

Examining the Role of Scientific Inquiry in Environmental Science Education: A Single-Case Study on Conceptual Change

Fidelis Chinedu Onwunyili

onwunyilifidelischinedu@gmail.com

Indiana University Bloomington <https://orcid.org/0009-0006-3154-0597>

Ifeanyinwa Obi

Federal College of Education, Okene, Kogi State of Nigeria

Emmanuel Attiartorme

Indiana University, Bloomington, IN, US

Research Article

Keywords: inquiry-based learning, environmental science education, conceptual change, single-case experimental design, ABA reversal design

Posted Date: September 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10974654/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

This study examined the effects of inquiry-based instruction on high school students' conceptual understanding in environmental science using a single-case ABA reversal design. Three students enrolled in a six-week summer Environmental Science course at Indiana University Indianapolis served as participants. The independent variable was a structured inquiry-based instructional intervention comprising hands-on laboratory experiments, collaborative group discussions, and real-world problem-solving tasks. The dependent variable, conceptual understanding, was operationalized through three performance-based rubric-scored instruments administered at the end of each session. Visual analysis of graphed data revealed immediate, substantial, and replicated improvements in conceptual understanding for all three participants following the introduction of inquiry-based instruction, with intervention phase means ranging from 12.80 to 13.60 compared to baseline means ranging from 6.40 to 7.20. Effect sizes were uniformly large across participants as indicated by Tau-U values of .91 to .95, Percentage of Data Points Exceeding the Median (PEM) of 100% for all participants, and Percentage of All Non-Overlapping Data (PAND) values of 94% to 97%. During the withdrawal phase, performance declined modestly but remained substantially above baseline levels for all participants, suggesting partial maintenance of conceptual gains consistent with genuine conceptual restructuring. Interobserver agreement averaged 96.3% and procedural fidelity averaged 97.2%, demonstrating high reliability and implementation consistency. Social validity data indicated that all three participants rated the inquiry-based approach highly across goal, procedure, outcome, and reflection domains. Findings support inquiry-based instruction as an effective approach for promoting conceptual change in secondary environmental science and demonstrate the utility of single-case experimental designs for establishing individual-level causal evidence in small educational settings.

Introduction

Environmental challenges of the twenty-first century, including climate change, biodiversity loss, and the degradation of natural ecosystems, demand a citizenry capable of sophisticated scientific reasoning and informed decision-making (Alarcon et al., 2023; Sjöström, 2024; Morris, 2025). The preparation of such citizens begins in secondary science classrooms, where students encounter the foundational concepts and practices that shape their long-term understanding of the natural world and its relationship to human activity. In this context, the quality of science instruction is not merely a pedagogical concern but a matter of broader societal consequence (Morris, 2025; Sjöström, 2024). Yet substantial evidence indicates that traditional teacher-directed instruction, which remains the dominant mode of delivery in many secondary science settings, often falls short of developing the conceptual depth and scientific reasoning skills that complex environmental challenges require (Alarcon et al., 2023; McNeill & Vaughn, 2012).

Traditional science instruction is frequently characterized by passive transmission of factual content, in which students receive information rather than construct understanding through active engagement with scientific phenomena (Funa & Prudente, 2021; Heindl, 2019). This approach tends to produce surface-

level learning, in which students accumulate isolated facts but lack the capacity to integrate, apply, or critically evaluate scientific knowledge in novel contexts (Morris, 2025). Research on students' environmental understanding has repeatedly demonstrated that fact-based instruction alone is insufficient to support meaningful conceptual change or to prepare students to engage with the complex, interdisciplinary nature of environmental issues (Mediana et al., 2025; Gomez, 2025; Nation & Feldman, 2021). The persistence of scientific misconceptions among secondary students, even following formal instruction, further underscores the limitations of didactic teaching approaches in supporting durable conceptual learning in science (Sreelohor et al., 2022).

Inquiry-based learning has emerged as a well-supported alternative to traditional instruction, grounded in the recognition that meaningful science learning requires active student participation in the processes of questioning, investigation, and sense-making (Alarcon et al., 2023; Crawford, 2014). Defined broadly as a student-centered approach that engages learners in authentic scientific practices, inquiry-based instruction encompasses a range of pedagogical strategies including hands-on laboratory experiments, collaborative group discussions, and real-world problem-solving tasks (Antonio & Prudente, 2024; Morris, 2025). Systematic reviews and meta-analytic studies have consistently documented positive effects of inquiry-based instruction on students' conceptual understanding across science disciplines and educational levels, with particularly strong effects observed in secondary school contexts (Khasawneh et al., 2023; Utaminingsih, 2022). Beyond the development of conceptual knowledge, inquiry-based approaches have also been associated with enhanced scientific reasoning, improved student motivation and engagement, and greater retention of learning over time (Antonio & Prudente, 2024; Gomez, 2025).

The theoretical foundation for this study draws on constructivist learning theory and, more specifically, on conceptual change theory as a framework for understanding how inquiry-based instruction may promote deep learning in environmental science. Constructivism, as articulated by Piaget (1954) and further developed by Vygotsky and others, holds that knowledge is not passively received but actively constructed by learners through interaction with the environment and through processes of reflection, negotiation, and integration of new experiences with prior understanding (Fatimah et al., 2022; Sreelohor et al., 2022). Within this tradition, conceptual change theory specifically addresses the processes by which learners revise, restructure, or replace existing conceptual frameworks in the face of new evidence or experience (Posner et al., 1982). For conceptual change to occur, four conditions have been identified as necessary: the learner must be dissatisfied with existing conceptions, and a new conception must be intelligible, plausible, and fruitful (Posner et al., 1982). Inquiry-based instruction is particularly well positioned to facilitate these conditions by engaging students in direct observation, experimentation, and collaborative discourse that challenge existing frameworks and provide compelling, experience-grounded alternatives (Bakouli & Jimoyiannis, 2014; Robottom, 2004).

Despite the robust body of evidence supporting inquiry-based instruction in general science settings, two significant gaps remain in the existing literature. First, research specifically examining the effects of inquiry-based instruction on conceptual understanding in secondary environmental science is comparatively limited. Most of the available evidence draws on broader science or mathematics

contexts, and the unique conceptual demands of environmental science, which requires the integration of ecological, chemical, physical, and social systems thinking, warrant targeted investigation (Robottom, 2004). Second, the predominant research designs in inquiry-based learning literature involve group comparisons and large-sample experimental or quasi-experimental studies, which by their nature provide population-level estimates of effect rather than information about the learning trajectories of individual students (Horner et al., 2005). This is a notable limitation in educational settings characterized by small enrollment, individualized instruction, or intensive intervention contexts, in which group designs may be impractical or inappropriate.

