ANALISTA: International Journal of Multidisciplinary Research (AIJMR)

Volume 4, Issue 1, June 2026 | ISSN: 3048-6130

EFFECTIVENESS OF THE TECHNOLOGY-AI-INQUIRY-BASED-LINEAR (TAIL) MODEL IN TEACHING EARTH SCIENCE AMONG GRADE 10 STUDENTS

Mark Joseph S. Ariola

Azpetia National High School, Agusan del Sur, Philippines

DOI:10.71182/aijmr.2606.0401.2002

A R T I C L E I N F O

Keywords:
TPCK
Artificial Intelligence in Education
Inquiry-Based Learning
Earth Science Achievement
Quasi-Experimental Design
Secondary Science
ANCOVA

Corresponding Author:
Mark Joseph S. Ariola

A B S T R A C T

Earth Science learning in secondary schools often suffers from abstract content, limited instructional scaffolds, and variable student readiness. Philippine large-scale assessment data also indicate persistent science achievement gaps, supporting the need for stronger classroom-level innovations. This quasi-experimental study used a non-equivalent control group pretest-posttest design with Grade 10 students in one school in Agusan del Sur, Philippines, SY 2025–2026. Two intact classes were assigned to an experimental group (TPCK-informed instruction, AI-supported tools, inquiry cycles, linear instructional sequencing) and a control group (conventional instruction). A 40-item Earth Science Test was administered before and after an 8-week intervention. Baseline equivalence was tested via an independent-samples t-test. Posttest differences were analyzed using ANCOVA with pretest as a covariate. Effect sizes were reported using partial eta squared and Cohen's d. Baseline pretest performance did not significantly differ between groups, $t(84) = 0.48$, $p = .633$, $d = 0.10$. *After controlling for pretest scores, the experimental group significantly outperformed the control group,* $F(1,83) = 34.52, p < .001$, *partial* $\eta^2 = .294$. *Adjusted mean difference was 5.31 points (95% CI [3.43, 7.18]), corresponding to a large effect* ($d = 0.93$). *The TAIL model produced statistically significant and educationally meaningful gains in Earth Science achievement. Integrating teacher knowledge (TPCK), guided AI use, inquiry pedagogy, and coherent sequencing can improve learning outcomes in secondary Earth Science.*

1. INTRODUCTION

Improving student achievement in Earth Science remains a persistent instructional priority, particularly because many core concepts in the discipline—such as geologic time, Earth systems interactions, and spatial–temporal processes. This requires forms of reasoning that learners often find cognitively demanding. These challenges are amplified in classrooms where students possess uneven prerequisite knowledge, limited motivation toward science, or restricted exposure to technology-supported learning environments. At the systems level, international assessment data continue to indicate that Filipino learners perform below the OECD average in science, underscoring the need for instructional innovations that are pedagogically robust yet feasible within real classroom conditions (OECD, 2023).

Contemporary research increasingly suggests that single-method approaches are insufficient for teaching complex scientific concepts. Rather, integrated instructional frameworks that combine multiple evidence-based elements tend to yield stronger learning outcomes. One such framework is Technological Pedagogical Content Knowledge (TPCK), which emphasizes the strategic alignment of content, pedagogy, and technology to support meaningful learning rather than tool-driven instruction (Mishra, 2019). Complementing this, systematic reviews of artificial intelligence (AI) in education indicate that AI can enhance learning when used as a scaffold for feedback, explanation support, and adaptive practice under explicit pedagogical guidance (Zawacki-Richter et al., 2019). Recent policy guidance further stresses that AI integration must remain human-centred, ethically governed, and instructionally purposeful to be educationally beneficial (UNESCO, 2023).

In parallel, inquiry-based learning continues to be recognized as a powerful mechanism for developing conceptual understanding and scientific reasoning, particularly when combined with explicit guidance and structured instructional design rather than implemented as minimally guided discovery (de Jong et al., 2024). This convergence of evidence suggests that effective science instruction may require deliberate synthesis of pedagogical frameworks, digital supports, inquiry processes, and coherent lesson sequencing.

The present study responds to these converging lines of evidence by testing the Technology–AI–Inquiry–Based–Linear (TAIL) model. TAIL is not simply another technology-rich lesson format. Its novelty lies in the deliberate integration of four elements that are often studied separately: (1) TPACK-informed lesson design, (2) teacher-guided use of generative AI for formative support, (3) inquiry tasks that position students as evidence users and explanation builders, and (4) a linear phase structure that manages cognitive load through a predictable progression from orientation to assessment. In contrast with widely used inquiry models such as 5E, which organize learning around engagement and exploration, TAIL gives a more explicit operational role to AI-mediated refinement and teacher verification. Compared with broader blended or comprehensive inquiry frameworks, TAIL more clearly specifies how classroom AI is embedded within teacher-led checkpoints and accuracy validation routines.

