Rima International Journal of Education (RIJE)
ISSN: 2756 – 6749 (p); 3141-2033 (e)
Volume 5, Issue 3, June. 2026; Page: 236—266 DOI: https://doi.org/10.65760/rijessu.v5.3.16

Evaluating the Effect of Predict-Observe-Explain Activities on Students' Critical Thinking in Acid-Base Reactions at Sokoto Technical Schools

***[1]Hassan Aliyu and [2]Mustapha Aliyu**

Department of Science Education, Faculty of Education, Sokoto State University, Sokoto State, Nigeria. **Email:** nagoronyo@gmail.com ORCID: https://orcid.org/0000-0003-4929-3126

Abstract

Critical thinking distinguishes students who genuinely understand chemistry from those who merely memorize facts, yet most instructional approaches fail to cultivate analytical skills in technical education contexts. Here we address the missing evidence on whether structured prediction-observation sequences can systematically build critical thinking capacities in acid-base chemistry. In a quasi-experimental study with 147 Nigerian technical college students, we demonstrate that Predict-Observe-Explain activities produce significantly larger gains in critical thinking compared to conventional instruction. Students in the POE group improved from 11.57 (SD = 3.31) to 20.92 (SD = 3.69) on a 75-point assessment, while controls improved from 11.73 (SD = 3.24) to only 15.82 (SD = 3.18). The group difference at posttest was statistically significant ($p < .001$) with a large effect size (Cohen's d = 1.24). Low achievers derived the greatest benefit, gaining 5.67 points more than their control group counterparts. Quality of POE engagement strongly predicted practical problem-solving performance (partial $r = .59$, $p < .001$) after controlling for prior ability. These findings establish that critical thinking can be systematically taught through cognitive conflict generated by prediction-observation discrepancies. For technical schools preparing students for laboratory careers, this work provides an accessible, resource-appropriate strategy for developing the analytical competence that industry demands.

Keywords: Predict-Observe-Explain; critical thinking; acid-base chemistry; technical education; cognitive conflict

Cite this as: Aliyu, H., & Aliyu, M. (2026). Evaluating the Effect of Predict-Observe-Explain Activities on Students' Critical Thinking in Acid-Base Reactions at Sokoto Technical Schools. *Rima International Journal of Education,* *5*(3), 236—266. DOI: https://doi.org/10.65760/rijessu.v5.3.16

Introduction

Critical thinking represents the core goal of chemistry education because it separates students who merely memorize facts from those who genuinely understand scientific principles. This distinction is particularly salient in acid-base chemistry, where students must connect observable phenomena—such as color changes or temperature variations—to molecular-level processes involving protons and equilibrium. Students who rely solely on memorization struggle to predict titration outcomes or understand how structural modifications affect acidity. The impact of this gap is evident in recent data. A 2023 study of first-year chemistry students across sub-Saharan Africa found that while 78% could define pH, only 23% could predict the behavior of an unfamiliar acid-base system (Okonkwo & Adebayo, 2023). These statistics underscore a fundamental failure to teach the analytical skills required for advanced university study or professional careers in the chemical industry.

The Predict-Observe-Explain (POE) strategy represents a robust pedagogical tool for addressing these learning gaps. White and Gunstone (1992) originally developed this approach to help students confront their own misconceptions. In a POE lesson, students must predict the outcome of an experiment, observe the result, and subsequently explain any discrepancies between their prediction and the observed outcome. Emerging evidence supports the efficacy of this approach in chemistry education. Treagust *et al.* (2021), for instance, found that POE implementation helped secondary students develop a more sophisticated understanding of acid strength by compelling them to articulate their mental models. Similarly, Cetin-Dindar and colleagues (2022) demonstrated that students retained titration concepts for longer periods when required to justify their predictions beforehand. The underlying mechanism is straightforward: when observations contradict expectations, cognitive conflict forces learners to reconsider their conceptual frameworks.

Despite this promising evidence, most research remains focused on whether students arrive at correct answers on assessments. We possess abundant data regarding "what" students know but insufficient information about "how" they think. Standardized tests typically obscure the processes by which students evaluate evidence or analyze variables—precisely the skills most needed in scientific practice.

A significant gap persists in our understanding of how POE activities actually cultivate critical thinking skills. The evidence base remains thin, with most studies either assuming that critical thinking develops as a byproduct of chemistry learning or employing generic tests unrelated to authentic scientific experimentation. For example, Karpudewan and colleagues (2020) demonstrated that inquiry-based chemistry instruction improved general thinking scores; however, their design could not isolate which component of the lesson—prediction, observation, or explanation—was responsible for the improvement. Yaman and Ayas (2022) examined students' explanations during POE activities but focused exclusively on answer accuracy rather than reasoning quality.

What remains lacking is direct evidence that the "cognitive struggle" inherent in POE lessons translates into specific scientific skills. We need to determine whether students can analyze variables before observing outcomes or evaluate competing theories when their predictions prove incorrect. This gap is particularly pronounced in technical education, where students are expected to graduate prepared for real-world analytical problem-solving, yet research has largely neglected this population.

The present study demonstrates that Predict-Observe-Explain activities significantly enhance specific critical thinking skills—including analysis, inference, and evaluation—among 147 technical college students in Sokoto State, Nigeria. Over a six-week period, we compared students who experienced POE instruction with those receiving conventional chemistry instruction. To measure progress, we employed a specialized rubric aligned with American Chemical Society standards, enabling assessment of whether students could apply pKa logic to predict reactions or use equilibrium principles to explain unexpected experimental outcomes.

This research advances chemistry education by demonstrating that critical thinking represents a teachable and measurable competency within specific disciplinary contexts. We focused on technical school students because they are frequently overlooked in educational research, despite being destined for laboratory and industrial careers. By equipping these students with analytical tools, we prepare them for authentic workplace challenges. If specific instructional components—such as predicting and reconciling errors—are responsible for skill development, then curricula can be designed accordingly. For teaching acid-base reactions, our findings suggest that the optimal approach for moving students from knowledge acquisition to genuine analysis is to require commitment to a prediction before observation occurs.

Objectives of the Study

This study investigated how Predict-Observe-Explain activities influence critical thinking among technical college students in acid-base chemistry. The primary aim was to determine whether this structured approach enhances students' ability to analyze chemical reactions and articulate their reasoning more effectively than traditional teaching methods. To accomplish this, we first assessed the baseline critical thinking levels of students in Sokoto technical schools, using discipline-specific indicators to evaluate their typical problem-solving approaches prior to intervention. We then examined the mechanisms through which POE steps operate, specifically investigating whether the cognitive conflict generated by incorrect predictions leads to deeper understanding of equilibrium and acid strength. Finally, we sought to connect classroom findings to practical applications by examining whether improvements in thinking translated to authentic chemistry tasks, such as interpreting titrations or predicting buffer and salt behavior. This approach enabled us to determine whether POE-generated thinking skills transfer to the types of professional tasks for which technical students are being prepared.

The specific objectives were to:

1. Establish the baseline critical thinking levels of students prior to POE intervention.
2. Identify which critical thinking components—analysis, inference, or evaluation—demonstrated the greatest change when students experienced cognitive conflict from incorrect predictions.
3. Explain how the requirement to justify observations facilitated the development of evidence-based arguments.
4. Examine the relationship between POE performance and practical problem-solving ability.
5. Determine whether POE benefits students uniformly or disproportionately advantages those with weaker initial understanding.

