Rima International Journal of Education (RIJE)
ISSN: 2756 – 6749 (p); 3141-2033 (e)
, Vol. 5, Issue 2, March. 2026; Page: 221—245 DOI: https://doi.org/10.65760/rijessu.v5.2.16

Delving Deeper into Chemistry Education: Understanding How Students Learn and How to Teach Effectively

***[1]Hassan Aliyu, [2]Amina M. Chado, [3]Umar Sarkin Bauchi Idris, [4]L. A. Fadipe and [5]Corrienna Abdul Talib**

*[*1]Department of Science Education, Faculty of Education, Sokoto State University (SSU), Sokoto.* ***Email**:* nagoronyo@gmail.com & aliyu.hassan@ssu.edu.ng **ORCID**: https://orcid.org/0000-0003-4929-3126

[2&3]Department of Science Education, School of Science and Technology Education (SSTE), Federal University of Technology Minna (FUTMinna), Niger State.

[4]Department of chemistry, School of Physical Science, Federal University of Technology Minna (FUTMinna), Niger State.
[5]Department of science, technology, mathematics and creative multimedia, Faculty of social science and humanities, Universiti Teknologi Malaysia (UTM), Johor, Malaysia

Abstract

Chemistry education directly influences workforce preparation for energy, medicine, and materials science. However, persistent difficulties in mastering abstract concepts such as thermodynamics and molecular interactions limit student success, with attrition rates exceeding 30% at many institutions. The cognitive mechanisms underlying these failures remain poorly specified, and instructional strategies that work in one context often fail in resource-constrained settings. Here we show that a 10-week intervention combining scenario-based problem solving with molecular visualization software significantly improves conceptual mastery. In a controlled trial with 324 preservice teachers, the experimental group achieved adjusted post-test scores 2.53 points higher than the control group (ANCOVA, F(1,318)=37.4, p<0.001, η^2=0.105). Gains concentrated on the most difficult concepts: pre-test correct rates below 8% for thermodynamics and intermolecular forces rose substantially in the experimental condition. Despite this improvement, over 57% of students never participated in collaborative problem-solving, and only 19.9% rated collaboration as highly effective compared to 71.5% for simulations. We conclude that representational bottlenecking, not general ability, drives chemistry learning failures, and that cognitive-conflict pedagogy targeting specific conceptual barriers produces measurable gains even in technology-limited environments.

Keywords*:* Chemistry education, Representational fluency, Cognitive conflict, Active learning, Sub-Saharan Africa

Cite this as: Aliyu, H., Chado,A. M., Idris, U. S. B., Fadipe, L. A. & Talib, C. A. (2026). Delving Deeper into Chemistry Education: Understanding How Students Learn and How to Teach Effectively. *Rima International Journal of Education, 5*(2), 221—245. DOI: https://doi.org/10.65760/rijessu.v5.2.16

Introduction

The attrition rate in undergraduate introductory chemistry courses remains persistently high, often exceeding 30-40% at many institutions (Fink *et al.*, 2020). This bottleneck directly constrains the pipeline of skilled graduates into critical fields such as renewable energy, pharmaceutical development, and materials science. The challenge is not merely one of content delivery, but of identifying and implementing pedagogical strategies that reliably transform novice learners into competent, conceptual thinkers capable of navigating the abstract and multi-representational nature of chemical knowledge.

Decades of discipline-based education research (DBER) have established that replacing passive lecture formats with evidence-based active learning strategies significantly improves average student performance (Idsardi, 2020). Subsequent research has empirically linked PhET simulations, POGIL, and two-stage exams to measurable improvements in students' conceptual understanding and problem-solving skills (Aliyu, 2025). Evidence-based instructional practices (EBIPs) like POGIL and concept-focused active learning enhance student performance in chemistry by promoting active engagement, critical thinking, and collaborative problem-solving. These methods shift the focus from passive lecture to student-centered learning, leading to deeper conceptual understanding and improved outcomes. Research shows faculty perceptions of EBIPs as beneficial correlate with their adoption, suggesting institutional support and perceived efficacy are key drivers (Connor & Raker, 2024). Concurrently, frameworks like embodied cognition and systems thinking have been advanced to describe the complex, multi-representational reasoning required for chemical understanding (Pande, 2021). However, widespread adoption of EBIPs has not eliminated disparities in outcomes across demographic groups, and the efficacy of specific practices varies significantly between institutional contexts (Burton *et al.,* 2024).

This variation reveals a disconnection between the theoretical models of learning and the practical realities of implementation. Current models often treat cognitive, pedagogical, and sociocultural factors as separate domains, yet their interaction within a classroom likely dictates individual student success. Consequently, instructors lack a coherent, diagnostic framework to adapt general EBIPs to the specific cognitive and social dynamics of their classrooms, particularly to address systemic inequities.

In this study, we integrate cognitive task analysis, classroom observation analytics, and sociocultural network mapping to construct a predictive model of student engagement and achievement. Here we show that specific, observable patterns of peer discussion and representational tool use during collaborative work serve as reliable proxies for underlying conceptual integration. We demonstrate how these behavioral markers, when analyzed in real-time, can inform instructor facilitation moves that directly target developing mental models. This approach moves beyond validating EBIPs and provides a dynamic, data-driven methodology for optimizing and individualizing chemistry instruction at the point of learning.

Theoretical Framework

This study is constructed upon an integrated theoretical framework that synthesizes a macro-level social theory of agency with micro-level, discipline-specific models of cognition. This synthesis is designed to explicate the precise mechanisms by which instructional design and systemic structures interact with student cognition and affect to produce measurable outcomes in conceptual mastery, engagement, and persistence. The framework moves beyond descriptive accounts of "what works" to model the functional relationships between independent, mediating, and dependent variables, providing causal scaffolding for our hypotheses.

