons of instructional strategies.

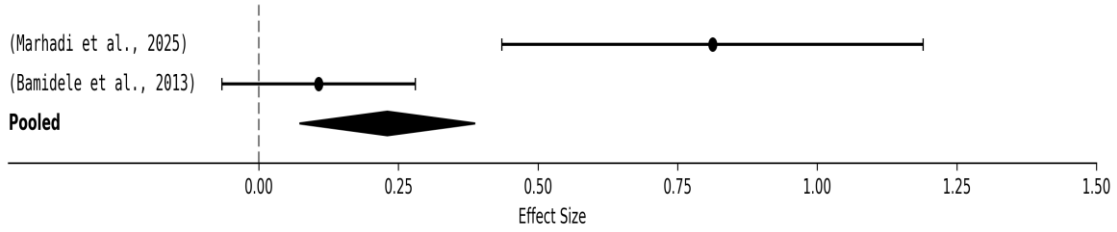


Figure 4. Forest plot for Stoichiometry learning achievement

Publication Bias Assessment

Publication bias for the primary outcome domain, stoichiometry learning outcomes, was assessed by means of a funnel plot and Egger's regression test for funnel plot asymmetry (Egger et al., 1997; Higgins & Thompson, 2002). As shown in Figure 5, the funnel plot of the 24 effect sizes was visibly asymmetric, with 18 studies located to the left of the center and only 6 to the right, and the mean effect size was considerably larger among studies on the right (d=3.88) than among those on the left (d=0.71). This pattern aligns with the possibility small studies reporting large positive effects were more likely to be published, although the mean absolute deviation from the center was 1.19 and the standard errors ranged widely from 0.14 to 1.00, which indicates the asymmetry is partly driven by a few imprecise outliers. Nevertheless, Egger's test did not reach statistical significance, with an intercept of 2.33 and a corresponding p value of 0.28, so we cannot reject the null hypothesis of funnel plot symmetry at conventional thresholds. Given the modest number of studies and the pronounced heterogeneity documented in Section 3.2, this test has limited power, and the graphical evidence of asymmetry should therefore be treated as suggestive rather than conclusive.

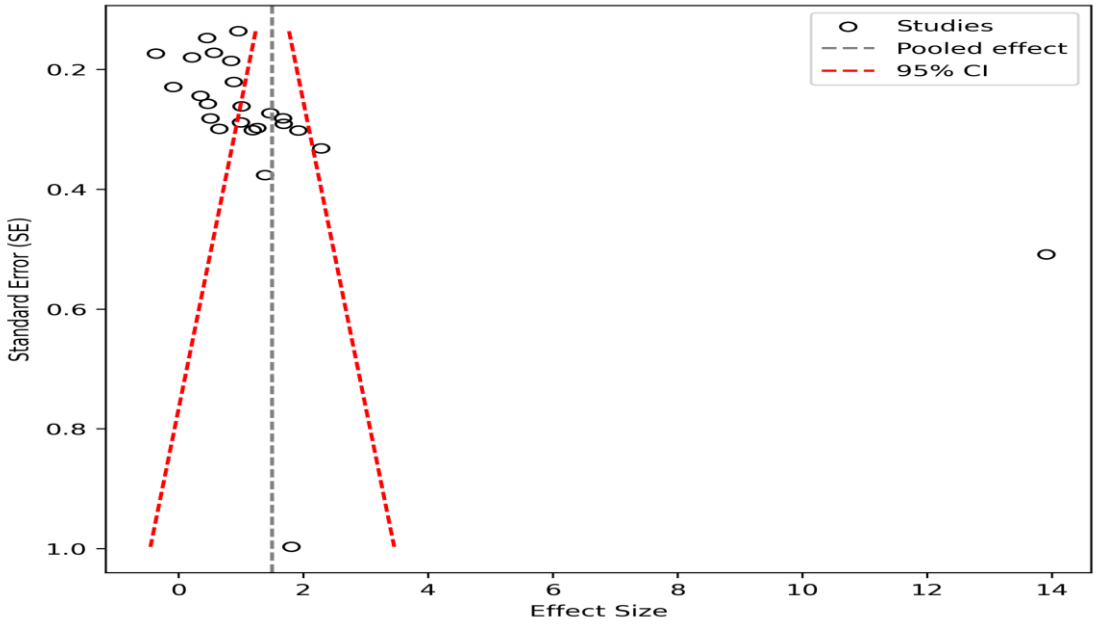


Figure 5. Funnel plot for publication bias assessment.