Research Questions

Based on the stated objectives, the following research questions guided this study:

I. What are the baseline critical thinking levels of technical college students in acid-base chemistry prior to POE intervention?

II. Which critical thinking components (analysis, inference, or evaluation) demonstrate the greatest improvement following POE instruction?

III. How does the requirement to justify observations during POE activities influence students' development of evidence-based arguments?

IV. What is the relationship between students' performance on POE activities and their ability to solve practical chemistry problems?

V. Does the effectiveness of POE instruction vary according to students' prior achievement levels?

Methodology

We employed a quasi-experimental pretest-posttest control group design to investigate how Predict-Observe-Explain activities influence students' critical thinking. This design was necessitated by the impracticality of random assignment of individual students to conditions, which would disrupt intact classes—a common constraint in educational research (Creswell & Guetterman, 2019). Chemistry education researchers frequently utilize this approach because it balances methodological rigor with the practical realities of school organization (Cohen et al., 2018). The design enabled direct comparison between groups while employing statistical techniques to account for pre-existing participant differences.

The investigation occurred over ten weeks from January to March 2025 at three government technical colleges in Sokoto State. This timeline accommodated a pretest, six weeks of instruction, an immediate posttest, and a delayed retention test administered four weeks later. This duration reflects research demonstrating that meaningful changes in cognitive processing require sustained intervention rather than brief exposure (Slisko & Cruz, 2013; Zahronada et al., 2025). Including multiple schools strengthened the generalizability of findings across the region.

This design was selected over alternatives for specific methodological reasons. A true experimental design would have offered greater internal validity but at the cost of unacceptable classroom disruption. Conversely, a single-group pretest-posttest design would confound instructional effects with maturation or testing effects (Shadish et al., 2002). The quasi-experimental approach thus represented an optimal compromise between scientific rigor and ecological validity. To strengthen causal inferences, we implemented several safeguards: verification of pretest equivalence, application of analysis of covariance to adjust for initial differences, employment of regular classroom teachers to control for novelty effects, and systematic observation to verify instructional fidelity.

The target population comprised all second-year senior chemistry students enrolled in government technical colleges in Sokoto State. These institutions specifically prepare students for technical careers where chemistry functions as an applied tool rather than abstract theory. Despite workplace demand for analytical competence, instruction in these settings predominantly relies on transmission-based methods where students passively receive information rather than actively construct understanding (Olateju et al., 2023).

School selection proceeded through purposive sampling. From the six technical colleges in the state, we selected three based on equivalent laboratory facilities, comparable enrollment numbers, and administrative willingness to participate. Functional laboratories were essential for conducting the acid-base experiments central to the POE intervention (Saunders et al., 2019). Within each selected school, we identified two existing chemistry classes and randomly assigned one to the experimental condition and the other to the control condition. This within-school randomization controlled for institutional variables such as resource availability and geographic location.

The final sample comprised 147 students: 73 in the experimental group and 74 in the control group. A priori power analysis using G*Power software indicated that detecting a medium effect size ($f = 0.25$) with 80% power at $\alpha = 0.05$ required a minimum of 128 participants (Faul et al., 2009), making the achieved sample size adequate. Participants ranged in age from 16 to 19 years. The gender distribution, while imbalanced, reflected the typical enrollment patterns of technical institutions in the region. Ethical protocols were rigorously followed: we obtained approval from the Ministry of Education, consent from school principals and parents, and voluntary assent from all students with assurance that withdrawal carried no academic penalty.

The experimental group engaged with acid-base chemistry through a structured Predict-Observe-Explain framework adapted from White and Gunstone (1992) and refined based on contemporary chemistry education research (Kala et al., 2013; Yaman et al., 2019). Over six weeks, students completed fourteen POE cycles spanning five core topics: acid-base definitions, pH and strength, titration, buffer solutions, and salt hydrolysis. Each cycle followed a consistent three-phase structure. During prediction, students responded to a specific scenario, documenting both their forecast and their reasoning. The observation phase involved teacher demonstration or student hands-on experimentation with careful recording of results. The explanation phase engaged students in small-group discussion to reconcile predictions with observations, followed by whole-class synthesis guided by the teacher.

The control group received instruction on identical content over the same six-week period using conventional teaching methods. Teachers presented concepts through lecture, wrote notes on the board, and conducted demonstrations while students observed. Students could ask questions but did not engage in structured prediction or systematic explanation of discrepancies. This approach represents the predominant mode of chemistry instruction in Nigerian technical schools, providing a realistic baseline against which to evaluate POE effectiveness.

Research assistants observed 30% of all class sessions using a structured protocol (Cetin-Dindar et al., 2022), verifying that experimental teachers completed all three POE phases and that control teachers adhered to conventional methods. Teacher logs and student work samples provided additional verification of adherence to assigned instructional conditions.

Data collection employed three quantitative instruments and two qualitative protocols developed specifically for this study.

Critical Thinking in Acid-Base Chemistry Assessment (CTABCA) served as the primary outcome measure. This 25-item constructed-response instrument assessed three components derived from Facione's (1990) framework: analysis of relationships among chemical variables, inference of logical conclusions from evidence, and evaluation of competing claims. Items presented authentic scenarios such as interpreting titration graphs or predicting buffer behavior, requiring students to both solve problems and justify their reasoning. Scoring followed a rubric adapted from American Chemical Society guidelines (ACS, 2018), with each component scored 0 to 3 for a maximum total of 75. Interrater reliability, assessed on a random sample of 30 protocols, yielded an intraclass correlation coefficient of 0.87, indicating strong agreement.

Content validity was established through expert review by three chemistry educators and two educational psychologists, who evaluated items for alignment with critical thinking definitions and the Nigerian technical college curriculum. Construct validity was examined through confirmatory factor analysis using pretest data, which supported the three-factor structure with acceptable fit indices ($\chi^2/df = 2.31$, CFI = 0.91, RMSEA = 0.06). Reliability was established through a pilot study with 45 students, yielding Cronbach's alpha of 0.82 for the total scale and 0.78, 0.80, and 0.75 for the analysis, inference, and evaluation subscales respectively.

Practical Problem-Solving Test (PPST) assessed application of acid-base concepts to authentic technical problems. This 10-item performance-based assessment presented scenarios from industrial and laboratory contexts requiring both correct solutions and written justifications. Scoring followed a rubric adapted from the Assessment of Practical Skills in Chemistry framework (Bennett & Kennedy, 2021).

POE Activity Response Sheets completed by experimental students during each cycle, provided process data on the evolution of student reasoning.

Semi-Structured Interviews with a purposive sample of 18 students representing varied achievement levels explored experiences with cognitive conflict, explanation construction, and perceived transfer of learning.

Data collection proceeded in four phases over ten weeks. Phase 1 involved CTABCA pretest administration to all participants under standardized conditions. Phase 2 comprised the six-week intervention with ongoing classroom observations and collection of POE response sheets. Phase 3 included posttest and PPST administration. Phase 4 involved retention testing and semi-structured interviews.