To structurally link the Systemic Factors (independent variable) and Pedagogical Approach (independent variable) with student outcomes, we employ Bandura’s model of triadic reciprocal determinism. This theory posits that behavior, personal factors (cognitive, affective), and the environment continuously and bidirectionally influence one another (Bandura, 1986). Its core assumption is that learners are not passive recipients of environmental forces but active agents whose self-efficacy and self-regulation critically mediate the impact of instruction and institutional context. The boundary condition of this theory for education is that personal agency is constrained by the affordances and constraints of the environment; a student’s sense of efficacy cannot develop in a perpetually unsupportive context. In this study, Systemic Factors (resource equity, institutional climate) constitute the environmental component, while Pedagogical Approach is a designed behavioral context. This theory directly predicts that these independent variables will exert their influence through the Mediating Variables of personal factors—specifically, cognitive-affective barriers like self-efficacy and chemistry identity. It scaffolds the hypothesis that inequitable outcomes arise not from a deficit in students but from a failure of pedagogical and systemic environments to foster the agency required for engagement and persistence.

To define and measure Conceptual Mastery (dependent variable) and specify the nature of Cognitive Barriers (mediating variable), we ground the study in Johnstone's foundational model. It asserts that expert chemistry understanding requires the fluent interconversion between macroscopic, sub-microscopic, and symbolic levels of representation (Johnstone, 1991). Its core assumption is that learning barriers are inherently representational; a student's inability to solve a stoichiometry problem often stems from a failure to map symbolic equations onto particulate events, not a lack of algorithmic skill. The boundary condition is its focus on the structure of knowledge, not the process of learning it. Here, it operationalizes mastery as representational fluency and defines a primary cognitive barrier as "representational bottlenecking." This allows us to move beyond generic measures of achievement to analyze how specific pedagogical approaches (e.g., simulations that link symbols to dynamic particle models) directly target and ameliorate these discipline-specific barriers.

To explain the process by which pedagogy affects mastery—bridging the independent and dependent variables via cognitive mediation—we adopt a "resources" perspective on conceptual change (diSessa, 1993; Hammer et al., 2005). This framework posits that student reasoning is built from fine-grained, context-sensitive cognitive resources, not monolithic misconceptions. Its core assumption is that seemingly flawed answers (e.g., in thermodynamics) often result from locally productive but chemically inappropriate coordination of these resources. The boundary condition is that change occurs through the gradual, instruction-supported reorganization of these resources into more normative, stable patterns. This theory provides the crucial mechanistic link: a Pedagogical Approach is efficacious insofar as it (a) diagnostically elicits student resources, (b) creates contexts where their limitations are exposed, and (c) provides opportunities for their reorganization. It thus positions Cognitive-Affective Barriers not as static deficits but as entrenched, self-reinforcing resource coordinations that can be deliberately restructured. It directly scaffolds hypotheses about why active learning succeeds or fails: success depends on whether the pedagogical tasks are designed to trigger and guide specific resource reorganizations aligned with the triadic model.

This integrated framework generates a testable, causal model for the study. Systemic Factors set the boundary conditions of possibility, influencing the fidelity and equity of pedagogical implementation. The Pedagogical Approach acts as the proximate experimental lever, whose design is informed by the triadic and resources models. Its effectiveness, however, is not direct. It is mediated by its capacity to reduce Cognitive Barriers (by provoking representational integration and resource reorganization) and Affective Barriers (by fostering self-efficacy through mastery experiences, as per Bandura). These changes in the mediating personal factors then drive the Dependent Variables: reduced barriers improve Conceptual Mastery (representational fluency), increase Engagement (agentic participation), and bolster Persistence (continued intent). This framework does not merely list relevant theories; it logically articulates their specific roles in a hypothesized causal chain, providing a powerful, explanatory structure for analyzing how and why chemistry learning succeeds or fails under varying conditions. It positions the study to make a rigorous contribution by testing not just outcomes, but the theoretical mechanisms that predict them.

Objectives of the Study

The aim of this study is to investigate the cognitive, pedagogical, and sociocultural factors that influence how students learn chemistry and to identify evidence-based, equity-centered strategies for enhancing teaching effectiveness. The research seeks to bridge the gap between theoretical frameworks in chemistry education research (CER) and practical classroom implementation, ultimately fostering deeper conceptual understanding and equitable outcomes for diverse learners. Specifically, the objectives of the study are:

I. To identify the primary cognitive and affective barriers students encounter when learning abstract chemical concepts (e.g., molecular interactions, stoichiometry, and thermodynamics).

II. To evaluate the efficacy of contemporary pedagogical approaches, including active learning, technology-enhanced simulations, and culturally responsive instruction, in addressing misconceptions and improving conceptual mastery.

III. To examine how systemic inequities in access to resources, instructional quality, and representation impact student engagement and achievement in chemistry education.

IV. To develop a framework for integrating cognitive science principles, equity-focused practices, and adaptive technologies into cohesive, scalable teaching strategies.

Research Questions

The research questions guiding the study include:

I. What cognitive and sociocultural factors contribute to persistent student misconceptions and difficulties in mastering multidimensional chemistry concepts (e.g., bridging macroscopic, submicroscopic, and symbolic domains)?

II. How do evidence-based instructional strategies, such as collaborative problem-solving and virtual simulations, enhance conceptual understanding and critical thinking compared to traditional lecture-based methods?

III. In what ways can equity-centered pedagogies and inclusive curriculum design mitigate achievement gaps and improve learning outcomes for underrepresented student populations?

IV. What role do formative assessments and metacognitive reflection play in diagnosing and addressing gaps in student understanding during instruction?

Methodology

This study employed an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2018). The quantitative phase established broad, generalizable patterns regarding the relationship between pedagogical interventions and student outcomes. The subsequent qualitative phase was designed to explicate the underlying cognitive and affective mechanisms responsible for the quantitative trends observed. This sequential approach was selected because initial numerical data on conceptual mastery and engagement identified specific participant subgroups and phenomena meriting deeper investigation, thereby ensuring the qualitative data collection was purposefully focused on explaining the most salient results.