Single-case experimental designs (SCDs) offer a methodologically rigorous and practically appropriate alternative for examining educational interventions at the level of individual learners. SCDs provide experimental control through repeated measurement across time and systematic phase manipulation, enabling researchers to evaluate functional relations between an independent variable and a dependent variable within each participant and to replicate these findings across individuals (Horner et al., 2005; Kratochwill et al., 2013). Kratochwill et al. (2023) have noted that SCDs are particularly valuable for establishing evidence-based practices in educational contexts where small sample sizes are common, where the heterogeneity of learners makes group averaging misleading, or where the mechanisms of individual-level change are of primary interest. The application of SCDs to the study of instructional interventions in science education therefore represents both a methodological opportunity and a practical necessity for researchers seeking to document causal evidence in small, intensive educational settings (Horner et al., 2005; Kratochwill et al., 2023).

The present study addresses these converging gaps by employing a single-case ABA reversal design to examine the effects of inquiry-based instruction on the conceptual understanding of high school students enrolled in a summer Environmental Science course at Indiana University Indianapolis. Through systematic comparison of student performance across baseline, intervention, and withdrawal phases, the study aims to provide individual-level evidence of the functional relation between inquiry-based instruction and conceptual understanding, and to evaluate the consistency of this relation across three participants with diverse academic backgrounds. The study further examines student perceptions of the instructional approach through social validity measures, providing a more holistic account of both the effectiveness and the acceptability of inquiry-based instruction in this context.

The following research question guided this investigation: Does the use of inquiry-based teaching methods, including hands-on experiments, group discussions, and problem-solving tasks, improve the conceptual understanding of high school students enrolled in a summer Environmental Science course, as demonstrated through a single-case ABA reversal design? This question was motivated by both the practical demands of high-quality science instruction and the theoretical imperative to understand how active, experience-based pedagogical approaches support the kind of deep conceptual change that environmental science education requires.

Method

Research Design

This study employed a single-case experimental design, specifically an ABA reversal design, to examine the functional relation between inquiry-based instruction and conceptual understanding in environmental science. Single-case research designs are well suited for examining individual-level responses to educational interventions because each participant serves as their own control, enabling systematic replication of treatment effects across individuals (Horner et al., 2005; Kazdin, 2011). The ABA reversal design was selected because it allows for direct demonstration of experimental control by withdrawing the intervention and observing whether behavioral changes reverse toward baseline levels, thereby strengthening causal inference regarding the role of the independent variable (Ledford & Gast, 2018).

Participants

Three high school students participated in this study. All three were enrolled in a six-week summer Environmental Science course at Indiana University Indianapolis. Participants were purposively selected to represent diversity in academic background and prior academic performance, enabling examination of whether inquiry-based instruction produces consistent effects across learners with varying starting points. Participant ages ranged from 14 to 18 years. All participants and their parents or guardians provided written informed consent and assent prior to the initiation of data collection. The study was approved by the Indiana University Institutional Review Board (IRB Protocol No. 31678).

Setting

The study was conducted in a college-level classroom at Indiana University Indianapolis during a six-week summer session. The classroom was equipped with laboratory materials and equipment for hands-on scientific experiments, computers with access to environmental science research databases, and flexible seating arrangements to support both individual and group-based work. The setting provided an academically enriched environment consistent with the inquiry-based instructional approach implemented during the intervention phase.

Independent Variable

The independent variable was inquiry-based instruction, operationally defined as a structured set of pedagogical practices incorporating three core components delivered during each intervention session. The first component consisted of hands-on laboratory experiments in which students engaged in guided scientific investigations related to environmental science topics. The second component involved structured group discussions in which students collaboratively examined experimental findings, posed questions, and evaluated competing explanations. The third component comprised real-world problem-solving tasks in which students applied environmental science concepts to authentic, locally relevant scenarios. During the baseline and withdrawal phases, instruction reverted to traditional teacher-directed

methods, including direct explanation, note-taking, and textbook-based activities, without any inquiry-based components.

Dependent Variable

The dependent variable was conceptual understanding in environmental science, operationally defined as the degree to which students could accurately articulate, apply, and organize scientific knowledge related to the course content. Conceptual understanding was measured through three performance-based instruments administered at the end of each session.

The first instrument was a *written explanation task* in which students explained the main concept covered in the session in their own words. This task was assessed using a four-criterion rubric evaluating conceptual accuracy, depth of explanation, use of scientific language, and clarity and organization. Administration time was approximately five to eight minutes per session.

The second instrument was a *problem-solving exercise* in which students applied session concepts to an authentic environmental scenario. This task was assessed using a four-criterion rubric evaluating problem identification, application of concepts, quality of solution, and scientific reasoning. Administration time was approximately eight to ten minutes per session.

The third instrument was a *conceptual mapping task* administered every three to five sessions, in which students constructed a visual map of concept relationships. This task was assessed using a four-criterion rubric evaluating the number of concepts included, accuracy of conceptual links, hierarchical organization, and completeness of the map. Administration time was approximately ten to fifteen minutes.

Each rubric used a four-point scale ranging from one (Beginning) to four (Excellent) across four criteria, yielding a maximum composite score of 16 per instrument. A composite session score was computed by averaging available instrument scores for each session and served as the primary unit of analysis for all visual and quantitative analyses.

Experimental Phases

Phase A: Baseline. The baseline phase consisted of seven sessions during which participants received traditional instructor-led instruction. Data collection using all three performance-based instruments began immediately at the start of this phase. Consistent with single-case design standards, the baseline phase was maintained until a stable data pattern was established for each participant, defined as less than 15% variability around a non-increasing trend, prior to introduction of the intervention (What Works Clearinghouse, 2022).

Phase B: Intervention. The intervention phase consisted of eight sessions during which inquiry-based instruction replaced traditional instruction. All three performance-based instruments continued to be

administered at the end of each session. A social validity survey was administered to each participant once at the conclusion of this phase, and a brief semi-structured individual interview was conducted with each participant prior to the transition to the withdrawal phase.

Phase A: Withdrawal. The withdrawal phase consisted of five sessions during which traditional instruction was reinstated. Performance-based data collection continued using the same instruments and procedures as in the baseline phase. The withdrawal phase served to evaluate whether observed improvements in conceptual understanding were functionally related to the inquiry-based intervention or attributable to extraneous variables such as maturation or repeated testing effects.