Quantitative data were analyzed using SPSS version 27. Preliminary analyses examined assumptions of normality and homogeneity of variance. Independent samples t-tests verified pretest equivalence. Mixed-design ANOVA assessed changes across time points with partial eta squared quantifying effect sizes. Analysis of covariance provided additional adjustment for pretest differences, with separate ANCOVAs for each critical thinking component. Qualitative data were analyzed thematically following the framework approach of Gale and colleagues (2013), involving familiarization, framework development, indexing, charting, and interpretation.

Result

The results are presented in five sections corresponding to the research objectives. First, descriptive statistics and preliminary analyses are reported, including baseline equivalence between groups. Second, the overall effects of POE instruction on critical thinking are examined using mixed between-within subjects ANOVA. Third, differential effects on analysis, inference, and evaluation components are analyzed through separate mixed ANOVAs and confirmed with ANCOVA. Fourth, variation in POE effectiveness across prior achievement levels is investigated using two-way ANOVA and simple effects analyses. Fifth, the relationship between POE performance and practical problem-solving ability is explored through Pearson correlations and partial correlations controlling for pretest scores. Finally, qualitative findings from POE response sheets and interviews are presented thematically to illuminate the cognitive processes underlying the quantitative patterns. All statistical analyses were conducted using SPSS version 27 with alpha set at .05.

Descriptive Statistics and Preliminary Analyses

Prior to conducting inferential analyses, the data were screened for missing values, outliers, and violations of parametric assumptions. No missing data were present, and inspection of boxplots revealed no extreme outliers requiring removal. Descriptive statistics for pretest and posttest critical thinking scores by group are presented in Table 1.

Table 1: Descriptive Statistics for Critical Thinking Scores by Group and Time

Group	Time	n	Mean	Standard Deviation
Control	Pretest	73	11.73	3.24
Control	Posttest	73	15.82	3.18
Experimental	Pretest	74	11.57	3.31
Experimental	Posttest	74	20.92	3.69

Table 1 presents the descriptive statistics for critical thinking scores across both groups and time points. At pretest, the control group ($n = 73$) recorded a mean score of 11.73 ($SD = 3.24$), while the experimental group ($n = 74$) recorded a comparable mean of 11.57 ($SD = 3.31$), indicating equivalence in baseline critical thinking abilities between the two groups prior to the intervention. At posttest, both groups demonstrated improvement, with the control group increasing to a mean of 15.82 ($SD = 3.18$). However, the experimental group exhibited a substantially larger gain, achieving a mean posttest score of 20.92 ($SD = 3.69$). These findings suggest that while both instructional approaches contributed to improved critical thinking, students exposed to Predict-Observe-Explain activities demonstrated markedly higher performance compared to those who received conventional instruction.

Baseline Equivalence of Groups

To address Objective 1 (establishing the starting level of critical thinking), an independent samples t-test was conducted to compare pretest critical thinking scores between the experimental and control groups. The analysis revealed no statistically significant difference between the experimental group ($M = 11.57$, $SD = 3.31$) and the control group ($M = 11.73$, $SD = 3.24$), $t(145) = -0.30$, $p = .767$. The magnitude of the difference in means was negligible (mean difference = -0.16, 95% CI [-1.22, 0.90]), confirming that the two groups demonstrated equivalent critical thinking abilities prior to the instructional intervention. This baseline equivalence strengthened the internal validity of the study by ruling out pre-existing differences as an alternative explanation for any post-intervention effects.

Effects of POE Activities on Critical Thinking

A mixed between-within subjects analysis of variance was conducted to evaluate the effect of group (experimental vs. control) and time (pretest, posttest) on students' critical thinking scores. The analysis addressed the primary research aim of determining whether POE activities enhanced critical thinking beyond conventional instruction. Results of the mixed ANOVA are presented in Table 2.

Table 2: Mixed Between-Within Subjects ANOVA for Critical Thinking Scores

Source	SS	df	MS	F	p	Partial η²
Between Subjects						
Group	574.18	1	574.18	27.64	<.001	0.16
Error (Between)	3012.45	145	20.78			
Within Subjects						
Time	3741.86	1	3741.86	602.17	<.001	0.806
Time × Group	452.37	1	452.37	72.8	<.001	0.334
Error (Within)	901.23	145	6.22			

Table 2 presents the results of the mixed between-within subjects analysis of variance conducted to examine the effects of group and time on students' critical thinking scores. The analysis revealed a statistically significant main effect for time, Wilks' $\Lambda = 0.194$, $F(1, 145) = 602.17$, $p < .001$, partial $\eta^2 = .806$, indicating that students across both groups demonstrated significant improvement from pretest to posttest. A significant main effect for group was also observed, $F(1, 145) = 27.64$, $p < .001$, partial $\eta^2 = .160$, with the experimental group ($M = 16.25$, $SE = 0.38$) achieving higher scores across both time points compared to the control group ($M = 13.77$, $SE = 0.38$). Critically, a significant time × group interaction emerged, Wilks' $\Lambda = 0.666$, $F(1, 145) = 72.80$, $p < .001$, partial $\eta^2 = .334$, indicating that the improvement trajectories differed between the two groups. The interaction effect, accounting for 33.4% of the variance, suggests that students in the experimental group experienced substantially greater gains in critical thinking compared to their control group counterparts.

Objective 2: Differential Effects on Critical Thinking Components

Prior to conducting the main analyses for Objective 2, preliminary checks were performed to verify that the assumptions for mixed ANOVA and analysis of covariance were satisfied for each critical thinking component. Levene's test for equality of variances was conducted for analysis, inference, and evaluation subscale scores at pretest and posttest. For the analysis subscale, Levene's test was not significant at pretest, $F(1,145) = 0.43$, $p = .514$, or at posttest, $F(1,145) = 2.01$, $p = .158$. For the inference subscale, the assumption was satisfied at pretest, $F(1,145) = 0.89$, $p = .347$, and at posttest, $F(1,145) = 1.56$, $p = .214$. For the evaluation subscale, Levene's test was also not significant at pretest, $F(1,145) = 1.23$, $p = .269$, or at posttest, $F(1,145) = 3.12$, $p = .079$. These results confirmed that the homogeneity of variance assumption was tenable for all subsequent analyses.

Box's test of equality of covariance matrices was examined for each component and was not significant, indicating that the assumption of homogeneity of intercorrelations was met. Examination of skewness and kurtosis values for each subscale at each time point revealed values within acceptable ranges, supporting the normality assumption for the planned parametric analyses.

Descriptive statistics for the three critical thinking components by group and time are presented in Table 3.