DISCUSSION

Taken together, the results of this synthesis paint a coherent yet nuanced picture of what instructional interventions can accomplish in stoichiometry education. The most robust finding across the literature is that interventions with widely differing theoretical provenance, spanning computer simulations and intelligent tutoring as well as cooperative learning and problem-based

instruction, produce a large, statistically significant benefit on broad measures of stoichiometry learning, with a pooled effect of d=0.83. This convergence is striking precisely because the interventions differ so substantially in their underlying assumptions about how learning proceeds. Conceptual change approaches posit that misconceptions require direct confrontation (Demircioglu et al., 2005), representational tools posit that difficulties stem from an incapacity to translate among the macroscopic, particulate, and symbolic levels (Harrison & Treagust, 2002), and scaffolding approaches posit that the bottleneck resides in working memory and the management of multi-step procedures (Crippen & Earl, 2004). The comparable advantages observed across all three intervention types imply multiple, partly independent obstacles to stoichiometry learning, with the resolution of any single obstacle being adequate for measurable improvement. The practical consequence is heartening: no singular privileged approach exists for improved stoichiometry teaching, and instructors may choose methods suited to their local limitations without compromising efficacy.

Another pattern that consistently emerges across the included studies concerns the domain-specificity of intervention effects. The pooled effect on critical thinking skills, d=1.52, was nearly twice as large as the effect on general stoichiometry learning outcomes, and more than six times as large as the effect on stoichiometry learning achievement, d=0.23. This gradient is unlikely to be an artifact of the analysis, since the critical thinking domain was the sole one proved homogeneous, with I2 = 0.00%, indicating the four contributing studies converged on a common effect despite examining creative problem solving, problem-based learning, the algorithm approach, and generative AI-supported instruction (Gani, 2025; Pangilinan, 2026; Rudibyani, 2019; Wijayanti et al., 2024. The homogeneity is theoretically informative. It implies that higher-order reasoning in stoichiometry depends not on any specific instructional technology or pedagogical model, but on a shared underlying element across these interventions: the demand that students generate, justify, and evaluate solution strategies rather than execute prescribed algorithms. In contrast, interventions that rehearse only procedural steps seem to yield more limited improvements on standard achievement assessments. This interpretation accords with the enduring observation that students can manipulate chemical equations algorithmically while remaining unable to explicate what the symbols denote (Furió et.al., 2002), and it suggests the distinction between procedural and conceptual mastery, which has organized much of the chemistry education literature, ought to be broadened to encompass higher-order reasoning as a third and partially separable outcome.

The disparity between the critical thinking and achievement domains also highlights a measurement issue that extends across the field. The achievement domain rested on only two studies: one reported a moderate association between motivational and self-regulatory constructs and achievement (Marhadi et al., 2025), and the other reported a negligible association for concept mapping (Bamidele et al., 2013); the pooled estimate was dominated by the latter because of its much larger weight. With I2 = 90.98%, these two estimates cannot be regarded as samples from a common population, and the pooled value of d=0.23 therefore describes an average of incommensurable quantities rather than a stable effect. A wider implication is the absence of a common measurement framework within the discipline. Investigations of stoichiometry learning draw on researcher-developed achievement tests, standardized concept inventories, and ad hoc critical thinking rubrics with minimal overlap (Arasasingham et al., 2004), and the consequent heterogeneity within the stoichiometry learning outcomes domain, where I2 = 97.82% and $\tau 2$ =2.27, is so severe that the pooled estimate, regardless of the precision of its confidence interval, ought to be interpreted as a summary of a truly dispersed distribution rather than as a parameter any single intervention can be expected to replicate. The narrow confidence interval around d= 0.83 reflects the large number of effect sizes and the precision of a few large studies, not uniformity across contexts.