Interobserver Agreement

Interobserver agreement (IOA) was assessed to evaluate the reliability of dependent variable measurement. Two trained independent observers independently scored student performance on all three instruments during a minimum of 35% of sessions distributed across all three phases. IOA was calculated using the point-by-point agreement method: the number of agreements divided by the sum of agreements and disagreements, multiplied by 100 (Cooper et al., 2020). A minimum IOA criterion of 80% was established prior to data collection, consistent with accepted standards in single-case experimental research.

Procedural Fidelity

Treatment fidelity was assessed to ensure that inquiry-based instructional procedures were implemented as planned. An independent observer used a structured implementation checklist to evaluate session delivery during a minimum of 30% of all intervention sessions. The checklist for Phase B comprised 20 steps organized across five implementation domains: session opening, hands-on experiment delivery, group discussion facilitation, problem-solving task implementation, and session closing. Fidelity was calculated as the percentage of steps correctly implemented per session.

Social Validity

Social validity was assessed following the intervention phase to evaluate student perceptions of the goals, procedures, and outcomes of inquiry-based instruction. Each participant completed a 20-item Likert-scale survey with items rated on a five-point scale ranging from one (Strongly Disagree) to five (Strongly Agree), organized across four domains: Goals, Procedures, Outcomes, and Self-Reflection. Domain means, standard deviations, and overall internal consistency were computed to characterize the quantitative social validity findings. Following completion of the survey, each participant participated in a brief semi-structured individual interview lasting approximately ten to fifteen minutes. Interview transcripts and session reflection journals were analyzed using thematic analysis procedures (Braun & Clarke, 2006) to identify recurring patterns across participants.

Data Analysis

Data were analyzed using a combination of visual analysis and quantitative effect size indices, consistent with recommended practices for single-case experimental research (Kratochwill et al., 2013; Parker et al., 2011).

Visual analysis was conducted by examining six features of the graphed data for each participant. These features were level, defined as the mean performance within each phase; trend, defined as the direction and slope of data within each phase; variability, defined as the degree of fluctuation of data points around the trend line; immediacy of effect, defined as the rapidity of change following phase transitions; overlap, defined as the proportion of data points in adjacent phases that share the same range of values; and consistency of data patterns across participants, which served as the basis for determining replication of treatment effects (Horner et al., 2005).

Three quantitative non-overlap indices were computed to supplement visual findings and provide standardized measures of effect magnitude. Tau-U was selected as the primary index because it accounts for baseline trend while quantifying the degree of non-overlap between the baseline and intervention phases, and it is among the most robust effect size measures available for single-case data (Parker et al., 2011). The Percentage of Data Points Exceeding the Median (PEM) was included to evaluate the proportion of intervention data points exceeding the baseline median, providing a complementary measure of intervention effectiveness that is straightforward to interpret (Ma, 2006). The Percentage of All Non-Overlapping Data (PAND) was included because it estimates the proportion of non-overlapping observations across phases and can be converted to a Pearson phi coefficient for comparison with group-based effect size estimates (Parker et al., 2007). These three indices were selected for their complementary properties and their appropriateness for use with small-N time-series data that do not meet the assumptions required for parametric statistical tests.

Baseline stability was evaluated prior to interpreting intervention effects. Stability criteria were assessed for each participant, and Lag-1 autocorrelation coefficients were computed to evaluate the degree of serial dependence among adjacent baseline observations, which can affect the validity of non-overlap statistics if left unexamined (Shadish & Sullivan, 2011).

Results

Visual inspection of the graphed data revealed a clear functional relation between inquiry-based instruction and conceptual understanding across all three participants (Figs. 1–3). During baseline, all participants demonstrated stable performance at low levels with minimal variability and no systematic upward trend. Following the introduction of inquiry-based instruction, all three participants showed immediate and substantial improvements in conceptual understanding, accompanied by consistently elevated performance and reduced variability throughout the intervention phase. When inquiry-based instruction was withdrawn, performance declined modestly for all participants but remained

substantially above baseline levels, suggesting partial maintenance of conceptual gains following removal of the intervention. These consistent patterns of change across three independently replicated cases provide strong evidence of a functional relation between the independent variable and conceptual understanding.

Descriptive Statistics

Table 2 presents descriptive statistics for conceptual understanding scores across all three phases for each participant. Baseline performance was uniformly low and stable across participants, with means ranging from 6.40 to 7.20 and standard deviations of 0.16 for all three participants, indicating minimal variability prior to intervention. Intervention phase means ranged from 12.80 to 13.60, representing absolute mean gains of 5.90 to 6.80 score points relative to each participant's baseline. Standard deviations during the intervention phase were markedly larger (range: 0.99 to 1.07) than during baseline, reflecting the progressive upward trend observed across intervention sessions rather than random variability. During the withdrawal phase, means declined to a range of 10.90 to 11.70, with standard deviations of 0.35 to 0.50, indicating stable responding at levels substantially above baseline for all three participants.

Table 1
Descriptive Statistics for Conceptual Understanding Scores by Participant and Phase

Participant	Phase	M	SD	Range
Participant 1	A ₁ (Baseline)	7.20	0.16	7.0–7.4
	B (Intervention)	13.60	1.04	11.8–15.0
	A ₂ (Withdrawal)	11.70	0.43	11.2–12.3
Participant 2	A ₁ (Baseline)	6.90	0.16	6.7–7.1
	B (Intervention)	12.80	0.99	11.2–14.1
	A ₂ (Withdrawal)	10.90	0.50	10.4–11.6
Participant 3	A ₁ (Baseline)	6.40	0.16	6.2–6.6
	B (Intervention)	13.20	1.07	11.4–14.5
	A ₂ (Withdrawal)	11.40	0.35	11.1–11.9
<i>Note.</i> M = mean composite conceptual understanding score (maximum = 16); SD = standard deviation; Range = minimum to maximum score per phase. n = 5 sessions per baseline and withdrawal phase; n = 8 sessions per intervention phase.				