Table 3: Descriptive Statistics for Critical Thinking Components by Group and Time

Component	Group	Time	Mean	Standard Deviation	n
Analysis	Control	Pretest	4.12	1.24	73
Analysis	Control	Posttest	5.48	1.18	73
Analysis	Experimental	Pretest	4.08	1.21	74
Analysis	Experimental	Posttest	7.21	1.32	74
Inference	Control	Pretest	3.95	1.18	73
Inference	Control	Posttest	5.31	1.22	73
Inference	Experimental	Pretest	3.89	1.15	74
Inference	Experimental	Posttest	6.94	1.28	74
Evaluation	Control	Pretest	3.66	1.08	73
Evaluation	Control	Posttest	5.03	1.14	73
Evaluation	Experimental	Pretest	3.6	1.05	74
Evaluation	Experimental	Posttest	6.77	1.25	74

Table 3 presents the descriptive statistics for the three critical thinking components—analysis, inference, and evaluation—across groups and time points. For the analysis component, the control group demonstrated an increase from pretest (M = 4.12, SD = 1.24) to posttest (M = 5.48, SD = 1.18), while the experimental group showed a substantially larger improvement from pretest (M = 4.08, SD = 1.21) to posttest (M = 7.21, SD = 1.32). Similarly, for the inference component, the control group improved from pretest (M = 3.95, SD = 1.18) to posttest (M = 5.31, SD = 1.22), whereas the experimental group exhibited more pronounced gains from pretest (M = 3.89, SD = 1.15) to posttest (M = 6.94, SD = 1.28). For the evaluation component, the control group increased from pretest (M = 3.66, SD = 1.08) to posttest (M = 5.03, SD

= 1.14), while the experimental group demonstrated the most substantial improvement from pretest (M = 3.60, SD = 1.05) to posttest (M = 6.77, SD = 1.25). These findings indicate that students exposed to POE instruction outperformed their control group counterparts across all three critical thinking components, with the experimental group consistently achieving higher posttest mean scores in analysis, inference, and evaluation.

Table 4: Mixed Between-Within Subjects ANOVAs for Critical Thinking Components

Component	Source	SS	df	MS	F	p	Partial η²
Analysis	Between Subjects						
	Group	97.42	1	97.42	32.18	<.001	0.182
	Error (Between)	438.91	145	3.03			
	Within Subjects						
	Time	423.67	1	423.67	412.56	<.001	0.74
	Time × Group	89.34	1	89.34	87.01	<.001	0.375
	Error (Within)	148.92	145	1.03			
Inference	Between Subjects						
	Group	81.23	1	81.23	28.45	<.001	0.164
	Error (Between)	414.18	145	2.86			
	Within Subjects						
	Time	378.45	1	378.45	398.37	<.001	0.733
	Time × Group	76.89	1	76.89	80.94	<.001	0.358
	Error (Within)	137.78	145	0.95			
Evaluation	Between Subjects						
	Group	74.56	1	74.56	26.89	<.001	0.156
	Error (Between)	402.23	145	2.77			
	Within Subjects						
	Time	356.78	1	356.78	412.34	<.001	0.74
	Time × Group	71.23	1	71.23	82.31	<.001	0.362
	Error (Within)	125.56	145	0.87			

Table 4 presents the results of separate mixed between-within subjects analyses of variance conducted for each critical thinking component—analysis, inference, and evaluation. For the analysis component, the analysis revealed a significant main effect for time, Wilks' $\Lambda = 0.260$, $F(1, 145) = 412.56$, $p < .001$, partial $\eta^2 = .740$, and a significant main effect for group, F(1,

145) = 32.18, $p < .001$, partial $\eta^2 = .182$. Critically, a significant time × group interaction was observed, Wilks' $\Lambda = 0.625$, $F(1, 145) = 87.01$, $p < .001$, partial $\eta^2 = .375$, indicating differential improvement trajectories between groups. For the inference component, significant main effects were found for time, Wilks' $\Lambda = 0.267$, $F(1, 145) = 398.37$, $p < .001$, partial $\eta^2 = .733$, and for group, $F(1, 145) = 28.45$, $p < .001$, partial $\eta^2 = .164$, with a significant time × group interaction, Wilks' $\Lambda = 0.642$, $F(1, 145) = 80.94$, $p < .001$, partial $\eta^2 = .358$. For the evaluation component, significant main effects emerged for time, Wilks' $\Lambda = 0.260$, $F(1, 145) = 412.34$, $p < .001$, partial $\eta^2 = .740$, and for group, $F(1, 145) = 26.89$, $p < .001$, partial $\eta^2 = .156$, with a significant time × group interaction, Wilks' $\Lambda = 0.638$, $F(1, 145) = 82.31$, $p < .001$, partial $\eta^2 = .362$. Comparison of interaction effect sizes revealed that the analysis component demonstrated the largest interaction effect (partial $\eta^2 = .375$), followed closely by evaluation (partial $\eta^2 = .362$) and inference (partial $\eta^2 = .358$), indicating that all three components were substantially responsive to the POE intervention, with analytical reasoning showing the greatest differential improvement relative to conventional instruction.

A series of one-way analyses of covariance was conducted to provide a confirmatory check on the robustness of the findings and to increase statistical power by controlling for pretest differences. For each critical thinking component—analysis, inference, and evaluation—posttest scores served as the dependent variable, group membership as the independent variable, and pretest scores on the respective component as the covariate. The assumption of homogeneity of regression slopes was tested and satisfied for each component, as the group × pretest interaction was not significant for analysis ($F(1, 143) = 0.89$, $p = .347$), inference ($F(1, 143) = 1.23$, $p = .269$), or evaluation ($F(1, 143) = 0.67$, $p = .414$), confirming that the relationship between pretest and posttest scores did not differ significantly between groups. As presented in Table 5, the ANCOVA results revealed significant main effects for group across all three components: analysis ($F(1, 144) = 93.67$, $p < .001$, partial $\eta^2 = .396$), inference ($F(1, 144) = 87.45$, $p < .001$, partial $\eta^2 = .378$), and evaluation ($F(1, 144) = 87.12$, $p < .001$, partial $\eta^2 = .377$), with the experimental group demonstrating substantially higher adjusted mean scores in each component compared to the control group, thereby confirming the differential responsiveness of the critical thinking components to POE instruction.

Table 5: Analysis of Covariance for Critical Thinking Components by Group with Pretest Covariate

Component	Source	SS	df	MS	F	p	Partial η^2
Analysis	Pretest (Covariate)	68.34	1	68.34	56.89	<.001	0.285
	Group	112.45	1	112.45	93.67	<.001	0.396
	Error	171.23	144	1.19			
Inference	Pretest (Covariate)	58.92	1	58.92	52.34	<.001	0.267
	Group	98.34	1	98.34	87.45	<.001	0.378
	Error	162.45	144	1.13			
Evaluation	Pretest (Covariate)	52.78	1	52.78	51.23	<.001	0.262
	Group	89.67	1	89.67	87.12	<.001	0.377
	Error	148.23	144	1.03			

Table 5 presents the results of the analysis of covariance conducted to examine the effects of group membership on each critical thinking component while controlling for pretest scores. For the analysis component, the pretest covariate was significantly related to posttest scores, $F(1, 144) = 56.89$, $p < .001$, partial $\eta^2 = .285$, and a significant main effect for group emerged, $F(1, 144) = 93.67$, $p < .001$, partial $\eta^2 = .396$, with the experimental group demonstrating a substantially higher adjusted mean ($M\_adj = 7.18$, $SE = 0.13$) compared to the control group ($M\_adj = 5.51$, $SE = 0.13$). For the inference component, the pretest covariate was significant, $F(1, 144) = 52.34$, $p < .001$, partial $\eta^2 = .267$, and a significant group effect was observed, $F(1, 144) = 87.45$, $p < .001$, partial $\eta^2 = .378$, with the experimental group ($M\_adj = 6.91$, $SE = 0.12$) outperforming the control group ($M\_adj = 5.34$, $SE = 0.12$). For the evaluation component, the pretest covariate was significant, $F(1, 144) = 51.23$, $p < .001$, partial $\eta^2 = .262$, and a significant group effect was found, $F(1, 144) = 87.12$, $p < .001$, partial $\eta^2 = .377$, favoring the experimental group ($M\_adj = 6.74$, $SE = 0.12$) over the control group ($M\_adj = 5.06$, $SE = 0.12$). The effect sizes indicate that the analysis component demonstrated the largest adjusted group effect (partial $\eta^2 = .396$), followed by inference (partial $\eta^2 = .378$) and evaluation (partial $\eta^2 = .377$), confirming the differential responsiveness of the critical thinking components with analysis being most substantially enhanced by POE instruction.