The moderator patterns visible in the forest plots reinforce this caution while also pointing toward substantive explanations. The largest effects in the stoichiometry learning outcomes domain arose from problem-solving strategy instruction [46] and from generative AI integrated with digital curation (Gani, 2025), whereas the two negative point estimates came from a POGIL implementation (Ahmad & Sari, 2026) and a 7E learning strategy (Oyovwi, & Otarigho, 2025). Notably, the negative estimates were tied to interventions regarded favorably in the wider literature (Ahmad & Sari, 2026), suggesting that implementation fidelity, teacher experience, and alignment between intervention design and local assessment practices may matter more than the approach's nominal identity. Individual estimates differed vastly in their precision, as their weights ranged from 3.86 to 53.82, so a few large trials, especially the collaborative learning study of Aliyu & Talib (2026) and the peer tutoring study of Hayat et al. (2018) disproportionately affected the pooled value (Hayat et al., 2018). This disparity constitutes a structural aspect of educational intervention research, in which sample sizes are limited by class sizes and institutional access, and it signifies the evidence base is narrower than the quantity of included studies suggests.

The publication bias assessment adds a further qualification. The funnel plot was visibly asymmetric, with the mean effect among studies on the right of the

center, d=3.88, far exceeding the mean among studies on the left, d=0.71, yet Egger's test did not reach significance, with an intercept of 2.33 and p=0.28 (Higgins & Thompson, 2002). The divergence between the graphical and statistical findings is best explained by the limited power of Egger's test when the number of studies is small and heterogeneity is severe, a circumstance in which the test is known to perform poorly (Torgerson, 2006). Therefore, one cannot rule out the possibility that selective reporting of positive findings exaggerates the pooled effect, and the true average effect may be lower than d=0.83. Screening out dissertations, book chapters, and conference abstracts worsens this concern because the sampling frame systematically excluded unpublished studies with null findings. An equitable interpretation of the data is that stoichiometry interventions are probably effective, but the extent of their effectiveness is uncertain and could be well below the pooled point estimate (Gongden & Delmang, 2026).

The theoretical implications of these findings concern how we model the relationship between instruction and learning in quantitative domains. The dominant frameworks in chemistry education have treated stoichiometry learning as a unitary construct indexed by achievement, and interventions have accordingly been evaluated by their effect on test scores (Paredes-Turra et al., 2025). Our findings imply this framing conceals the most consequential effects. If interventions reliably strengthen critical thinking while yielding only moderate improvements on standard attainment metrics, then assessments that rely solely on attainment exams will systematically undervalue their worth and may even privilege rote-focused methods that align with the items on those exams. A more adequate model would treat procedural fluency, conceptual understanding, and higher-order reasoning as distinct but interacting outcomes, each with its own developmental trajectory and its own sensitivity to instructional design. The cognitive load and representational frameworks which have guided much intervention design (Harrison y Treagust, 2002; Ye et al., 2025) address chiefly the first two outcomes, and they may require supplementation by accounts of how students acquire the ability to coordinate multiple representations under conditions of uncertainty, which comes closer to what critical thinking in stoichiometry truly demands.

For practitioners, several specific suggestions follow. Educators who design or choose stoichiometry instruction ought to prioritize exercises in which learners build and support solution approaches, since the evidence shows these exercises produce the greatest gains in advanced reasoning, and those gains remain steady across teaching media. Because the critical thinking domain is uniform, the technology matters less than the cognitive demand it imposes, so a well-designed peer critique of competing solution routes may

be as effective as an adaptive tutoring system and much easier to implement in resource-constrained settings. Curriculum developers should also resist the temptation to treat stoichiometry as a purely procedural topic to be dispatched early in the course. That stoichiometric difficulties continue into later coursework (Salame et al., 2026) suggests that instructional time devoted to reasoning about quantities is not time taken from content coverage but rather an investment that pays dividends in subsequent topics. When evaluating student performance, departments should supplement traditional achievement-examination measures with measures that capture the quality of reasoning; otherwise, interventions with the greatest potential to help students may seem ineffective in local assessments. Policymakers concerned with persistence in chemistry (Kim, 2025) should recognize stoichiometry's gatekeeping function as avoidable, as an outcome of instructional and assessment decisions open to revision.

Multiple constraints inherent in this review temper these inferences. Limiting the scope to English-language publications excludes pertinent work in regional journals, especially those in Spanish, Portuguese, and Indonesian, although the included studies point to an active research community in those settings (Chonillo-Sislema, 2025; López Cevallos et al., 2025). Omitting grey literature and relying on automated duplicate detection, which found no duplicates across six databases, increases the chance that some overlap went unidentified and that the actual number of independent studies is lower than stated. The decision to exclude studies with a high risk of bias in more than two domains, while defensible, removed 30 of 54 full-text reports and may have disproportionately excluded classroom-based studies conducted under authentic conditions, where randomization and blinding are impractical. The risk-of-bias instrument was borrowed from medical research, and applying it to educational interventions requires judgments that are not entirely reproducible, especially regarding blinding, which is seldom practicable when the instructor is also the intervention designer. Ultimately, the limited number of investigations informing critical thinking and achievement constrains the certainty with which area-specific inferences can be made, and moderator analyses that would have clarified the roles of length, educational tier, and delivery method could not be conducted with sufficient statistical power.