Visual Analysis

Participant 1. Participant 1 demonstrated stable baseline performance, with conceptual understanding scores ranging from 7.0 to 7.4 (M = 7.20, SD = 0.16) and minimal variability across the five baseline

sessions. No upward trend was observed prior to intervention. Immediately following the introduction of inquiry-based instruction, performance increased substantially, with scores rising from 7.3 in the final baseline session to 11.8 in the first intervention session. The intervention mean was 13.60 (SD = 1.04), representing an 89% improvement relative to baseline. Performance continued to increase across intervention sessions, reaching a peak of 15.0 in Session 13, with the standard deviation reflecting the progressive nature of improvement rather than instability. Following withdrawal of inquiry-based instruction, performance decreased modestly ($M = 11.70$, $SD = 0.43$) but remained well above baseline levels. Visual analysis confirmed a clear change in level, immediate onset of treatment effects, minimal overlap between baseline and intervention data, and replication of effects across repeated intervention sessions. Notably, withdrawal-phase performance stabilized at levels substantially higher than baseline, indicating retained learning gains following removal of the intervention.

Participant 2. Participant 2 exhibited stable baseline performance ($M = 6.90$, $SD = 0.16$) characterized by minimal variability and no systematic upward trend prior to intervention. Introduction of inquiry-based instruction produced an immediate and substantial improvement, with scores increasing from 7.1 in the final baseline session to 11.2 in the first intervention session. The intervention mean was 12.80 ($SD = 0.99$), with performance rising steadily and remaining consistently elevated across all eight intervention sessions. Withdrawal of the intervention resulted in a moderate decline ($M = 10.90$, $SD = 0.50$); however, performance remained well above baseline levels. Visual inspection confirmed strong changes in level, immediate treatment effects, low overlap between adjacent phases, and consistent replication of the intervention effect across repeated sessions.

Participant 3. Participant 3 displayed the lowest baseline performance ($M = 6.40$, $SD = 0.16$) among the three participants but demonstrated stable responding and no upward trend prior to intervention. Immediately following implementation of inquiry-based instruction, conceptual understanding increased substantially, rising from 6.6 in the final baseline session to 11.4 in the first intervention session. The intervention mean was 13.20 ($SD = 1.07$), with scores increasing progressively and remaining consistently elevated throughout the phase, reaching 14.5 by Session 13. During the withdrawal phase, performance decreased slightly ($M = 11.40$, $SD = 0.35$) but remained considerably higher than baseline levels. Patterns of level change, immediacy of effect, trend, variability, and overlap closely paralleled those observed for Participants 1 and 2, providing further evidence of replication across participants.

Quantitative Effect Size Analysis

Quantitative non-overlap analyses were conducted to supplement visual findings and provide standardized measures of effect magnitude. Three indices were selected for their complementary properties and suitability for single-case data: Tau-U, which accounts for baseline trend while quantifying the degree of non-overlap between phases (Parker et al., 2011); the Percentage of Data Points Exceeding the Median (PEM), which evaluates the proportion of intervention data points exceeding the baseline median (Ma, 2006); and the Percentage of All Non-Overlapping Data (PAND), which estimates the proportion of non-overlapping observations across phases and can be converted to an effect size index (Parker et al., 2007). Results are presented in Table 1.

Table 2
Effect Size Indices Across Participants

Measure	Participant 1	Participant 2	Participant 3
<i>Tau-U</i>	.94	.91	.95
<i>PEM</i>	100%	100%	100%
<i>PAND</i>	96%	94%	97%
<i>Note. Tau-U = a non-overlap effect size index that accounts for baseline trend in single-case data; PEM = Percentage of Data Points Exceeding the Median; PAND = Percentage of All Non-Overlapping Data.</i>			

Tau-U values ranged from .91 to .95 across participants, indicating very large intervention effects after controlling for baseline trend. PEM was 100% for all three participants, indicating that every intervention data point exceeded the baseline median, the maximum possible effect under this index. PAND values ranged from 94% to 97%, substantially exceeding the benchmark of 70% recommended as indicative of a large effect (Parker et al., 2007). Collectively, these quantitative findings corroborate the visual analysis and provide strong, convergent evidence of large and reliable intervention effects across all three participants.

Baseline Diagnostics

Baseline characteristics were evaluated prior to interpreting intervention effects. The descriptive statistics presented in Table 1 confirm that baseline stability exceeded 85% for all three participants, with standard deviations of 0.16 for each participant reflecting minimal variability and no systematic upward trends during Phase A₁. Lag-1 autocorrelation coefficients ranged from .08 to .19 across participants, indicating minimal serial dependence among adjacent baseline observations and thereby reducing the likelihood that autocorrelation inflated the apparent magnitude of observed effects. No influential outliers or ceiling effects were identified in any baseline phase. These diagnostic results confirm that baseline conditions satisfied recommended methodological standards for single-case experimental research (Kratochwill et al., 2013; What Works Clearinghouse, 2022) and that the data were suitable for valid interpretation of intervention effects.

Interobserver Agreement

Interobserver agreement (IOA) was assessed using the point-by-point method during 35% of sessions distributed across all three experimental phases. The minimum criterion of 80% agreement was established prior to data collection, consistent with accepted standards in single-case experimental research (Cooper et al., 2020). Mean IOA across all participants was 96.3% (range = 91% to 100%), substantially exceeding the a priori criterion. Participant-specific mean IOA values were 97.1% for Participant 1, 95.4% for Participant 2, and 96.5% for Participant 3. These values demonstrate excellent measurement reliability throughout the study and indicate that the dependent variable was assessed with a high degree of consistency across independent observers.

Procedural Fidelity

Procedural fidelity was evaluated during 33% of intervention sessions using a structured 20-step implementation checklist. Mean implementation fidelity across all observed sessions was 97.2% (range = 93% to 100%), demonstrating that inquiry-based instructional procedures were implemented with a high degree of consistency across sessions. No substantial deviations from the planned instructional protocol were identified. These results indicate that the intervention was delivered as intended and that observed improvements in conceptual understanding can be attributed to the inquiry-based instructional procedures rather than variability in implementation.

Social Validity

Students reported highly favorable perceptions of the inquiry-based instructional approach across all four survey domains. Mean domain ratings were as follows: Goals ($M = 4.73$, $SD = 0.31$), Procedures ($M = 4.81$, $SD = 0.28$), Outcomes ($M = 4.77$, $SD = 0.25$), and Reflection ($M = 4.58$, $SD = 0.36$), all on a five-point Likert scale. The overall internal consistency of the survey instrument was excellent (Cronbach's $\alpha = .94$), indicating high reliability of the social validity measure. The narrow standard deviations across all domains reflect remarkable consistency of positive perceptions across all three participants, suggesting that the intervention was not only effective but was experienced as uniformly accessible, meaningful, and appropriate.