Objectives 5: Variation of POE Effects by Prior Achievement Level

To address Objective 5, researchers classified participants using pretest critical thinking scores. They calculated the 33rd and 67th percentiles from the full sample, resulting in cut scores of 10 and 13. This produced low (n=48), medium (n=51), and high (n=48) achiever groups with approximately equal sizes suitable for factorial analysis. Gain scores represented improvement magnitude from pretest to posttest. Preliminary examination revealed no substantial normality violations. Levene's test for equality of variances was not statistically significant, $F(5,141)=1.87$, $p=.104$, confirming homogeneity of variance across the six group-achievement combinations.

Descriptive statistics for gain scores by group and achievement level are presented in Table 6.

Table 6: Mean Gain Scores by Group and Prior Achievement Level

Achievement Level	Group	n	Mean Gain	Standard Deviation
Low	Control	24	3.79	2.13
Low	Experimental	24	9.46	2.67
Medium	Control	26	4.23	1.96
Medium	Experimental	25	8.92	2.34
High	Control	23	4.91	2.08
High	Experimental	25	8.56	2.41

Table 6 presents the mean gain scores in critical thinking by group and prior achievement level. For low achievers, the experimental group demonstrated a substantially higher mean gain ($M = 9.46$, $SD = 2.67$) compared to the control group ($M = 3.79$, $SD = 2.13$), representing a difference of 5.67 points. Among medium achievers, the experimental group ($M = 8.92$, $SD = 2.34$) outperformed the control group ($M = 4.23$, $SD = 1.96$) by 4.69 points. For high achievers, the experimental group ($M = 8.56$, $SD = 2.41$) exceeded the control group ($M = 4.91$, $SD = 2.08$) by 3.65 points. These findings indicate that while students across all achievement levels benefited from POE instruction, low achievers derived the greatest relative benefit, as evidenced by the largest gain difference between experimental and control groups at this level. The pattern suggests that POE activities may function as an equalizing pedagogical strategy, providing greater support to students with weaker initial critical thinking abilities.

Researchers conducted a two-way between-subjects ANOVA with group and prior achievement level as independent variables and gain score as the dependent variable, with results presented in Table 7.

Table 7: Two-Way Analysis of Variance for Gain Scores by Group and Achievement Level

Source	SS	df	MS	F	p	Partial η^2
Group	789.34	1	789.34	156.23	<.001	0.526
Achievement Level	12.45	2	6.23	1.23	0.295	0.017
Group × Achievement Level	41.67	2	20.84	4.12	0.018	0.055
Error	712.89	141	5.06			

Table 7 presents the results of the two-way analysis of variance conducted to examine the main and interaction effects of group and prior achievement level on students' critical thinking gain scores. The analysis revealed a statistically significant main effect for group, $F(1, 141) = 156.23$, $p < .001$, partial $\eta^2 = .526$, indicating that students in the experimental group ($M = 8.98$, $SE = 0.26$) demonstrated substantially greater gains in critical thinking compared to those in the control group ($M = 4.30$, $SE = 0.26$). The main effect for prior achievement level was not statistically significant, $F(2, 141) = 1.23$, $p = .295$, partial $\eta^2 = .017$, suggesting that gain scores did not differ across achievement levels when collapsing across instructional groups. However, a statistically significant group × achievement level interaction emerged, $F(2, 141) = 4.12$, $p = .018$, partial $\eta^2 = .055$, indicating that the magnitude of POE effectiveness depended on students' prior achievement level. This modest interaction effect, accounting for 5.5% of the variance in gain scores, demonstrates that POE instruction interacts with entering ability to produce differential gains, with the experimental group significantly outperforming the control group across all achievement levels, though the size of this advantage varied.

Researchers conducted simple main effects analyses with a Bonferroni adjustment, setting alpha at .017 for each test to control Type I error, and presented the results in Table 8.

Table 8: Simple Main Effects of Group Within Each Achievement Level

Achievement Level	Comparison	Mean Difference	SE	F	df	p	Partial η^2
Low	Experimental vs. Control	5.67	0.65	76.23	1, 141	<.001	0.351
Medium	Experimental vs. Control	4.69	0.63	55.34	1, 141	<.001	0.282
High	Experimental vs. Control	3.65	0.64	32.45	1, 141	<.001	0.187

Table 8 presents the results of the simple main effects analysis examining the differences between experimental and control groups within each achievement level. For low achievers, the experimental group demonstrated a significantly higher mean gain compared to the control group, with a mean difference of 5.67 points (SE = 0.65), $F(1, 141) = 76.23$, $p < .001$, partial $\eta^2 = .351$, representing a large effect size. Among medium achievers, the experimental group significantly outperformed the control group with a mean difference of 4.69 points (SE = 0.63), $F(1, 141) = 55.34$, $p < .001$, partial $\eta^2 = .282$, also indicating a large effect. For high achievers, the experimental group showed a significant advantage over the control group with a mean difference of 3.65 points (SE = 0.64), $F(1, 141) = 32.45$, $p < .001$, partial $\eta^2 = .187$, representing a medium to large effect. These findings demonstrate that the experimental group significantly outperformed the control group at every achievement level, with the magnitude of the experimental advantage being largest for low achievers (partial $\eta^2 = .351$) and progressively decreasing for medium (partial $\eta^2 = .282$) and high achievers (partial $\eta^2 = .187$), indicating that POE instruction was particularly beneficial for students with weaker initial critical thinking skills.

Objective 4: Relationship Between POE Performance and Practical Problem-Solving

To address Objective 4, which sought to examine whether students' engagement with and performance on POE activities predicted their ability to solve authentic practical problems in acid-base chemistry, a series of correlation analyses was conducted. These analyses were restricted to participants in the experimental group ($n = 74$), as only these students completed POE activity response sheets during the intervention. For each student in the experimental group, an average POE performance score was computed by rating their responses across all fourteen POE cycles using a rubric adapted from the CTABCA criteria. This rubric assessed the quality of predictions, the accuracy of observations, and the depth of explanations on a four-point scale, yielding a composite score ranging from 0 to 12 per cycle. These cycle scores were then averaged to produce a single POE performance score for each student.

The Practical Problem-Solving Test (PPST) was administered at posttest and assessed students' ability to apply acid-base concepts to authentic technical problems, including titration curve interpretation, buffer capacity prediction, and salt hydrolysis analysis. PPST scores ranged from 0 to 40, with higher scores indicating greater practical problem-solving competence.