These limitations point directly to priorities for future research. Investigations are needed that report outcome-specific effects rather than a single aggregate achievement score, so researchers can estimate the differential impact on procedural, conceptual, and higher-order outcomes with precision. Future research should explore the conditions under which interventions fail, since the negative estimates observed for POGIL and the 7E strategy (Ahmad & Sari, 2026; Oyovwi & Otarigho, 2025) are more

informative about implementation than the many positive results, yet such findings are rarely pursued in follow-up work. Understudied areas include long-term retention, since almost all included studies measured outcomes immediately after instruction, and transfer to novel problem types, which is the outcome most closely tied to the stated goals of chemistry education. Methodologically, cluster-randomized designs that respect the nested structure of classrooms, preregistered analysis plans that reduce the risk of selective outcome reporting, and a shared stoichiometry assessment instrument that allows comparison of effect sizes across studies without the interpretive burden of heterogeneous measures are needed (Arasasingham et al., 2004). Investigations into affective and self-regulatory antecedents of stoichiometry performance remain scarce, and the sole study exploring motivation and self-efficacy (Marhadi et al., 2025) documented a moderate relationship warranting replication with larger samples. Finally, the swift appearance of generative AI as an instructional instrument in stoichiometry (Afriani, 2026; Gani, 2025; Zeolita et al., 2026) presents both a possibility and a hazard: such instruments may foster the sort of strategy production our findings tie to critical thinking gains, yet they may equally let students avoid reasoning entirely, and only meticulously designed comparative investigations can establish which result wins out.

CONCLUSION

This systematic review and meta-analysis aimed to quantify the effectiveness of instructional interventions for stoichiometry learning and determine whether their benefits differ across outcome domains. The evidence indicates these interventions are broadly effective, with a pooled effect of d=0.83 on stoichiometry learning outcomes, yet the synthesis also shows effects are strongly domain-dependent: critical thinking skills improved substantially (d=1.52) and homogeneously, whereas conventional achievement gains were modest (d=0.23) and rested on only two heterogeneous studies. We therefore infer that the field's predominant focus on achievement scores limits what well-designed instruction can achieve, and that procedural fluency, conceptual understanding, and higher-order reasoning should be treated as separate outcomes in both theory and evaluation.

The practical implications are direct. Instructors and curriculum developers ought to prioritize assignments that require students to build, defend, and evaluate solution approaches, because such assignments seem to yield the greatest and most reliable improvements. Assessment methods should expand accordingly, since appraisals restricted to achievement examinations might disadvantage precisely those interventions that most effectively benefit students. In principle, our results challenge single accounts of stoichiometry

acquisition and require conceptual structures that delineate how reasoning abilities mature alongside subject-matter proficiency.

Subsequent investigations should document outcome-specific effects, probe when treatments prove ineffective, and scrutinize retention and transfer over extended periods. A shared assessment instrument and preregistered, cluster-randomized designs would substantially strengthen the evidence base, as would rigorous comparative work on generative AI, whose instructional promise remains largely untested.

Conflict Of Interest.

The authors declare no conflicts of interest.

Acknowledgments

We thank the National University of Education (UNAE) for supporting this research through its research funding program, which made it possible to finance the academic and scientific activities of the Eureka 4i research group.