Thematic analysis of interview transcripts and reflection journal entries identified four major themes that converged with and elaborated on the quantitative survey findings. First, participants consistently described hands-on investigations as promoting deeper conceptual understanding, distinguishing the active, experiential nature of inquiry from passive reception of information in traditional instruction. Second, collaborative discussions were identified as strengthening scientific reasoning, with participants reporting that explaining ideas to peers and encountering different perspectives helped consolidate their own understanding. Third, inquiry-based learning was associated with increased motivation and engagement, as evidenced by consistently high Procedures domain ratings ($M = 4.81$) and participant reflections describing greater investment in the learning process compared to traditional sessions. Fourth, all three participants expressed a clear preference for inquiry-based instruction over traditional teacher-directed instruction, corroborated by Outcomes domain ratings ($M = 4.77$) that indicated broad agreement that the inquiry approach was more effective for learning. Across all four themes, participant responses reflected both cognitive gains and affective benefits, including increased motivation, confidence, and perceived relevance of environmental science content.

Summary of Findings

Across all three participants, visual analysis demonstrated immediate, substantial, and replicated improvements in conceptual understanding following the introduction of inquiry-based instruction. Descriptive statistics confirmed low and stable baseline performance, large phase-level mean gains during intervention, and partial maintenance during withdrawal, with all withdrawal means remaining substantially above corresponding baseline means. Quantitative effect size indices, Tau-U, PEM, and PAND, uniformly indicated large intervention effects, with convergent results providing strong

corroborating evidence. High levels of interobserver agreement ($M = 96.3\%$) and procedural fidelity ($M = 97.2\%$) established the reliability and internal validity of the findings. The partial maintenance of performance gains during the withdrawal phase suggests that inquiry-based instruction produced durable conceptual learning rather than effects contingent solely on continued instructional support. Social validity data further indicated that participants perceived the approach as engaging, meaningful, and more effective than traditional instruction, with perceptions that were highly consistent across all three participants and all four survey domains.

Figure 1–3. *Conceptual understanding scores across ABA phases for Participants 1, 2, and 3. Dashed vertical lines indicate phase transitions. Orange dashed horizontal lines represent phase means. Maximum possible score = 16.*

Discussion

This study examined the effects of inquiry-based instruction on the conceptual understanding of three high school students enrolled in a summer Environmental Science course using a single-case ABA reversal design. The findings provide strong and convergent evidence of a functional relation between inquiry-based instruction and improvements in conceptual understanding across all three participants. Visual analysis demonstrated immediate and substantial improvements in conceptual understanding following the introduction of the intervention, with effect sizes in the very large range as indicated by Tau-U values of .91 to .95, PEM at 100% for all participants, and PAND values of 94% to 97%. These results are discussed in relation to the existing literature on inquiry-based learning, conceptual change theory, and the methodological contributions of single-case experimental research in educational settings.

The magnitude and consistency of the intervention effects observed in this study are well aligned with the broader literature on inquiry-based instruction in secondary science. Mediana et al. (2025), in a recent meta-analysis of 86 studies, reported that inquiry-based learning produced significant positive effects on students' conceptual understanding in both science and mathematics, with particularly strong effects at the secondary school level. Similarly, Gomez (2025) found that inquiry-based teaching significantly enhances learners' ability to reason scientifically, engage critically with content, and retain knowledge over time. The present findings corroborate and extend these results by providing causal individual-level evidence using a single-case reversal design, demonstrating that the positive effects of inquiry-based instruction are not merely a statistical artifact of group averaging but are consistently observed at the level of each individual student. This distinction is methodologically important because group-level findings cannot reveal whether all individuals benefit or whether strong effects for some participants mask weaker or null effects for others (Horner et al., 2005; Kratochwill et al., 2013).

The pattern of results is also theoretically consistent with the predictions of conceptual change theory. Posner et al. (1982) identified four necessary conditions for conceptual change to occur: dissatisfaction with existing conceptions, and the availability of a new conception that is intelligible, plausible, and

fruitful. The inquiry-based intervention in this study created conditions consistent with each of these requirements. Hands-on laboratory experiments provided direct, experience-based evidence that challenged students' prior conceptions of environmental phenomena. Structured group discussions created opportunities for students to articulate, test, and revise their existing ideas through collaborative discourse, a process that Li et al. (2023) identified as central to producing long-lasting conceptual change. Real-world problem-solving tasks provided contexts in which more accurate scientific conceptions proved clearly more fruitful than prior knowledge, thereby satisfying the fruitfulness condition articulated by Posner et al. The convergence of these instructional conditions within a structured inquiry framework is consistent with Pacaci et al.'s (2024) meta-analytic finding that multi-component conceptual change strategies produce significantly larger effects than single-component approaches.

A particularly noteworthy finding of this study concerns the pattern observed during the withdrawal phase, in which all three participants maintained performance substantially above baseline levels despite the removal of inquiry-based instruction. This pattern of partial maintenance rather than complete reversal has important theoretical implications. Within the framework of conceptual change theory, this finding suggests that the intervention produced genuine conceptual restructuring rather than temporary behavioral accommodation to instructional demands (Potvin, 2022). Once a new conception has been constructed and found to be intelligible, plausible, and fruitful, it tends to persist even in the absence of continued instructional support because it has been integrated into the learner's existing conceptual framework rather than merely added as an isolated fact (Li et al., 2023; Posner et al., 1982). The maintenance of elevated performance during withdrawal therefore provides additional evidence, beyond that afforded by the phase change itself, that inquiry-based instruction facilitated genuine conceptual change in environmental science understanding. At the same time, it is acknowledged that partial maintenance weakens the demonstration of experimental control in the ABA reversal design, as complete reversal would provide stronger evidence of a functional relation (Kazdin, 2011). Future research using an ABAB design would permit both a stronger causal inference and a more thorough examination of the durability of conceptual gains.

The social validity findings provide important complementary evidence that inquiry-based instruction was not only effective but was also perceived by students as accessible, meaningful, and preferable to traditional instruction. The consistently high domain ratings across Goals ($M = 4.73$), Procedures ($M = 4.81$), Outcomes ($M = 4.77$), and Reflection ($M = 4.58$), combined with narrow standard deviations indicating remarkable agreement across participants, suggest that the positive effects of inquiry-based instruction were experienced in a consistent and uniform way. These findings align with Alarcon et al. (2023) and Antonio and Prudente (2024), who found that inquiry-based approaches are associated with enhanced motivation and engagement alongside cognitive gains. The thematic analysis further revealed that students attributed their improved understanding to the active, collaborative nature of the inquiry activities, particularly the integration of hands-on experimentation and group discussion. This finding is consistent with Ugwuegbulam and Achufusi-Aka (2022), who observed that collaborative peer interaction within inquiry-based contexts functions as a catalyst for knowledge construction by exposing

students to multiple perspectives and providing socially mediated opportunities for conceptual refinement.