Prior to conducting correlation analyses, the data were screened for violations of the assumptions underlying Pearson product-moment correlation. Inspection of scatterplots for the relationship between POE performance and PPST scores revealed a linear pattern with no evidence of curvilinearity. Examination of boxplots identified no extreme outliers requiring removal. Shapiro-Wilk tests of normality were not statistically significant for POE performance ($W = 0.98$, $p = .234$) or PPST scores ($W = 0.97$, $p = .189$), supporting the normality assumption. Descriptive statistics for the variables included in the correlation analyses are presented in Table 9.

Table 9: Descriptive Statistics for POE Performance, PPST Scores, and Pretest Scores (Experimental Group Only)

Variable	n	Mean	Standard Deviation	Minimum	Maximum
POE Performance Score	74	8.47	1.82	4.21	11.83
Practical Problem-Solving Test (PPST)	74	32.16	5.34	18	40
Critical Thinking Pretest	74	11.57	3.31	4	20

Table 9 presents the descriptive statistics for POE performance scores, Practical Problem-Solving Test (PPST) scores, and critical thinking pretest scores for the experimental group participants ($n = 74$). The mean POE performance score was 8.47 ($SD = 1.82$), with scores ranging from 4.21 to 11.83, indicating considerable variability in students' engagement quality with POE activities. The Practical Problem-Solving Test yielded a mean score of 32.16 ($SD = 5.34$), with scores ranging from 18 to 40, reflecting a relatively wide distribution of practical problem-solving abilities among experimental group students. The critical thinking pretest scores showed a mean of 11.57 ($SD = 3.31$), with scores ranging from 4 to 20, suggesting that participants represented a diverse range of initial critical thinking abilities prior to the intervention. These descriptive statistics provide the foundation for examining the relationships between POE engagement, practical problem-solving performance, and prior critical thinking ability.

Table 10: Pearson Product-Moment Correlations Among POE Performance, PPST Scores, and Pretest Scores

Variable	POE Performance	PPST Score	Pretest Score
POE Performance	—		
PPST Score	.68***	—	
Pretest Score	.42***	.58***	—
*** p<.001p<.001			

Table 10 presents the Pearson product-moment correlation coefficients examining the relationships among POE performance scores, Practical Problem-Solving Test (PPST) scores, and critical thinking pretest scores for the experimental group participants ($n = 74$). The analysis revealed a strong positive correlation between POE performance and PPST scores, $r(72) = .68$, $p < .001$, indicating that students who demonstrated deeper engagement with POE activities achieved higher scores on practical problem-solving tasks, with shared variance of approximately 46%. A moderate positive correlation was observed between POE performance and pretest critical thinking scores, $r(72) = .42$, $p < .001$, accounting for 18% of shared variance. Additionally, a strong positive correlation emerged between pretest scores and PPST scores, $r(72) = .58$, $p < .001$, explaining 34% of the variance. These findings suggest that while prior critical thinking ability is associated with both POE engagement and practical problem-solving performance, the strong relationship between POE performance and PPST scores indicates that quality engagement with prediction, observation, and explanation processes substantially contributes to students' practical competence in acid-base chemistry.

Table 11: Partial Correlation Between POE Performance and PPST Scores Controlling for Pretest

Control Variable	Variables	Partial Correlation	df	p
Pretest Score	POE Performance & PPST	0.59	71	<.001

Table 11 presents the results of the partial correlation analysis examining the relationship between POE performance and Practical Problem-Solving Test (PPST) scores while controlling for pretest critical thinking scores. After accounting for the influence of initial critical thinking ability, the partial correlation between POE performance and PPST scores remained significant and strong, $r\_partial(71) = .59$, $p < .001$. This finding indicates that even after removing the variance attributable to pretest scores, POE performance uniquely accounted for approximately 35% of the variance in practical problem-solving scores. The sustained strength of this relationship demonstrates that quality engagement with prediction, observation, and explanation processes predicts practical competence in acid-base chemistry beyond what entering critical thinking levels would suggest, confirming that the POE intervention itself contributed substantially to the development of students' practical problem-solving abilities.

Table 12: Partial Correlations Between POE Performance and Posttest Critical Thinking Components Controlling for Respective Pretest Components

Component	Partial Correlation with POE Performance	df	p
Analysis (Posttest)	0.54	71	<.001
Inference (Posttest)	0.51	71	<.001
Evaluation (Posttest)	0.56	71	<.001

Table 12 presents the partial correlations between POE performance and each posttest critical thinking component—analysis, inference, and evaluation—after controlling for the respective pretest component scores. The analysis revealed that POE performance demonstrated significant positive partial correlations with all three critical thinking components: analysis, $r\_partial(71) = .54$, $p < .001$; inference, $r\_partial(71) = .51$, $p < .001$; and evaluation, $r\_partial(71) = .56$, $p < .001$. Evaluation showed the strongest unique association with POE engagement, accounting for approximately 31% of the variance, followed by analysis at 29% and inference at 26%. These findings indicate that deeper engagement with POE activities directly contributes to the development of specific critical thinking skills in acid-base chemistry, with evaluation being most strongly associated with POE performance, followed closely by analysis and inference, even after accounting for students' initial abilities in each respective component.

Qualitative Findings: Cognitive Processes Underlying POE Effects

Thematic analysis of POE response sheets and interviews with 18 purposefully selected students (six each from low, medium, and high prior achievement levels) elucidated the mechanisms behind the intervention's quantitative effects. Following the framework approach of Gale and colleagues (2013), the analysis revealed three primary themes explaining student experiences. Cognitive conflict acted as a catalyst for analytical engagement. Students progressed from simple description to constructing mechanistic reasoning in their explanations. Furthermore, participants reported transferring reasoning skills learned through POE activities to practical chemistry problems. These qualitative findings provide experiential evidence for how POE activities enhance critical thinking in acid-base chemistry.

Theme 1: Cognitive Conflict as a Catalyst for Analytical Engagement

The first theme addressed Objective 2 by revealing how cognitive conflict catalyzes analytical reasoning. Students consistently described prediction-observation discrepancies as memorable and productive experiences that forced deeper engagement with acid-base concepts. A low-achieving student detailed his experience during an acid-base titration cycle. Direct observation contradicted his prediction that hydrochloric acid-ammonia neutralization would yield pH 7. The student engaged in checking procedures, sought peer verification, and ultimately revised his mental model about base strength. He explained that confusion prompted him to question whether all bases function identically, transforming previously memorized information into understood knowledge. A high-achieving student described a similar experience with buffer capacity demonstration. The near-constant pH reading conflicted with his prediction of measurable change. This discrepancy connected the symbolic Henderson-Hasselbalch equation to observable phenomenon, giving quantitative meaning to the concept of resistance. Both accounts demonstrate that cognitive conflict triggers analytical processes regardless of achievement level, though resolution pathways differ.

Analysis of POE response sheets provided written evidence of these analytical processes. In a salt hydrolysis cycle, one student initially predicted ammonium chloride solution would test neutral. Observing acidic results prompted mechanistic reasoning about ion behavior and proton transfer. The student progressed from a simple categorical rule to constructing an explanation based on parent acid and base identities. This written response illustrates conceptual reorganization rather than mere correction.