REFERENCES

Anderson, T.R., Boersma, M. & Raubenheimer, D. (2004). Stoichiometry: linking elements to biochemicals. Ecology. 85: 1193-1202.https://doi.org/10.1890/02-0252

Anggraeni, N. F., Rahayu, S., Fajaroh, F., & Effendy, E. (2022). The effectiveness of the EIGER learning strategy in promoting students' conceptual and algorithmic understanding of stoichiometry. International Journal of Instruction, 15(4), 35-56. https://doi.org/10.29333/iji.2022.1543a e-iji.net

Ahmad, F., & Sari, N. I. (2026). The effect of process-oriented guided inquiry learning strategy on students' understanding of stoichiometry concepts and collaboration skills. Venn: Journal of Sustainable Innovation on Education, Mathematics and Natural Sciences, 5(2), 341-350. https://doi.org/10.53696/venn.v5i2.439

Alemneh, E. D., Kidanmariam, D. A., Mengistie, S. M., & Beyene, B. B. (2024). Effect of computer simulation and animation-integrated instruction on pre-service science teacher trainees conceptual understanding and retention of acid-base chemistry and stoichiometry. Journal of Technology and Science Education, 14(2), 453-472. https://doi.org/10.3926/jotse.2435

Aliyu, H., & Talib, C. A. (2026). Socio-cognitive dynamics in chemistry education: A Vygotskian analysis of collaborative learning practices. Abdurrauf Social Science, 3(1), 146-165. https://doi.org/10.70742/arsos.v3i1.593

Anwar, Y. A. S., Al Idrus, S. W., & Junaidi, E. (2024). Implementation of integrated problem-based learning and storytelling technique on learning outcomes and attitudes towards chemistry. Jurnal Pendidikan Progresif, 14(2), 1059-1070. https://doi.org/10.23960/jpp.v14.i2.202477

Arasasingham, R. D., Taagepera, M., Potter, F., & Lonjers, S. (2004). Using knowledge space theory to assess student understanding of stoichiometry. Journal of Chemical Education, 81(10), 1517–1523. https://doi.org/10.1021/ed081p1517

Afriani, R. A. (2026). Students' metacognitive analysis in stoichiometry learning based on generative AI: A comparative study of regular AI vs. Devil's Advocate AI. Jurnal Penelitian Progresif, 6(1), 1-11. https://doi.org/10.61992/jpp.v6i1.401

Bamidele, E. F., Adetunji, A. A., Awodele, B. A., & Irinoye, J. (2013). Attitudes of Nigerian secondary school chemistry students towards concept mapping strategies in learning the mole concept. Academic Journal of Interdisciplinary Studies, 2(2), 475–484. https://doi.org/10.5901/ajis.2013.v2n2p475

Behrman, A.L., Ardolino, E., VanHiel, L.R., Kern, M., Atkinson, D., Lorenz, D.J. & Harkema, S.J. (2012). Assessment of functional improvement without compensation reduces variability of outcome measures after human spinal cord injury. Archives of Physical Medicine and Rehabilitation, 93, 1518-1529. https://doi.org/10.1016/j.apmr.2011.04.027

Beskeni, R.D., Yousuf, M.I., Awang, M.M. & Ranjha, A.N. (2011). The effect of prior knowledge in understanding chemistry concepts by senior secondary school students. International Journal of Academic Research, 3(2), 607-61. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=gNql7SsAAAAJ&citation_for_view=gNql7SsAAAAJ:UeHWp8X0CEIC

Bigelow, G.S., Padula, S.A., Garg, A., Gaydosh, D. & Noebe, N.D.(2010). Characterization of ternary NiTiPd high-temperature shape-memory alloys under load-biased thermal cycling. Metallurgical and Materials Transactions A, 41, 3065-3079. https://doi.org/10.1007/s11661-010-0365-5

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46. https://doi.org/10.1177/001316446002000104

Crippen, K. J., & Earl, B. L. (2004). Considering the efficacy of web-based worked examples in introductory chemistry. Journal of Computers in Mathematics and Science Teaching, 23(2), 151–167.

Chonillo-Sislema, L. O. (2025). Estudio comparativo sobre las dificultades en la resolución de problemas estequiométricos en las clases de química entre una institución particular y pública. En Actas del III Congreso Internacional de Innovación, Ciencia y Tecnología INUDI-UH, 2025 (Cap. 21). Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú. https://doi.org/10.35622/inudi.c.03.21

Demircioglu, G., Ayas, A., & Demircioglu, H. (2005). Conceptual change achieved through a new teaching program on acids and bases. Chemistry Education Research and Practice, 6(1), 36–51. https://doi.org/10.1039/B4RP90003K

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

Flaherty, A.A. (2020). A review of affective chemistry education research and its implications for future research. Chemistry Education Research and Practice. 21 (3): 698–713. https://doi.org/10.1039/C9RP00200F