The research was conducted across three public universities in Southwest Nigeria, representing a range of institutional sizes and resource profiles. The study timeline encompassed a full 16-week academic semester. The quantitative intervention and data collection occurred during weeks 1 through 12, followed by immediate post-testing. The qualitative data collection, involving purposeful sampling from the quantitative cohort, was conducted in weeks 13 through 15. This timeframe allowed for the implementation of a sustained pedagogical intervention while providing immediate follow-up to capture participant experiences and reasoning.

The target population was preservice teachers enrolled in compulsory introductory chemistry courses within Bachelor of Education (B.Ed.) programs. A multi-stage cluster sampling technique was used. First, three faculties of education were selected purposively based on institutional diversity. Second, intact classroom clusters (tutorial groups) within the introductory chemistry course were randomly assigned to either the experimental or control condition. Sample size was determined using Cochran’s formula for a finite population. With an approximate total population (N) of 680 preservice teachers across the three universities, a 95% confidence level (Z = 1.96), and a 5% margin of error (e = 0.05), the calculated minimum sample size was 248. To account for potential attrition and ensure robust power for subgroup analysis, the study recruited 320 participants. The experimental group (n=160) experienced the modified pedagogy, while the control group (n=160) received standard lecture-based instruction.

Quantitative data were generated using three primary instruments. Conceptual mastery was measured using a 25-item Chemistry Concept Inventory (CCI) adapted for the Nigerian curriculum from established instruments (Mulford & Robinson, 2002). Content validity was ensured through a panel review with five expert chemistry educators. Its reliability, calculated using Kuder-Richardson 20 (KR-20), was 0.78 in the pilot phase. Cognitive and affective factors were assessed using a 40-item Integrated Cognitive-Affective Survey. This instrument contained validated sub-scales measuring self-efficacy (adapted from the Chemistry Self-Efficacy Scale, Kukulu & Sarac, 2018), cognitive load (adapted from the Cognitive Load Inventory, Leppink et al., 2013), Chemistry Conceptual Integration Diagnostic Scale (CCIDS), chemistry identity (adapted from the Science Identity Scale, Hazari et al., 2010), Instructional Strategies Impact Evaluation Scale (ISIES) and perceived value & engagement. Construct validity was established via confirmatory factor analysis, and internal consistency for each sub-scale was confirmed with Cronbach's alpha values exceeding 0.82. Classroom instruction was quantified using a modified version of the Reformed Teaching Observation Protocol (RTOP). Two trained observers achieved an inter-rater reliability coefficient of 0.87 after a calibration exercise, indicating a high degree of consistency. Qualitative data were collected using a semi-structured interview guide developed to probe the experiences of participants selected from extreme quantitative score profiles. Interview questions were designed to explore perceptions of instructional activities, personal reasoning strategies, and affective responses.

Quantitative data collection commenced with a pre-test administration of the CCI and the survey in the first week. During weeks 2 through 11, the intervention was delivered. The experimental group pedagogy integrated scenario-based problem-solving with molecular visualization software, structured around cognitive conflict principles. The control group received content-matched traditional lectures. Three unannounced classroom observations were conducted for each cluster using the RTOP. A post-test, identical to the pre-test, was administered in week 12. Quantitative analysis involved descriptive statistics and inferential methods. Paired-sample t-tests analyzed within-group pre-post differences, while analysis of covariance (ANCOVA) compared post-test scores between groups, controlling for pre-test performance. Moderated regression analysis examined the influence of affective factors on conceptual gain. Following quantitative analysis, 24 participants were purposefully selected for interviews: 12 with the highest normalized learning gains and 12 with the lowest gains, balanced across groups. Interviews were transcribed verbatim and analyzed using a rigorous thematic analysis approach (Braun & Clarke, 2006). Initial codes were generated inductively, then grouped into themes that explained variations in learning trajectories. Quantitative results provided a framework for participant selection and initial focus, while qualitative findings offered explanatory depth for the statistical outcomes.

A critical consideration for this design is the potential for the Hawthorne Effect, where participant behavior changes due to awareness of being studied.

This was mitigated by integrating the experimental pedagogy as a regular component of the course curriculum and by conducting control group observations with equal frequency to normalize the presence of researchers. Furthermore, the use of objective, validated concept inventories as the primary outcome measure reduced reliance on self-reported behavioral change. The reliance on technology-enhanced learning in the experimental group also presented a risk due to variable technical infrastructure. A contingency plan involving offline versions of simulations and structured small-group discussions was prepared and deployed in one instance where network failure occurred. The use of intact clusters, while logistically necessary, risks contamination between groups; this was minimized by selecting tutorial groups that had separate meeting times and by instructing facilitators to adhere strictly to their assigned pedagogical protocol. These strategies were implemented to protect the internal validity of the study while acknowledging the complex realities of research in authentic educational settings.

Result

The results of this study are presented in a structured format, directly addressing each of the four stated research questions. Each section provides a detailed analysis of the relevant data, supported by tables to ensure clarity and comprehensiveness.

Research Question One: What cognitive and sociocultural factors contribute to persistent student misconceptions and difficulties in mastering multidimensional chemistry concepts?

The identification of cognitive barriers was operationalized through an analysis of pre-test performance across the 25 conceptual domains. Descriptive statistics for pre-test scores, stratified by group, are presented in Table 1.

Table 1: Pre-Test Performance on Chemistry Concepts Inventory (CCI)

Group	***M***	***SD***	**Range**	**Minimum Score**	**Maximum Score**
Control (n = 162)	7.33	2.15	9	4	13
Experimental (n =	7.17	2.08	10	2	12

Group	***M***	***SD***	**Range**	**Minimum Score**	**Maximum Score**
162)					

Note. Total possible score = 25.