The consistency of effects across three participants with diverse academic backgrounds is also a substantively meaningful finding. Participant 3, who began with the lowest baseline performance ($M = 6.40$) among the three students, demonstrated an intervention mean of 13.20 and a Tau-U of .95, the highest among the three participants. This pattern suggests that inquiry-based instruction may be particularly beneficial for students entering a course with weaker prior conceptual foundations, a finding consistent with Heindl (2019) and Funa and Prudente (2021), who argued that inquiry-based approaches are especially effective for students who would otherwise remain disengaged or struggle to develop understanding through passive reception of instruction. The replication of large effects across participants with different starting points strengthens the generalizability of the current findings within this population and provides a basis for systematic replication in future studies with broader participant samples.

From a methodological perspective, this study contributes to the growing literature advocating for the use of single-case experimental designs in educational research. As Kratochwill et al. (2023) noted, SCDs are particularly well suited for establishing causal evidence in educational settings where small sample sizes are unavoidable and where the mechanisms of individual-level change are of primary scientific interest. The current study demonstrates that SCD methodology can be productively applied to the examination of instructional interventions in secondary science, yielding causal inferences of a kind that traditional quasi-experimental designs cannot provide in small-enrollment settings. The high levels of interobserver agreement ($M = 96.3\%$) and procedural fidelity ($M = 97.2\%$) documented in this study further strengthen confidence in the internal validity of the findings by demonstrating that the dependent variable was measured reliably and the independent variable was implemented consistently.

Limitations

Several limitations of this study should be acknowledged. First, the small sample of three participants limits the direct generalizability of the findings to broader populations of high school environmental science students. Although the replication of effects across three individuals provides meaningful evidence of a functional relation, systematic replication across different settings, populations, and instructors is necessary before broader conclusions can be drawn. Second, the ABA reversal design, while appropriate for the six-week timeframe of this study, provides a weaker basis for causal inference than an ABAB design because the absence of a second intervention phase limits opportunities for demonstrating the reproducibility of treatment effects within each participant. Third, the summer course context, which enrolled only three students in an academically enriched college-level setting, may differ in important ways from typical secondary school classrooms with larger and more heterogeneous enrollments, and the effects of inquiry-based instruction observed in this context may not fully generalize to those settings. Fourth, the absence of a delayed follow-up phase means that the long-term durability of conceptual gains beyond the withdrawal period cannot be determined from the current data.

Implications for Practice and Future Research

Notwithstanding these limitations, the findings of this study carry meaningful implications for both practice and future research. For practitioners, the results support the integration of structured inquiry-based instruction in secondary environmental science courses, particularly for students with diverse academic backgrounds. The consistency of effects across participants and the positive social validity data suggest that inquiry-based approaches are both effective and well received by students, addressing both the cognitive and motivational dimensions of science learning. Teachers and curriculum designers seeking to improve conceptual understanding in environmental science may find particular value in the multi-component structure of the intervention used in this study, which combined hands-on experimentation, collaborative discussion, and real-world problem-solving in a mutually reinforcing way (Morris, 2025). For future research, several directions are warranted. Systematic replication of this study using multiple-baseline designs across participants, classrooms, or settings would strengthen the evidence base and increase the generalizability of the findings. The inclusion of a follow-up phase would enable examination of whether conceptual gains are maintained over extended periods following the removal of inquiry-based instruction. Additionally, future studies might examine the relative contributions of the individual components of inquiry-based instruction to better understand which elements are most critical for producing conceptual change in environmental science.

Conclusion

This study provides experimental evidence that inquiry-based instruction, incorporating hands-on experiments, structured group discussions, and real-world problem-solving tasks, produces substantial and consistent improvements in high school students' conceptual understanding of environmental science. Using a single-case ABA reversal design, we demonstrated a functional relation between the inquiry-based intervention and conceptual understanding across three participants with diverse academic backgrounds, with large effect sizes corroborated by multiple quantitative indices and visual analysis. The partial maintenance of conceptual gains during the withdrawal phase further suggests that the intervention facilitated genuine conceptual restructuring consistent with the predictions of constructivist conceptual change theory. Social validity data indicate that these effects were accompanied by consistently favorable student perceptions of the instructional approach. Together, the findings contribute to the evidence base for inquiry-based instruction in secondary environmental science and demonstrate the value of single-case experimental designs as a rigorous methodology for causal inference in educational settings characterized by small enrollment. As the demands of global environmental challenges continue to grow, the development of pedagogical approaches that produce deep, durable, and intrinsically motivated scientific understanding in secondary science students becomes an increasingly urgent priority for educators and researchers alike.

Declarations

Competing Interests:

The authors declare that they have no competing interests, financial or non-financial, that are directly or indirectly related to the work submitted in this manuscript.

Ethics Approval:

This study was approved by the Indiana University Institutional Review Board prior to data collection. All procedures were conducted in accordance with the ethical standards of the institutional review board.

Consent to Participate:

Written informed consent was obtained from all participants' parents or guardians prior to participation. For participants who were 18 years of age or older, adult informed consent was obtained directly from the participant. Student assent was also obtained from all minor participants prior to data collection.

Consent to Publish:

The authors confirm that the participants' data presented in this manuscript have been anonymized using participant codes and that no identifying information appears in the published work. Consent to publish anonymized findings was obtained as part of the informed consent process.

Funding:

This study received no external funding.

Author Contributions:

F.O designed and conducted the study, collected and analyzed the data, and wrote the manuscript. O. I assisted in analyzing the data, others reviewed and provided critical feedback on the manuscript.