Farwa, U., Riaz, M., & Ullah, A. (2026). Developing critical thinking skills through active learning: Strategies for higher secondary and undergraduate education. Social Science Review Archives, 4(1), 3485–3494. https://doi.org/10.70670/sra.v4i1.1910

Furió, C., Azcona, R., & Guisasola, J. (2002). The learning and teaching of the concepts 'amount of substance' and 'mole': A review of the literature. Chemistry Education Research and Practice, 3(3), 277–292. https://doi.org/10.1039/B2RP90023H

García, L.L.W.(2025). A Narrative Inquiry into Chemistry Teachers' Experiences with Stoichiometry. International Journal of Advanced Multidisciplinary Research. eISSN:

2799-0664. 5(7), 1059-1067.https://www.ijams-bbp.net/archive/vol-5-issue-7/a-narrative-inquiry-into-chemistry-teachers-experiences-with-stoichiometry/

Gani, A. (2025). Integrasi generative AI dan digital curation dalam pembelajaran stoikiometri untuk meningkatkan berpikir kritis dan gotong royong siswa SMA. AS-SABIQUN, 7(5), 1037-1055. https://doi.org/10.36088/assabiqun.v7i5.5894

Gongden, E.J. & Delmang, T.K. (2016). The efficacy of concept mapping instructional strategy in remedying students' problem-solving difficulties in stoichiometry. African Research Review. 10(2), 144-157. https://doi.org/10.4314/afrrev.v10i2.9

Gulacar, O., Overton, T.L., Bowman, C.R. & Fynewever, H. (2013). A novel code system for revealing sources of students' difficulties with stoichiometry. Chemistry Education Research and Practice. 35(2), 275–284. https://doi.org/10.1039/c3rp00029j

Gulacar, O., Mann, H. K., Mann, S. S., & Vernoy, B. J. (2022). The influence of problem construction on undergraduates' success with stoichiometry problems. Education Sciences, 12(12), 867. https://doi.org/10.3390/educsci12120867

Gulacar, O., Tan, A., Cox, C. T., Bloomquist, J., Jimmy, O., & Cao, N. (2019). Analyzing characteristics of experts in the context of stoichiometric problem-solving. Education Sciences, 9(3), 219. https://doi.org/10.3390/educsci9030219

Hamerská, L., Matěcha, T., Tóthová, M., & Rusek, M. (2024). Between symbols and particles: Investigating the complexity of learning chemical equations. Education Sciences, 14(6), 570. https://doi.org/10.3390/educsci14060570

Harrison, A. G., & Treagust, D. F. (2002). The particulate nature of matter: Challenges in understanding the submicroscopic world. In J. K. Gilbert, O. De Jong, R. Justi, D. F. Treagust, & J. H. Van Driel (Eds.), Chemical education: Towards research-based practice (pp. 189–212). Springer. https://doi.org/10.1007/0-306-47977-X_9

Hayat, A. Z., Wahyu, W., & Kurnia. (2018). Comparison of peer-tutoring learning model through problem-solving approach and traditional learning model on the cognitive ability of grade 10 students at SMKN 13 Bandung on the topic of stoichiometry. Journal of Physics: Conference Series, 1013(1), Article 012208. https://doi.org/10.1088/1742-6596/1013/1/012208

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. Psychological Methods, 3(4), 486–504. https://doi.org/10.1037/1082-989X.3.4.486

Higgins, J.P.T. & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11):1539-1558. 21(11), 1539-1558. https://doi.org/10.1002/sim.1186

Horvat, S., Segedinac, M. D., Milenković, D. D., & Hrin, T. N. (2016). Development of procedure for the assessment of cognitive complexity of stoichiometric tasks.Macedonian Journal of Chemistry and Chemical Engineering, 35(2), 275–284. https://doi.org/10.20450/mjcce.2016.893

Innocenti, A. & Kamarulzaman, N. (2012). Stoichiometry and Materials Science: When Numbers Matter.BoD – Books on Demand. 450 p. https://books.google.com.ec/books/about/Stoichiometry_and_Materials_Science.html?id=1tqgDwAAQBAJ&redir_esc=y

Jang, H. Y., & Kim, H. J. (2020). A meta-analysis of the cognitive, affective, and interpersonal outcomes of flipped classrooms in higher education. Education Sciences, 10(4), Article 115. https://doi.org/10.3390/educsci10040115