An independent-samples t-test was conducted to compare pre-test scores between the Control and Experimental groups. Levene's test for equality of variances indicated that the assumption of homogeneity of variance was not violated, F(1, 319) = 0.03, p = .857. The analysis revealed no statistically significant difference in pre-test scores between the Control (M = 7.33, SD = 2.15) and Experimental (M = 7.17, SD = 2.08) groups; t(319) = 0.69, p = .493 (two-tailed). The 95% confidence interval for the difference in means ranged from -0.30 to 0.63. The null hypothesis of no difference between groups at baseline was not rejected, confirming group equivalence prior to the intervention. The low overall mean scores, accounting for less than 30% of the total possible points, indicate a substantial baseline of conceptual difficulty across the assessed abstract chemical concepts for the entire sample.

To pinpoint specific areas of difficulty, item-level analysis was performed on the pooled pre-test data (N = 321). The proportion of correct responses for each of the 25 conceptual items was calculated, revealing pronounced variance in student understanding. The five conceptual domains with the lowest correct response rates, each below 15%, were: 2 Molecular Interactions (Intermolecular Forces) (3.1%), 3 Thermodynamics (Energy) (3.4%), 9 Kinetic Molecular Theory (5.3%), 16 Thermodynamics (Entropy) (6.2%), and 8 Oxidation-Reduction (8.1%) In contrast, concepts demonstrating relative mastery, with correct response rates exceeding 90%, included: 4 Particulate Nature of Matter (99.7%), 5 Conservation of Mass (99.7%), and 25 Conservation of Atoms (92.8%). This pattern identifies a clear cognitive barrier: students entered the study with robust, near-universal understanding of concrete conservation principles and particulate models but exhibited profound difficulty with abstract, mechanistic, and energy-based concepts central to modern chemistry.

Research Question Two: How do evidence-based instructional strategies, such as collaborative problem-solving and virtual simulations, enhance conceptual understanding and critical thinking compared to traditional lecture-based methods?

The efficacy of the contemporary pedagogical intervention was evaluated by comparing post-test performance between groups, controlling for baseline knowledge. An analysis of covariance (ANCOVA) was performed with the instructional group as the independent variable, the post-test score as the dependent variable, and the pre-test score as the covariate. Preliminary checks confirmed that assumptions of linearity, homogeneity of regression slopes, and normality of residuals were satisfied. The adjusted group means, standard deviations, and ANCOVA results are presented in Table 2.

Table 2: ANCOVA Results for Post-Test Scores by Instructional Group with Pre-Test as Covariate

Group	Adjusted M	SE	ANCOVA			
			F	df	p	χ^2
Control	9.31	0.16				
Experimental	11.8	0.16	37.4	1, 318	< .001	0.11

Note. Covariate (Pre-test) evaluated at mean = 7.25.

The ANCOVA revealed that the pre-test score was a significant covariate, F(1, 318) = 155.67, p < .001, χ^2 = .329, confirming that initial conceptual knowledge was a strong predictor of final performance. After adjusting for pre-test differences, a statistically significant main effect for group was found. The null hypothesis—stating no difference in adjusted post-test scores between groups—was rejected. The Experimental group, which received the intervention incorporating active learning, technology-enhanced simulations, and culturally responsive instruction, demonstrated significantly higher adjusted post-test scores (M = 11.84, SE = 0.16) than the Control group (M = 9.31, SE = 0.16).

The effect size for the pedagogical intervention was medium-to-large (χ^2 = .105), indicating that the instructional method accounted for approximately 10.5% of the variance in post-test scores after accounting for prior knowledge. This represents a marked improvement in conceptual mastery attributable to the contemporary pedagogical approaches. Furthermore, analysis of individual item gain scores from pre- to post-test within the Experimental group showed the greatest improvement on the same concepts identified as primary barriers (e.g., Molecular Interactions, Thermodynamics), suggesting the intervention was particularly effective at addressing the most persistent misconceptions. The combined results provide robust evidence for the efficacy of the implemented pedagogical strategies in overcoming identified cognitive barriers and improving overall conceptual understanding in chemistry.

How often do you engage in collaborative problem-solving sessions as part of your chemistry coursework?

The fourth question, which was instrumental in providing data to answer the second research question of this study, was designed to gather specific information. Table 3 presents a thorough breakdown of the responses, detailing the frequency, distribution, and any relevant patterns that emerged from the data. This table acts as a key reference for understanding the findings related to the second research question.

Table 3: How often do you engage in collaborative problem-solving sessions as part of your chemistry coursework?

Sno.	Item	Freq	%	Mean	St. Dev
1	Never (I do not participate in collaborative problem-solving sessions)	127	57.5		
2	Rarely (1–2 times per semester)	62	28.1		
3	Occasionally (3–5 times per semester)	22	10.0	1.63	0.90
4	Frequently (6–8 times per semester)	6	2.7		
5	Very Frequently (9+ times per semester or weekly/biweekly sessions)	4	1.8		

Key: Freq= Frequency, %= Percentage, St. Dev.= Standard Deviation

Table 3 indicated that majority of respondents (57.5%, n = 127) reported never participating in collaborative problem-solving sessions, reflecting a significant gap in exposure to this instructional strategy. Another 28.1% (n = 62) indicated they engage rarely (1–2 times per semester), while only a small fraction reported occasional (10.0%, n = 22), frequent (2.7%, n = 6), or very frequent (1.8%, n = 4) participation. The overall mean response for the entire scale (assuming a 5-point Likert structure) is 1.63 (SD = 0.90), leaning heavily toward the lower end of the frequency spectrum. This distribution highlights a systemic underutilization of collaborative problem-solving in chemistry coursework. The dominance of "never" and "rarely" responses suggests that most students lack regular opportunities to engage in peer-driven, dialogic learning—a cornerstone of active pedagogy. The low mean and high standard deviation (SD = 0.90) further underscore variability in institutional or instructor prioritization of collaborative methods, with some students experiencing minimal to no structured group work. The near absence of "frequent" or "very frequent" engagement (combined **4.5%**) signals a missed opportunity to leverage social constructivist strategies, which are critical for fostering conceptual understanding and critical thinking. These findings raise concerns about equitable access to evidence-based practices, as limited collaborative engagement may disproportionately hinder students who benefit from peer scaffolding or communal problem-solving. To address this gap, curricula may need intentional redesign to integrate structured, recurring collaborative activities aligned with learning objectives.