A contrasting pattern emerged in approximately 12 percent of low-achiever response sheets, predominantly during early intervention cycles. These students attributed discrepancies to procedural error rather than engaging conceptual implications. However, longitudinal analysis showed these same students demonstrating substantive engagement in later cycles, suggesting that resistance to cognitive conflict diminishes with sustained POE exposure. The qualitative data thus illuminate the mechanisms through which prediction-observation discrepancies transform student thinking about acid-base chemistry.

Theme 2: Evolution of Explanation Construction from Description to Mechanistic Reasoning

The second theme addressed Objective 3 by demonstrating how the explanation requirement transformed students' evidence-based argumentation. Analysis of POE response sheets across the six-week intervention revealed a clear progression from descriptive accounts to mechanistic reasoning. Early explanations typically restated observations or invoked memorized rules without elaboration. Later explanations increasingly incorporated chemical principles, referenced specific concepts such as equilibrium and proton transfer, and demonstrated reasoning about underlying mechanisms. A low-achieving student's responses illustrated this progression. In Week 2, the student explained that hydrochloric acid produced a brighter light than acetic acid because it makes more ions. By Week 5, the same student explained buffer action by describing how acetate ions react with added H+ ions, connecting the observation to equilibrium shifting according to Le Chatelier's principle. This later explanation invoked specific mechanisms, referenced chemical theory, and contrasted buffer behavior with non-buffer systems.

Interview data provided insight into how students perceived this evolution. One medium-achieving student reflected that early explanations simply recorded observations. Teacher questioning about the meaning of observations prompted thinking about connections between color changes and chemistry. By the end of intervention, explaining became automatic, with students immediately considering which principle applied. This reflection highlighted the role of teacher scaffolding and repeated practice in shaping explanatory habits. The requirement to explain, reinforced through multiple cycles, shifted student orientation from passive recording to active interpretation. A high-achieving student described developing a mental checklist for explanation construction. This process involved identifying discrepancies, determining which chemical principle applied, and tracing causal chains from principle to observation. When explanation failed, it revealed gaps in understanding, making the explanation process a tool for diagnosing conceptual limits.

Response sheet analysis revealed qualitative differences in explanation patterns across achievement levels. High achievers frequently invoked multiple principles in their explanations, connected observations to mathematical relationships such as relating pH changes to logarithmic scale, and explicitly addressed why alternative explanations were inadequate. Medium achievers typically provided correct explanations focused on a single principle. Low achievers' explanations, though improving over time, often remained focused on observable features with limited mechanistic depth. These patterns aligned with quantitative findings showing differential gains across achievement levels. The progression from descriptive to mechanistic reasoning thus represents a key mechanism through which POE activities enhanced students' critical thinking about acid-base chemistry.

Theme 3: Perceived Transfer of POE-Learned Reasoning to Practical Problems

The third theme addressed Objective 4 by revealing how POE engagement facilitated transfer to practical problem-solving. Interview data demonstrated that students connected their POE experiences directly to solving authentic acid-base chemistry problems. Students transferred both specific conceptual knowledge and general reasoning strategies to novel contexts. One medium-achieving student described encountering a titration curve with an equivalence point near pH 9 on the Practical Problem-Solving Test. Before the intervention, the student would have dismissed the graph as incorrect. The memorable prediction error from titrating ammonia with HCl provided a retrieval cue for weak base-strong acid recognition. This account directly linked cognitive conflict during a POE cycle to subsequent problem-solving performance.

A high-achieving student described transferring the POE cognitive routine itself. The student internalized the predict-observe-explain sequence as a mental habit applied to novel problems. This metacognitive strategy, once established, enabled identification of wrong assumptions through systematic comparison of expectations with evidence. A low-achieving student connected POE buffer activities to practical laboratory applications. The experience of observing minimal pH changes and explaining buffer mechanisms prepared the student to test unknown solutions by adding acid and measuring response. POE activities, by simulating investigative processes, helped students develop practical reasoning skills applicable outside academic contexts.

Three of eighteen interviewees expressed uncertainty about transfer. These students reported learning topics better but struggling to apply specific examples to different situations. Analysis revealed alignment between transfer perceptions and PPST scores. The six students most articulately describing transfer processes scored above 35, placing them in the experimental group's top third. Two of three students expressing transfer uncertainty scored below 25, placing them in the bottom third. This pattern, while not generalizable given sample size, aligned with quantitative findings of strong correlation between POE performance and PPST scores.

Integration of Qualitative and Quantitative Findings

The qualitative themes provide explanatory depth for the quantitative patterns reported earlier. The centrality of cognitive conflict in students' accounts aligns with the quantitative finding that analysis—the component most directly engaged when reconciling prediction with observation—showed the largest intervention effect. Students' descriptions of moving from descriptive to mechanistic explanations illuminate the processes underlying gains in evaluation, which required judging the adequacy of explanations. The perceived transfer reported by students corresponds to the strong partial correlation between POE performance and PPST scores after controlling for pretest ability.

Together, these qualitative findings suggest that POE activities work not merely by providing more engaging instruction, but by systematically engaging students in cognitive processes—confronting discrepancies, constructing mechanistic explanations, applying reasoning to new contexts—that are fundamental to critical thinking development. The voices of students themselves, captured in their response sheets and interviews, reveal that the structured sequence of prediction, observation, and explanation creates conditions for meaningful conceptual change and analytical skill development in acid-base chemistry.

Discussion

The findings provide compelling evidence that Predict-Observe-Explain activities enhance critical thinking among technical college students studying acid-base chemistry. The demonstration that POE instruction produced substantially larger gains compared to conventional methods aligns with research supporting inquiry-based approaches in chemistry education. The effect sizes observed, particularly the large interaction effect with partial eta squared of .334 and the substantial group difference at posttest of 5.10 points, are consistent with recent findings from Indonesian secondary schools where POE implementation in acid-base topics yielded significant improvements in learning outcomes and student activeness (Lia *et al.,* 2025). The consistency of these effects across different educational contexts suggests that POE effectiveness may transcend cultural and curricular boundaries. The differential effects across critical thinking components, with analysis showing the largest intervention effect at partial eta squared of .375, followed closely by evaluation at partial eta squared of .362 and inference at partial eta squared of .358, provide novel insights into the mechanisms through which POE operates. This pattern suggests that prediction-observation discrepancy most directly engages students' analytical reasoning, forcing them to dissect relationships between variables and identify sources of conflict between mental models and empirical evidence. The cognitive conflict process model offers a useful explanatory framework, as anomaly-induced cognitive conflict proves most productive when accompanied by supportive processing activities and when conceptual alternatives remain available to learners. The structured explanation phase following observation provided this supportive processing, while peer discussion and teacher guidance offered conceptual alternatives for students to consider.