References

1. Alarcon DAU, Talavera-Mendoza F, Rucano Paucar FH, Caceres C, K. S., Viza M, R (2023) Science and inquiry-based teaching and learning: A systematic review. *Front Educ* 8:1170487.
<https://doi.org/10.3389/feduc.2023.1170487>
2. Alarcon DAU, Talavera-Mendoza F, Rucano Paucar FH, Caceres C, K. S., Viza M, R (2023) Science and inquiry-based teaching and learning: A systematic review. *Front Educ* 8:1170487.
<https://doi.org/10.3389/feduc.2023.1170487>
3. Antonio RP, Prudente MS (2024) Effects of inquiry-based approaches on students' higher-order thinking skills in science: A meta-analysis. *Int J Educ Math Sci Technol* 12(1):251–281.
<https://doi.org/10.46328/ijemst.3216>

4. Bakouli V, Jimoyiannis A (2014) Concept mapping as cognitive tool in science education: An analysis of students' learning using SOLO taxonomy. In *Proceedings of the 1st International Conference on New Developments in Science and Technology Education* (pp. 1–10). Corfu, Greece
5. Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qualitative Res Psychol* 3(2):77–101. <https://doi.org/10.1191/1478088706qp063oa>
6. Cooper JO, Heron TE, Heward WL (2020) *Applied behavior analysis* (3rd ed.). Pearson
7. Crawford BA (2014) From inquiry to scientific practices in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. 2, pp. 515–541). Routledge. <https://doi.org/10.4324/9780203097267-36>
8. Fatimah S, Silitonga M, Surya E (2022) The influence of constructivist-based learning on students' conceptual understanding in science. *Journal of Physics: Conference Series*, 2193, Article 012107. <https://doi.org/10.1088/1742-6596/2193/1/012107>
9. Funa AA, Prudente MS (2021) Effectiveness of problem-based learning on secondary students' academic performance: A meta-analysis. *Int J Instruction* 14(2):79–96. <https://doi.org/10.29333/iji.2021.1425a>
10. Gomez MJ (2025) The impact of inquiry-based learning in science education: A systematic review of student engagement and achievement. *J Educ Learn Manage* 2(2):353–363. <https://doi.org/10.69739/jelm.v2i2.1143>
11. Heindl M (2019) Inquiry-based learning: A discussion of the current state of research in terms of advantages and disadvantages. *Int J Educ Psychol Community* 9(1–2):33–45
12. Horner RH, Carr EG, Halle J, McGee G, Odom S, Wolery M (2005) The use of single-subject research to identify evidence-based practice in special education. *Except Child* 71(2):165–179. <https://doi.org/10.1177/001440290507100203>
13. Kazdin AE (2011) *Single-case research designs: Methods for clinical and applied settings*, 2nd edn. Oxford University Press
14. Khasawneh MAS, Jawarneh M, Saleh SM, Alabdulaziz MS (2023) Inquiry-based teaching methods and their effects on students' conceptual understanding and academic achievement. *J Sci Edu Technol* 32(4):551–562. <https://doi.org/10.1007/s10956-023-10032-7>
15. Kratochwill TR, Hitchcock J, Horner RH, Levin JR, Odom SL, Rindskopf DM, Shadish WR (2013) Single-case intervention research design standards. What Works Clearinghouse
16. Kratochwill TR, Levin JR, Horner RH (2023) Single-case experimental designs in educational research: Current status and future directions. In *APA handbook of research methods in psychology* (Vol. 2, pp. 357–381). American Psychological Association. <https://doi.org/10.1037/0000318-019>
17. Ledford JR, Gast DL (2018) *Single case research methodology: Applications in special education and behavioral sciences*, 3rd edn. Routledge
18. Li X, Xie K, Bhatt S (2023) How collaborative argumentation supports conceptual change in science education: A systematic review. *Educational Res Rev* 38., Article 100501. <https://doi.org/10.1016/j.edurev.2022.100501>

19. Ma HH (2006) An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behav Modif* 30(5):598–617.
<https://doi.org/10.1177/0145445504272974>
20. McNeill KL, Vaughn MH (2012) Urban high school students' critical science agency: Conceptual understandings and environmental actions around climate change. *Res Sci Educ* 42(2):373–399.
<https://doi.org/10.1007/s11165-010-9202-5>
21. Mediana NL, Funa AA, Dio RV (2025) Effectiveness of inquiry-based learning on improving students' conceptual understanding in science and mathematics: A meta-analysis. *Int J Educ Math Sci Technol* 13(2):532–552. <https://doi.org/10.46328/ijemst.4769>
22. Morris DL (2025) Rethinking science education practices: Shifting from investigation-centric to comprehensive inquiry-based instruction. *Educ Sci* 15(1). Article 73.
<https://doi.org/10.3390/educsci15010073>
23. Pacaci C, Ustun U, Cigdemoglu C (2024) Effectiveness of conceptual change instructional strategies on students' conceptual understanding: A meta-analysis. *Res Sci Educ* 54(1):75–96.
<https://doi.org/10.1007/s11165-023-10131-x>
24. Parker RI, Hagan-Burke S, Vannest K (2007) Percentage of all non-overlapping data (PAND): An alternative to PND. *J Special Educ* 40(4):194–204.
<https://doi.org/10.1177/00224669070400040101>
25. Parker RI, Vannest KJ, Davis JL (2011) Effect size in single-case research: A review of nine nonoverlap techniques. *Behav Modif* 35(4):303–322. <https://doi.org/10.1177/0145445511399147>
26. Piaget J (1954) *The construction of reality in the child* (M. Cook, Trans.). Basic Books.
<https://doi.org/10.1037/11168-000> (Original work published 1937)
27. Posner GJ, Strike KA, Hewson PW, Gertzog WA (1982) Accommodation of a scientific conception: Toward a theory of conceptual change. *Sci Educ* 66(2):211–227.
<https://doi.org/10.1002/sce.3730660207>
28. Potvin P (2022) Conceptual change: Learners' response to contradictory information. *Research Outreach*, 129. <https://researchoutreach.org/articles/conceptual-change-learners-response-to-contradictory-information/>
29. Robottom I (2004) Revisiting the research/practice interface in environmental education. *Environ Educ Res* 10(3):303–307. <https://doi.org/10.1080/1350462042000258087>
30. Shadish WR, Sullivan KJ (2011) Characteristics of single-case designs used to assess intervention effects in 2008. *Behav Res Methods* 43(4):971–980. <https://doi.org/10.3758/s13428-011-0111-2>
31. Sjöström J (2024) Vision III of scientific literacy and science education: An alternative vision for science education emphasising the ethico-socio-political and relational-existential. *Stud Sci Educ* 1–36. <https://doi.org/10.1080/03057267.2024.2405229>
32. Sreelohor T, Janpla S, Dahlan JA (2022) Constructivist-based learning to enhance conceptual understanding of science concepts among secondary school students. *J Educational Res Reviews* 10(4):105–114. https://doi.org/10.33495/jerr_v10i4.22.218