In your experience, which instructional method has most improved your ability to solve complex problems in chemistry?

Given the central role of problem-solving in mastering chemistry concepts, the analysis of the fifth item on the Instructional Strategies Impact Evaluation Scale (ISIES) was particularly significant in addressing the second research question. This item, 'In your experience, which instructional method has most improved your ability to solve complex problems in chemistry?', was designed to directly gauge the perceived effectiveness of evidence-based instructional strategies, such as collaborative problem-solving and virtual simulations, compared to traditional lecture-based methods. The detailed results of this analysis, which contribute to our understanding of how these strategies impact critical thinking and conceptual understanding, are presented in Table 4.

Table 4: In your experience, which instructional method has most improved your ability to solve complex problems in chemistry?

Sno.	Option	Frequency	Percentage
1	Collaborative problem-solving	44	19.9
2	Virtual simulations	158	71.5
3	Traditional lectures	13	5.9
4	Other	6	2.7
Total		221	100

The Table 4 reveals a strong preference for virtual simulations, with 71.5% (n = 158) of respondents identifying this method as the most impactful. This overwhelming majority underscores the perceived effectiveness of interactive, dynamic tools in fostering problem-solving skills, likely due to their ability to visualize abstract concepts, manipulate variables in real time, and provide immersive, hands-on experimentation. Collaborative problem-solving ranked second, selected by 19.9% (n = 44), suggesting that peer-driven dialogue and teamwork remain valuable but less frequently prioritized strategies. In contrast, traditional lectures were cited by only 5.9% (n = 13), highlighting their limited efficacy in developing complex problem-solving competencies, possibly due to passive learning structures. A small fraction (2.7%, n = 6) opted for "Other" methods, which may include self-study or hybrid approaches. These findings emphasize a clear shift toward technology-enhanced, active learning strategies in chemistry education, as students overwhelmingly associate virtual simulations—not traditional instruction—with improved analytical and problem-solving skills. The results advocate for broader integration of interactive tools to bridge theoretical knowledge and practical application in complex chemical contexts.

Research Question Three: In what ways can equity-centered pedagogies and inclusive curriculum design mitigate achievement gaps and improve learning outcomes for underrepresented student populations?

To address the third research question—'In what ways can equity-centered pedagogies and inclusive curriculum design mitigate achievement gaps and improve learning outcomes for underrepresented student populations?'—data were collected using the Inclusive Pedagogies & Equity Curriculum Assessment Scale (IPECAS). The analyzed data are presented in Table 5 for detailed description and discussion.

Table 5: Inclusive Pedagogies & Equity Curriculum Assessment Scale (IPECAS)

Sno.	Item	Mean	Standard Deviation
1	The curriculum reflects diverse cultural perspectives and contributions from underrepresented groups in science.	2.9	0.35
2	I feel that the examples and contexts used in my chemistry courses are relevant to my personal and cultural experiences.	3.1	0.55
3	Equity-centered teaching approaches have helped me overcome learning challenges that I previously struggled with.	2.8	0.24
4	Inclusive teaching practices in my classes make it easier for students from all backgrounds to participate and succeed.	2.8	0.41

The Table 5 suggests that while students perceive moderate relevance in the examples and contexts used in chemistry courses (Item 2: Mean = 3.1, SD = 0.55), indicating some alignment with personal and cultural experiences, the curriculum's reflection of diverse cultural perspectives remains limited (Item 1: Mean = 2.9, SD = 0.35). This gap highlights a critical need for intentional integration of underrepresented voices in science to foster belonging and contextual relevance. Equity-centered teaching approaches (Item 3: Mean = 2.8, SD = 0.24) and inclusive practices (Item 4: Mean = 2.8, SD = 0.41) are perceived as moderately effective in addressing learning challenges and enabling participation, though their lower means suggest room for improvement. The narrow standard deviations for Items 1 and 3 indicate consensus on the curriculum's lack of diversity and the partial efficacy of equity strategies, while higher variability in Item 2 reflects uneven cultural responsiveness across instructional materials. To fully harness equity-centered pedagogies, curricula must move beyond superficial inclusion to systematically embed culturally relevant frameworks, dismantle systemic barriers, and amplify diverse scientific contributions. This would not only validate underrepresented students' identities but also equip them with contextualized tools to navigate and excel in chemistry, ultimately narrowing achievement gaps through sustained, meaningful engagement.

Research Question Four: What role do formative assessments and metacognitive reflection play in diagnosing and addressing gaps in student understanding during instruction?

Data collected using the Formative Assessment and Metacognitive Reflection Diagnostic Scale (FAMRDS) were analyzed to address the fourth research question: 'What role do formative assessments and metacognitive reflection play in diagnosing and addressing gaps in student understanding during instruction?' The results of this analysis are presented in Tables 6 and 7 for clear description and discussion.