The finding that evaluation and inference also showed substantial improvements, though slightly less pronounced than analysis, indicates that POE benefits extend beyond immediate conflict resolution to broader critical thinking capacities. Students who repeatedly engaged in explaining discrepancies developed stronger abilities to assess claim credibility and draw logical conclusions from evidence. This finding resonates with research on the Predict-Explain-Observe-Explain model, which enhances students' learning outcomes in Nigerian secondary schools when implemented with collaborative elements (Ajayi, 2025). The Collaborative-Predict-Explain-Observe-Explain instructional package similarly emphasized social knowledge construction through team-based problem-solving, paralleling the small-group explanation phase in the current study. The significant group by achievement level interaction with partial eta squared of .055 revealed that low achievers derived the greatest relative benefit from POE activities, gaining 5.67 points more than control group counterparts compared to 3.65 points for high achievers. This pattern indicates POE may function as an equalizing pedagogical strategy, helping students with weaker initial understanding catch up to peers. Qualitative data illuminate this mechanism, as low achievers described cognitive conflict forcing them to question assumptions previously accepted without examination. For these students, the discrepancy between prediction and observation exposed understanding gaps that might otherwise have remained hidden. This finding carries important implications for equity in chemistry education, as technical colleges in Nigeria enroll students with diverse educational backgrounds. Instructional strategies that disproportionately benefit already-advantaged learners risk widening achievement gaps, while strategies providing greater support to struggling learners promote more equitable outcomes. The present findings suggest POE falls into the latter category, supported by recent research from Indonesia indicating that the POE2WE model incorporating prediction, observation, explanation, elaboration, writing, and evaluation particularly benefited students with lower critical thinking skills (Haryanti *et al.,* 2024).

The strong correlation observed between POE performance and practical problem-solving scores at r equals .68, which remained substantial even after controlling for pretest ability with partial r equals .59, addresses a central concern in chemistry education research regarding transfer of classroom learning to authentic problem contexts. Students who engaged more deeply with POE activities, making thoughtful predictions, attending carefully to observations, and constructing reasoned explanations, were precisely those who later demonstrated stronger performance on tasks requiring titration interpretation, buffer analysis, and salt hydrolysis reasoning. This finding extends previous research on technology-supported constructivist models in Nigerian chemistry education, which demonstrated significant improvements in pre-service teachers' teaching quality and students' learning outcomes (Ajayi *et al.,* 2025). The present study adds to this literature by documenting a direct relationship between engagement quality with inquiry activities and subsequent problem-solving competence. The partial correlation findings prove particularly instructive, as POE performance uniquely accounted for approximately 35 percent of variance in Practical Problem-Solving Test scores after removing variance attributable to pretest scores. This indicates the POE process itself contributed to practical problem-solving development rather than merely reflecting pre-existing differences among students. This finding supports the theoretical claim that critical thinking is not merely a general aptitude that students either possess or lack, but rather a teachable competency cultivated through appropriately structured instructional experiences. Qualitative data reinforce this interpretation, as students described internalizing the POE cognitive routine and applying it to novel problem contexts. The progression from descriptive accounts to mechanistic reasoning observed across the six-week intervention demonstrates that repeated engagement with explanation requirements gradually shapes students' explanatory habits.

Conclusion

The study demonstrates that Predict-Observe-Explain activities significantly enhance critical thinking in acid-base chemistry. The experimental group outperformed the control group by more than five points at posttest, with analysis showing the largest gains. Low achievers derived greatest benefit, gaining nearly twice as much as control group counterparts. POE engagement strongly predicted practical problem-solving performance, with a correlation of $r = .68$ that remained substantial after controlling for initial ability. These findings confirm that structured prediction-observation sequences, when combined with guided explanation, cultivate analytical competence. The approach proves particularly valuable for students with weaker initial understanding and requires no expensive technology, making it accessible for resource-constrained educational contexts.

Reference

Ajayi, V. O. (2025). Effect of Collaborative-Predict-Explain-Observe-Explain (CPEOE) instructional package on students' learning outcomes in chemistry in secondary schools in Benue State, Nigeria. *International Journal of Chemistry Education Research, 9*(1), 45-62.

Ajayi, V. O., Ameh, P. O., & Negedu, S. A. (2025). Technology-supported constructivist model for pre-service chemistry teachers' teaching quality and students' learning outcomes in Nigerian secondary schools. *Journal of Science Teacher Education, 36*(2), 112-134.

American Chemical Society. (2018). *ACS guidelines for assessing analytical reasoning in chemistry education*. American Chemical Society, Division of Chemical Education.

Bennett, J., & Kennedy, D. (2021). Assessment of practical skills in chemistry: A framework for authentic evaluation. *Studies in Educational Evaluation, 68*, 100-118.

Cetin-Dindar, A., Kala, N., Yaman, F., & Ayas, A. (2022). The effect of POE activities on students' conceptual understanding and retention in acid-base chemistry. *Chemistry Education Research and Practice, 23*(3), 567-582.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction—The Delphi report*. The California Academic Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), 117-124.

George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference* (16th ed.). Routledge.

Haryanti, A., Suwono, H., & Ibrahim, I. (2024). The POE2WE learning model: Enhancing critical thinking skills in chemistry for Indonesian secondary students. *Journal of Turkish Science Education, 21*(1), 78-95.

Kala, N., Yaman, F., & Ayas, A. (2013). The effectiveness of Predict-Observe-Explain technique in probing students' understanding about acid-base chemistry: A case for the concept of pH and pOH. *Chemistry Education Research and Practice, 14*(4), 575-586.

Karpudewan, M., Roth, W. M., & Abdullah, M. N. S. (2020). Enhancing primary school students' critical thinking skills through inquiry-based chemistry learning. *Research in Science Education, 50*(5), 1821-1841.

Lia, R. M., Siahaan, P., & Haris, V. (2025). Implementation of Predict-Observe-Explain (POE) learning model to improve students' learning outcomes and activeness in acid-base titration material. *Jurnal Pijar MIPA, 20*(1), 112-118.

Okonkwo, C. A., & Adebayo, S. A. (2023). Critical thinking deficits in first-year chemistry students across sub-Saharan Africa: A diagnostic assessment. *African Journal of Chemical Education, 13*(2), 45-62.

Olateju, T. T., Salami, I. A., & Ogunleye, A. O. (2023). Instructional practices in Nigerian technical colleges: A survey of chemistry teachers' pedagogical approaches. *Nigerian Journal of Science Education, 24*(1), 78-94.

Rahmawati, Y., Ridwan, A., & Faustine, S. (2021). Developing students' critical thinking skills through integration of socio-scientific issues in POE-based chemistry learning. *Journal of Physics: Conference Series, 1836*(1), 012-028.

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students* (8th ed.). Pearson.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Slisko, J., & Cruz, A. C. (2013). Helping students to recognize and correct their wrong ideas: The role of POE activities in developing assumption-awareness. *European Journal of Physics Education, 4*(3), 35-48.

Treagust, D. F., Won, M., & Duit, R. (2021). POE in chemistry education: A review of research on conceptual change through prediction-observation sequences. *International Journal of Science Education, 43*(8), 1245-1268.

White, R., & Gunstone, R. (1992). *Probing understanding*. Falmer Press.

Yaman, F., & Ayas, A. (2022). Investigating the quality of students' explanations during POE activities in acid-base chemistry. *Research in Science & Technological Education, 40*(2), 189-208.

Yaman, F., Ayas, A., & Çalık, M. (2019). Facilitating conceptual change in acid-base chemistry through POE activities supported by animated movies. *Chemistry Education Research and Practice, 20*(3), 589-603.

Zahronada, A., Cahyono, E., & Sumarti, S. S. (2025). The effectiveness of POE learning model with hands-on activities on students' critical thinking skills in chemistry. *International Journal of Instruction, 18*(1), 221-238.