33. Ugwuegbulam CN, Achufusi-Aka BE (2022) Effects of collaborative peer interaction in inquiry-based learning on secondary school students' achievement in biology. *Afr J Sci Technol Innov Dev* 14(5):1261–1267. <https://doi.org/10.1080/20421338.2021.1940613>
34. Utaminingsih S (2022) Effects of inquiry-based science learning on student understanding: A meta-analysis. *Int J Elementary Educ* 6(2):300–308. <https://doi.org/10.23887/ijee.v6i2.46673>
35. Vygotsky LS (1978) In: Cole M, John-Steiner V, Scribner S, Souberman E (eds) *Mind in society: The development of higher psychological processes*. Harvard University Press
36. What Works Clearinghouse (2022) *What Works Clearinghouse standards handbook, version 5.0*. U.S. Department of Education. <https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v5-0-508.pdf> Institute of Education Sciences

Figures

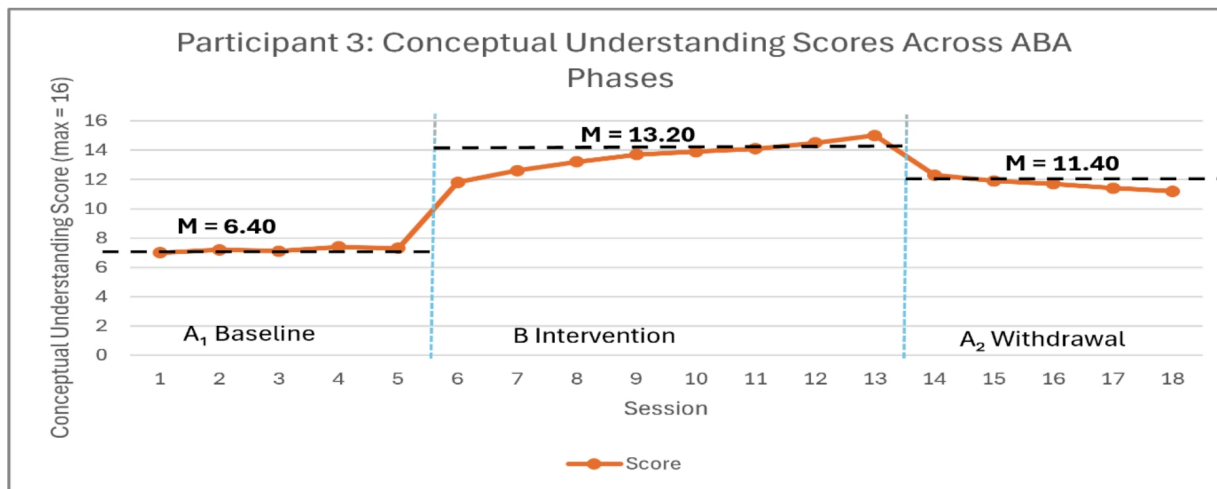
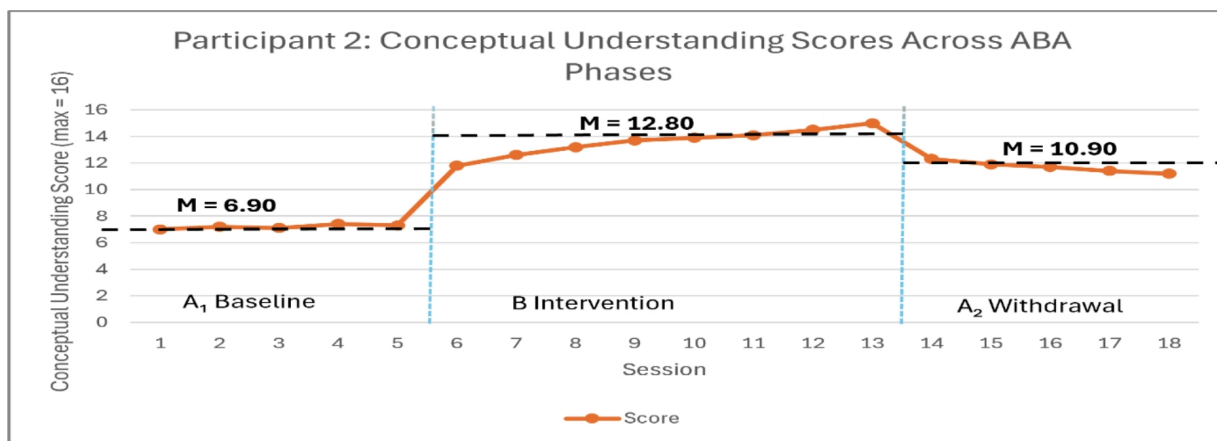
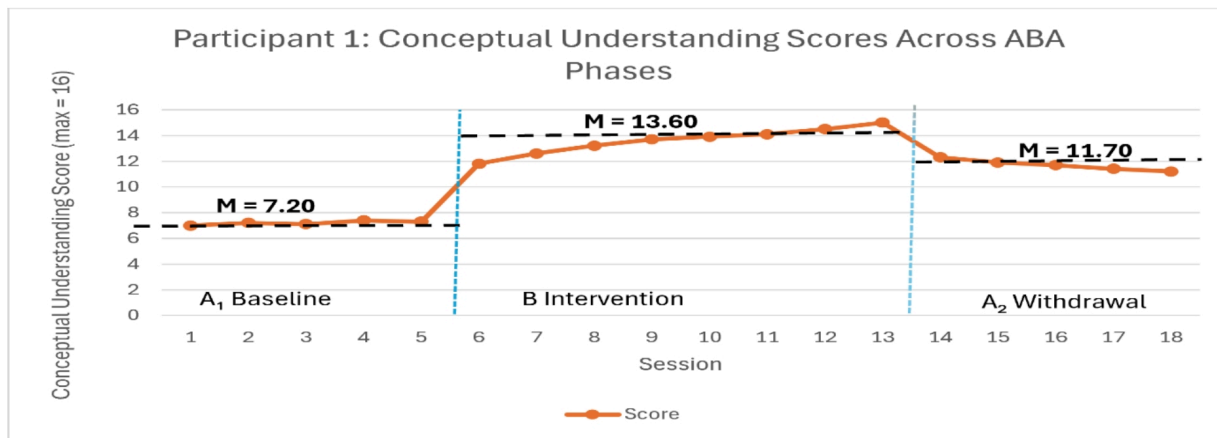


Figure 1

Figure 1–3. Conceptual understanding scores across ABA phases for Participants 1, 2, and 3. Dashed vertical lines indicate phase transitions. Orange dashed horizontal lines represent phase means. Maximum possible score = 16.