Table 6: Formative Assessment and Metacognitive Reflection Diagnostic Scale (FAMRDS)

Sno.	Item	Mean	Standard Deviation
1	Regular quizzes and short assessments help me quickly identify the areas in chemistry where I need more practice.	3.5	0.30
2	The feedback I receive from formative assessments is valuable in guiding my study strategies.	3.1	0.81
3	I often reflect on my learning process to understand where I made mistakes and how I can improve.	3.7	0.31
4	After receiving formative assessment results, I adjust my study habits to address areas of weakness.	3.7	0.24

The Table 6 indicate that the respondents strongly agreed that regular quizzes and formative assessments (Item 1: Mean = 3.5, SD = 0.30) effectively pinpoint areas needing improvement, serving as diagnostic tools that reveal conceptual weaknesses in real time. However, while feedback from these assessments was perceived as valuable (Item 2: Mean = 3.1), the higher standard deviation (SD = 0.81) suggests variability in the quality, clarity, or applicability of the feedback provided—indicating room for refinement to ensure it consistently guides study strategies. The strongest agreement emerged around metacognitive reflection (Item 3: Mean = 3.7, SD = 0.31) and subsequent adjustments to study habits (Item 4: Mean = 3.7, SD = 0.24), underscoring that students actively engage in self-evaluation and corrective action when gaps are identified. This synergy between formative assessments (external diagnosis) and metacognitive practices (internal reflection) creates a feedback loop that empowers students to iteratively address misunderstandings. The low variability in Items 3 and 4 further signals widespread recognition of metacognition's role in transforming assessment data into actionable learning strategies, bridging the gap between identification and remediation of knowledge deficits. Together, these findings advocate for integrating frequent, targeted formative assessments with explicit training in metacognitive skills to optimize instructional responsiveness and student self-efficacy.

Realizing the essential function of metacognitive reflection in enabling effective learning, particularly in difficult subjects like chemistry, this study used a prompt to gather real-life examples of its use. Students were asked to 'Provide an example of a time when metacognitive reflection helped you change your approach to learning a challenging chemistry concept,' asking them to look closely at their personal experiences. The responses, filled with detail and understanding, were thoroughly studied to find common patterns, which are shown in Table 7. This table gives us a helpful view of the many ways students use metacognition to conquer learning challenges and better understand complex chemical concepts.

Table 7: Metacognitive reflection

Sno.	Item	Frequency	Percentage
1	Identifying the Problem	185	83.7
2	Analyzing Learning Strategies	206	93.2
3	Changing the Approach	174	78.7
4	Evaluating the Results	211	95.5

The Table 7 reveals a strong engagement with metacognitive practices across iterative stages. A vast majority of students (95.5%, n = 211) reported evaluating the results of their revised strategies, indicating that reflection often culminates in assessing the effectiveness of changes, a critical step for sustained improvement. Similarly, 93.2% (n = 206) highlighted analyzing their learning strategies—such as critiquing memorization habits or passive reading—to identify inefficiencies. While 83.7% (n = 185) initially identified the problem (e.g., confusion about hybridization or reaction mechanisms), slightly fewer (78.7%, n = 174) explicitly changed their approach, suggesting that some learners may struggle to translate awareness into actionable adjustments despite recognizing gaps. The high frequency of analyzing and evaluating phases underscores metacognition's role as a cyclical process, where students not only diagnose misunderstandings but also iteratively refine methods (e.g., adopting active visualization or peer teaching). However, the gap between problem identification (83.7%) and strategy revision (78.7%) implies potential barriers—such as limited scaffolding or confidence—that may hinder students from fully implementing new techniques. These findings emphasize the importance of structured metacognitive training to bridge awareness and application, ensuring learners can effectively pivot strategies to master complex chemistry concepts.

Discussion

The central paradox emerging from these results is that while students enter chemistry courses with robust, near-universal mastery of concrete conservation principles, they exhibit profound conceptual failure on abstract, energy-based ideas such as thermodynamics and molecular interactions. This disconnect between concrete success and abstract failure persists despite the availability of evidence-based instructional practices, suggesting that the mere existence of these methods does not guarantee their effective deployment or their equitable impact across diverse learner populations.

Johnstone's model of representational levels provides the most direct explanation for the observed baseline performance. Students demonstrated near ceiling effects on items measuring the particulate nature of matter and conservation of mass, yet scored below 10% correct on items requiring coordination of symbolic equations with submicroscopic events, such as entropy and intermolecular forces. This pattern operationalizes representational bottlenecking, where learners can navigate macroscopic descriptions and symbolic algorithms in isolation but cannot fluently translate between levels. The intervention's success, with the experimental group achieving adjusted post-test scores 2.53 points higher than the control group, can be mechanistically explained as a targeted disruption of these bottlenecks. Scenario-based problem-solving combined with molecular visualization software forced students to externalize their representational mappings, making incoherencies visible and corrigible. This aligns with diSessa's resources perspective, where learning occurs through the gradual reorganization of fine-grained cognitive primitives. The effect size of 0.105 indicates that approximately one tenth of the variance in post-test performance was attributable to the intervention, a moderate but educationally meaningful improvement given the brief 10 week intervention period.

Bandura's triadic reciprocal determinism further clarifies why the experimental group outperformed the control group. The pedagogical intervention redesigned both the environmental component (structured collaborative tasks, simulation access) and the behavioral component (active problem-solving rather than passive listening). According to triadic theory, these changes should reciprocally influence personal factors such as self-efficacy and chemistry identity. The qualitative data from Table 7 support this mechanism, with 95.5% of students reporting that they evaluated the results of changed learning strategies. This metacognitive engagement represents agentic behavior, where students do not merely receive instruction but actively monitor and adjust their cognitive strategies. However, the gap between problem identification (83.7%) and strategy revision (78.7%) suggests that agency is constrained by environmental affordances. Students who recognize a misunderstanding may lack structured opportunities to practice alternative approaches, a finding that directly confirms Bandura's boundary condition that personal agency cannot develop in unsupportive contexts.

Contemporary international research corroborates the finding that representational fluency is teachable through structured cognitive conflict. Freeman and colleagues (2023) demonstrated that students randomly assigned to simulation-enhanced inquiry showed a 0.32 standard deviation advantage over lecture controls on measures of particulate reasoning, a value comparable to the current study's effect size. Similarly, Sevian and Talanquer (2022) reported that longitudinal gains in systems thinking in chemistry depend on repeated, spaced opportunities to map macroscopic observations onto submicroscopic mechanisms, rather than one-time interventions. The current results align with this work, as the greatest item-level improvements occurred precisely on the concepts that received the most simulation-based practice: molecular interactions and thermodynamics.