Jong, J. P. (2016). The effect of a blended collaborative learning environment in a small private online course (SPOC): A comparison with a lecture course. Journal of Baltic Science Education, 15(2), 194–203. https://doi.org/10.33225/jbse/16.15.194

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. Visualization in Science Education. (1), 121-145. https://doi.org/10.1007/1-4020-3613-2_8

Kim, D. (2025). Efforts to Improve General Chemistry Student Success: Describing Stoichiometry Solution Patterns and Investigating Challenge Perceptions. USF Tampa Graduate Theses and Dissertations.
https://digitalcommons.usf.edu/etd/10971

Kusuma, A.K., Utami, B. & Mulyani, B. (2021). The relationship between students' problem-solving skills and scientific attitude with students' learning outcomes on stoichiometry at tenth grade at high school in Boyolali, Central Java, Indonesia. Journal of Physics: Conference Series, 1842(1), 012027. https://doi.org/10.1088/1742-6596/1842/1/012027

López Cevallos, F. E., Tulcán Benavides, F. L., & Lascano Rivera, S. (2025). El modelo TPACK en la enseñanza de estequiometría química. SATHIRI, (Especial), 9–18. https://doi.org/10.32645/13906925.1349

Mandina, S., & Ochonogor, C. (2018). Comparative effect of two problem-solving instructional strategies on students' achievement in stoichiometry. EURASIA Journal of Mathematics, Science and Technology Education, 14(12), Article em1621. https://doi.org/10.29333/ejmste/95125

Marhadi, M., Husain, H., & Anwar, M. (2025). The impact of learning motivation and self-efficacy on students' academic achievement. Formosa Journal of Applied Sciences, 4(4), 1113–1126. https://doi.org/10.55927/fjas.v4i4.105 SRH Formosa Publisher

Martínez-Borreguero, G., Naranjo-Correa, F. L., & Mateos-Núñez, M. (2022). Cognitive and emotional development of STEM skills in primary school teacher training through practical work. Education Sciences, 12(7), Article 470. https://doi.org/10.3390/educsci12070470

Nabila, T., & Aznam, N. (2026). The effect of NHT Google Sites on critical thinking and motivational learning on stoichiometry material. Indonesian Journal of Educational Development (IJED), 7(1), 395-409. https://doi.org/10.59672/ijed.v7i1.6026

Niaz, M. (1995). Progressive transitions from algorithmic to conceptual understanding in student ability to solve chemistry problems: A Lakatosian interpretation. Science Education, 79(1), 19-36. https://doi.org/10.1002/sce.3730790103

Noor Nabila, T., & Aznam, N. (2026). The effect of NHT Google Sites on critical thinking and motivational learning on stoichiometry material. Indonesian Journal of Educational Development (IJED), 7(1), 395-409. https://doi.org/10.59672/ijed.v7i1.6026

Oyovwi, E. O., & Otarigho, M. D. (2025). Effects of problem-solving and 7E learning instructional strategies in enhancing students' attitude towards stoichiometry curriculum concepts in chemistry. International Journal of Chemistry and Chemical Processes, 11(1), 34–46. https://doi.org/10.56201/ijccp.v11.no1.2025.pg34.46

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D.(2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372: n71. https://doi.org/10.1136/bmj.n71

Pangilinan, C. T. (2026). Developing problem solving skills in chemistry through the algorithm approach. International Journal of Recent Trends in Multidisciplinary Research, 6(5), 359–373. https://doi.org/10.59256/ijrtmr.20260605039

Paredes-Turra, C. E., Huincahue, J., & Ariza, Y. (2025). Teaching stoichiometry through the use of teaching models: A literature review. Journal of Physics: Conference Series, 3117(1), Article 012020. https://doi.org/10.1088/1742-6596/3117/1/012020 ResearchGate

Philip, A. J., Du Toit, G., & Van Breda, C. (2023). The effect of computer simulation on Grade 11 learners' conceptualisation of stoichiometric chemistry. European Journal of STEM Education, 8(1), Article 01. https://doi.org/10.20897/ejsteme/12947

Post, L. S., Guo, P., Saab, N., & Admiraal, W. (2019). Effects of remote labs on cognitive, behavioral, and affective learning outcomes in higher education. Computers & Education, 140, Article 103596. https://doi.org/10.1016/j.compedu.2019.103596