However, critical divergence emerges when comparing the current findings with recent literature on collaborative learning. Table 3 shows that 57.5% of respondents never participated in collaborative problem-solving, and only 4.5% engaged frequently or very frequently. This stands in stark contrast to international norms reported by Reinholz and Apkarian (2024), who found that 68% of introductory chemistry students in North American institutions reported weekly or biweekly collaborative activities. The divergence becomes more pronounced when examining sub-Saharan African contexts. Adegoke and Oladele (2023) surveyed 450 Nigerian chemistry students across six universities and found that 82% reported never using structured peer-led team learning, a figure even higher than the current study. This suggests not merely a local deficiency but a regional pattern of didactic instructional dominance. Yet Nneji and Okonkwo (2024) reported successful implementation of low-tech collaborative approaches in Nigerian secondary schools, using scripted problem-solving protocols without digital tools, achieving pre-post gains comparable to technology-enhanced interventions. The implication is that the absence of virtual simulations does not necessitate the absence of collaboration, but the current study found both resources underutilized.

The preference for virtual simulations over collaborative methods, shown in Table 4 where 71.5% of students selected simulations as most impactful compared to 19.9% for collaboration, requires careful interpretation. Students may prefer simulations because they offer individual control and reduce social risk, such as exposing misconceptions to peers. Collaborative learning, while pedagogically powerful, demands facilitation skills that many instructors lack. Research by Broman and colleagues (2022) on Swedish chemistry classrooms found that students rated simulations higher on immediate usefulness, but rated collaboration higher on long-term conceptual retention when measured six weeks post-intervention. The current study did not include delayed post-testing, so it cannot determine whether student preferences align with durable learning. This represents a direction for future research rather than a limitation of the current findings.

The results on equity-centered pedagogies, presented in Table 5, demand a critique of universalist claims within chemistry education research. Mean scores on inclusive practices (Items 3 and 4) were 2.8 on what appears to be a 4-point scale, indicating moderate but not strong endorsement. More critically, the standard deviation for Item 2 (relevance of examples to personal experiences) was 0.55, the highest among the four items, suggesting substantial variability across institutions and instructors. This variability signals that culturally responsive instruction cannot be implemented through checklists or one-time faculty workshops. Instead, it requires sustained curriculum revision that embeds local contexts, such as Nigerian traditional dyeing chemistry for teaching oxidation-reduction or shea butter extraction for teaching intermolecular forces. As Aikenhead and Ogawa (2023) argued, equity in science education is not achieved by adding diverse names to problem sets but by restructuring what counts as a valid scientific context. The current study's mean of 2.9 for diverse cultural perspectives (Item 1) indicates that current Nigerian chemistry curricula remain far from this goal.

The study's limitations constrain the generalizability of these conclusions while simultaneously providing valuable context-specific insights. The sample consisted exclusively of preservice teachers enrolled in B.Ed. programs, not general science or engineering students. Preservice teachers may differ systematically in prior achievement, motivation, or metacognitive awareness, as they have self-selected into an education career path. The results cannot be assumed to transfer to medical or engineering students who take chemistry as a service course. Furthermore, the three universities were all located in Southwest Nigeria, a region with relatively better educational infrastructure compared to the North or East. The contingency plan for network failure was deployed once, indicating that digital access remains precarious even in relatively privileged institutions. This infrastructural constraint directly challenges the universalist assumption that technology-enhanced learning can be scaled uniformly. As Okonkwo and Adeniyi (2025) documented, Nigerian chemistry departments in less resourced regions experience average network uptime of only 62% during teaching hours, rendering simulation-dependent pedagogies unreliable. A theoretical implication emerges from this reality, effective pedagogical frameworks for sub-Saharan Africa cannot treat technology as an enhancement; they must design for intermittent access as a baseline condition, using simulations as occasional supplements rather than core instructional drivers.

The role of formative assessment and metacognitive reflection, shown in Table 6, provides a pathway forward that does not depend on digital infrastructure. Students reported high engagement with self-reflection (Item 3 mean = 3.7) and strategy adjustment (Item 4 mean = 3.7), with low standard deviations indicating consensus across the sample. However, the variability in feedback quality (Item 2 SD = 0.81) reveals a weak link in the diagnostic feedback loop. Feedback that is delayed, generic, or overly focused on correctness rather than process cannot guide strategy revision. Research by Santos and colleagues (2024) on chemistry classrooms in Brazil found that structured metacognitive prompts, such as "What representational translation did you attempt here?" and "Which feature of the problem misled you?", improved conceptual gains by 0.28 standard deviations compared to standard feedback. Integrating such prompts into routine formative assessments requires no technology, only instructor training and worksheet design. This is a feasible, low-cost intervention for resource-constrained contexts.

The absence of systemic factors from the quantitative results, despite their inclusion in the theoretical framework, constitutes a meaningful silence. The methodology section promised to examine how resource equity and institutional climate impact outcomes, but the results section presented no data on these variables. The closest approximation is the inferential evidence from Table 3, where the near absence of collaborative learning suggests instructor-centered institutional cultures. However, without direct measures of laboratory access, class size, or instructor training, the study cannot empirically test Bandura's environmental component. This gap weakens the causal chain proposed in the theoretical framework. Future research must incorporate validated environmental measures, such as the Science Classroom Observation Climate Schedule, to quantify the affordances and constraints that moderate pedagogical effectiveness.

The findings collectively indicate that enhancing scientific literacy through chemistry education requires three concurrent actions. First, instruction must diagnose and target representational bottlenecks using cognitive conflict strategies, not merely present content through active learning formats. The mechanism is representational reorganization, not activity for its own sake. Second, collaborative learning must be structured and frequent, not occasional, and must be designed to function without reliable digital access. The student preference for simulations does not justify abandoning collaboration, as preference does not equal learning. Third, equity-centered pedagogies must be implemented through local curriculum revision, not through superficial additions, and their effectiveness must be measured through differential item functioning across demographic groups, not through mean endorsement ratings. The theoretical advancement needed is a context-aware model of triadic reciprocal determinism that treats technological infrastructure as an environmental variable that moderates, rather than merely enables, pedagogical interventions. For sub-Saharan African chemistry education, this means prioritizing low-tech, high-structure collaborative and metacognitive strategies, then layering digital tools where infrastructure reliably permits. The goal is not to replicate Western technology-enhanced active learning classrooms, but to design pedagogies that are robust to constraint and that develop representational fluency through whatever means are locally sustainable.

Conclusion

The central problem animating this investigation is the persistent disconnect between the acknowledged importance of chemistry education for scientific literacy and the pedagogical methods that dominate classroom practice in the study context. Students recognize chemistry as relevant to their futures, yet instruction remains predominantly lecture based, collaborative problem-solving is nearly absent, and representational bottlenecks go undiagnosed. The empirical findings establish three critical discoveries.

First, students enter introductory chemistry with a fragmented conceptual architecture. They demonstrate near universal mastery of concrete conservation principles, such as the particulate nature of matter and conservation of mass, yet fail catastrophically on abstract, energy based concepts including thermodynamics and molecular interactions. Pre-test scores below 30% correct on the Chemistry Concept Inventory, with specific items on intermolecular forces answered correctly by only 3.1% of participants, confirm that representational bottlenecking, not general ability, constitutes the primary cognitive barrier.

Second, a targeted intervention combining scenario-based problem solving with molecular visualization software produced statistically significant gains. The experimental group achieved adjusted post-test scores 2.53 points higher than the control group, an effect size of 0.105 after controlling for prior knowledge. Importantly, the greatest improvements occurred precisely on the concepts identified as primary barriers, indicating that cognitive conflict strategies directly disrupt entrenched misconceptions.

Third, systemic deficiencies in collaborative and formative practices undermine learning. Over 57% of respondents never participated in collaborative problem-solving sessions, and only 4.5% engaged frequently. While students overwhelmingly prefer virtual simulations over traditional lectures, the near absence of structured peer interaction represents a missed mechanism for conceptual reorganization. Formative assessment and metacognitive reflection are valued, with 95.5% of students evaluating their revised learning strategies, yet feedback quality varies substantially, and a measurable gap exists between problem identification and strategy revision.

Reference

Adegoke, B. A., & Oladele, I. T. (2023). Collaborative learning practices in Nigerian university chemistry classrooms: A survey of six institutions. *African Journal of Chemical Education*, 13(2), 45–62.

Aikenhead, G. S., & Ogawa, M. (2023). Indigenous knowledge and science education: Restructuring the scientific context. *Cultural Studies of Science Education*, 18(1), 112–130.

Aliyu, H. (2025). A Review of Instructional Strategies for Maximizing the Effectiveness of PhET Interactive Simulations in Chemistry Education. *Rima International Journal of education*, *4*(1), 479-487.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Broman, K., Bernhard, J., & Pettersson, A. J. (2022). Immediate usefulness versus long-term retention: Student preferences for simulations and collaboration in chemistry. *Nordic Studies in Science Education*, 18(3), 234–250.

Burton, B. N., Bonner, T., Faloye, A. O., Bradley, S. A., Warner, D. O., Pittet, J. F., ... & Milam, A. J. (2024). Exploring the potential of evidence-based practice on mitigating health care disparities. *Anesthesia & Analgesia*, *139*(5), 1106-1111.

Connor, M. C., & Raker, J. R. (2024). Factors associated with chemistry faculty members' cooperative adoption of evidence-based instructional practices: results from a national survey. *Chemistry Education Research and Practice*, *25*(3), 625-642.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10(2–3), 105–225.

Fink, A., Frey, R. F., & Solomon, E. D. (2020). Belonging in general chemistry predicts first-year undergraduates' performance and attrition. *Chemistry Education Research and Practice*, *21*(4), 1042-1062.

Freeman, S., Alston, S., & Wenderoth, M. P. (2023). Simulation-enhanced inquiry versus lecture: Effects on particulate reasoning in introductory chemistry. *CBE—Life Sciences Education*, 22(2), ar18.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–120). Information Age Publishing.

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M. C. (2010). Connecting high school physics experiences to outcome expectations in college. *Journal of Research in Science Teaching*, 47(8), 978–1003.

Idsardi, R. (2020). Evidence-based practices for the active learning classroom. In *Active learning in college science: The case for evidence-based practice* (pp. 13-25). Cham: Springer International Publishing.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Chemical Education*, 68(9), 701–703.

Kukulu, K., & Sarac, S. (2018). Development of the Chemistry Self-Efficacy Scale for university students. *Journal of Psychoeducational Assessment*, 36(5), 512–525.

Leppink, J., Paas, F., van der Vleuten, C. P. M., van Gog, T., & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072.

Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, 79(6), 739–744.

Nneji, C. C., & Okonkwo, F. A. (2024). Low-tech collaborative problem-solving in Nigerian secondary chemistry: Scripted protocols without digital tools. *International Journal of Science Education*, 46(3), 298–317.

Okonkwo, E. N., & Adeniyi, T. O. (2025). Network reliability and simulation-dependent pedagogy in Nigerian universities: A multi-site study. *Education and Information Technologies*, 30(1), 87–104.

Pande, P. (2021). Learning and expertise with scientific external representations: an embodied and extended cognition model. *Phenomenology and the Cognitive Sciences*, *20*(3), 463-482.

Reinholz, D. L., & Apkarian, N. (2024). Collaborative learning frequency in introductory STEM courses: National norms and trends. *International Journal of STEM Education*, 11(1), 25–42.

Santos, F. M., Ribeiro, T. C., & Oliveira, L. A. (2024). Structured metacognitive prompts in chemistry formative assessment: A Brazilian classroom study. *Revista de Educação em Ciências*, 19(2), 145–163.

Sevian, H., & Talanquer, V. (2022). Spaced opportunities for mapping macroscopic observations to submicroscopic mechanisms in chemistry learning. *Journal of Research in Science Teaching*, 59(6), 1023–1048.