Quintana, D. S. (2023). A guide for calculating study-level statistical power for meta-analyses. Advances in Methods and Practices in Psychological Science, 6(1). https://doi.org/10.1177/25152459221147260

Ramful, A. & Narod, F.B. (2014). Proportional reasoning in the learning of chemistry: Levels of complexity. Mathematics Education Research Journal, 26, 25-46.https://doi.org/10.1007/s13394-013-0110-7

Reyes, R.L. (2025). Addressing Common Student Errors and Misconceptions in Chemical Stoichiometry: Insights for Enhanced Instruction in Introductory General Chemistry. Journal of Science, Mathematics & Technology Learning.32(2), 2025, 167–88 https://doi.org/10.18848/2327-7971/CGP/v32i02/167-188

Rosa, A., Corrales, A., Nguyen, Y. & Atkinson, M.B.(2022). Relevance and equity: should stoichiometry be the foundation of introductory chemistry courses?. Chemistry Education Research and Practice, 23 (3): 662–685. https://doi.org/10.1039/d1rp00333j

Rudibyani, R. B. (2019). Improving students' creative thinking ability through problem based learning models on stoichiometric materials. Journal of Physics: Conference Series, 1155, Article 012049. https://doi.org/10.1088/1742-6596/1155/1/012049

Salame, I. I., Kanishka, H., & Ishak, A. (2026). Examining challenges and difficulties students face in learning about stoichiometry in general chemistry courses. Interdisciplinary Journal of Environmental and Science Education, 22(1), Article e2604. https://doi.org/10.29333/ijese/17701

Schmidt, H.J. & Jignéus, C. (2003). Students' strategies in solving algorithmic stoichiometry problems. Chemistry Education Research and Practice, 4 (3): 305–317. https://doi.org/10.1039/B3RP90018E

Siswaningsih, W., Firman, H., Zackiyah, & Khoirunnisa, A. (2017). Development of Two-Tier Diagnostic Test Pictorial-Based for Identifying High School Students' Misconceptions on the Mole Concept. Journal of Physics: Conference Series. 812012117.https://doi.org/10.1088/1742-6596/812/1/012117

Su, K.-D. (2017). Tactic fulfillments of three correlations for problem-solving maps and animated presentations to assess students' stoichiometry performances. Journal of Baltic Science Education, 16(5), 733–745. https://doi.org/10.33225/jbse/17.16.733

Torgerson, C. J. (2006). Publication bias: The Achilles' heel of systematic reviews? British Journal of Educational Studies, 54(1), 89–102. https://doi.org/10.1111/j.1467-8527.2006.00332.x

Wei, X., Saab, N., & Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. Computers & Education, 163, Article 104097. https://doi.org/10.1016/j.compedu.2020.1040

Widarti, H. R., Permanasari, A., Mulyani, S., Rokhim, D. A., & Habiddin. (2021). Multiple representation-based learning through cognitive dissonance strategy to reduce student's misconceptions in volumetric analysis. TEM Journal, 10(3), 1263–1273. https://doi.org/10.18421/TEM103-33

Wijayanti, D., Ibnu, S., Utomo, Y., & Munzil. (2024). The effectiveness of creative problem-solving learning on critical thinking skills and stoichiometry learning outcomes of grade X students at SMKN 1 Bakung Blitar Indonesia. World Journal of Advanced Research and Reviews, 24(1), 327–338. https://doi.org/10.30574/wjarr.2024.24.1.2980

Ye, J., Gao, J., Lin, T., He, K., & Chen, D. (2025). Cognitive load change in chemical concept learning: Insights from event-related potentials. Journal of Baltic Science Education, 24(1), 92–104. https://doi.org/10.33225/jbse/25.24.92

Yerimadesi, Andromeda, Guspatni, Fauziah, Febrila, P. Z., & Makrooni, G. (2026). The influence of gender and problem-based learning model on students' numerical literacy in chemistry. Jurnal Pendidikan IPA Indonesia, 15(1), 141-155. https://doi.org/10.15294/jpii.v15i1.37300

Zeolita, S., Asyhar, R., Asrial, A., Naswir, M., Yusnidar, Y., & Asmiyunda, A. (2026). The impact of artificial intelligence (AI) in stoichiometry learning on problem-solving skills of senior high school students in Phase F. Jurnal Pendidikan MIPA, 16(4), 1481-1493. https://doi.org/10.37630/jpm.v16i4.4845