1.1 Research Questions

1. Is there a significant difference in pretest Earth Science performance between the experimental and control groups?
2. Is there a significant difference in posttest Earth Science performance between the experimental and control groups after controlling for baseline performance?
3. What is the effect size of the TAIL model on students' Earth Science achievement?

2. LITERATURE REVIEW

2.1 TPACK as an Instructional Design Backbone

The Technological Pedagogical Content Knowledge (TPCK) framework conceptualizes effective technology integration as the dynamic intersection of teachers' content expertise, pedagogical knowledge, and technological competence within authentic instructional contexts. Rather than emphasizing technology as an isolated tool, TPCK foregrounds pedagogically meaningful integration aligned with disciplinary goals and learner needs (Mishra, 2019). Recent empirical syntheses indicate that professional development and instructional interventions grounded in TPCK can positively influence instructional planning, classroom implementation, and student learning outcomes, particularly when technology use is intentional, curriculum-aligned, and cognitively purposeful (Ning et al., 2022; Wilson et al., 2020). These findings suggest that TPCK functions not merely as a conceptual model but as an actionable design framework capable of guiding complex instructional innovations in science education.

2.2 AI-Supported Learning in Science Classrooms

Artificial intelligence (AI) has emerged as a promising educational support mechanism, particularly for formative feedback, adaptive learning, and personalized scaffolding. Systematic reviews consistently report that AI applications can enhance learning efficiency, support differentiated instruction, and provide real-time diagnostic feedback when integrated within pedagogically sound environments (Zawacki-Richter et al., 2019; Wang et al., 2024). However, the effectiveness of AI-supported learning depends strongly on instructional alignment, ethical governance, and active teacher mediation. International policy guidance stresses that AI implementation in schools must remain human-centered, transparent, and developmentally appropriate, ensuring that technological affordances enhance rather than replace pedagogical decision-making (UNESCO, 2023). Emerging studies also indicate that AI-supported professional learning environments can strengthen teachers' technological–pedagogical competencies, thereby reinforcing TPCK development in authentic classroom contexts (Xie et al., 2024).

Moreover, recent research shows both promise and caution regarding AI use in science learning. A systematic review of several empirical studies in science education concluded that AI tools can support engagement, personalization, assessment, and feedback, but that impact depends on pedagogical alignment and responsible teacher mediation (Almasri, 2024). In physics-related contexts, ChatGPT and similar systems have been found capable of generating explanations, solving some structured problems, and supporting question practice; however, studies also show accuracy limitations and a tendency for learners to over-trust incorrect answers, highlighting the need for explicit verification routines (Ding et al., 2023; Liang et al., 2023; Tong et al., 2024).

Emerging work is especially relevant for secondary science inquiry. Min et al. (2025), for example, analyzed how middle school students used a generative AI assistant while designing open-ended physics experiments. Students used AI to ask about variables, materials, procedures, and analysis, and their interaction patterns varied according to topic familiarity and their perceptions of AI. These findings suggest that AI can function productively as a formative partner in inquiry when learners are given task boundaries and when teacher guidance remains visible. Collectively, these studies support the use of AI not as a substitute for disciplinary teaching but as a structured aid for questioning, explanation, checking, and feedback.

2.3 Instructional Sequencing and Blended Inquiry Frameworks

Current evidence rejects a simplistic "inquiry vs direct instruction" dichotomy and supports well-designed blends of inquiry with explicit guidance depending on learning goals and learner readiness. This has direct relevance for Earth Science, where abstract concepts often require both student investigation and explicit conceptual consolidation.

Also, some current studies increasingly reject a simplistic "inquiry versus direct instruction" dichotomy. De Jong et al. (2024) argue that impactful science learning opportunities are best designed through context-sensitive combinations of inquiry and explicit support. This position is reinforced by Zhang and Sweller (2024), who found that the effectiveness of sequencing investigations and explicit instruction depends on learners' prior knowledge and task complexity. For less knowledgeable learners dealing with high element interactivity, explicit instruction before investigations may be advantageous; for more knowledgeable learners, investigation-first sequences may become productive. Such evidence provides a strong rationale for TAIL's phase-based sequencing and its emphasis on teacher-guided transitions.

Broader integrated inquiry frameworks offer useful comparison points for justifying TAIL's novelty. Morris (2025) proposed the Comprehensive Inquiry-Based Science Education (CIBSE) framework to move inquiry beyond investigation-centric teaching by balancing exploration, explanation, argumentation, and reflective support. Similarly, blended learning research in Philippine K–12 settings suggests that well-designed technology-supported learning communities can strengthen interaction and flexible access when pedagogy remains purposeful. TAIL builds on these insights but adds a distinct feature set: the planned use of generative AI within teacher-controlled checkpoints, a verification routine for scientific accuracy, and an explicitly linear structure intended to reduce fragmentation during conceptually demanding Earth Science lessons.

2.4 Earth Science Instructional Context

Earth Science education presents distinctive pedagogical demands that differentiate it from other science domains. Many topics require students to reason across vast temporal scales, visualize subsurface processes, and understand interactions among Earth systems, all of which impose substantial cognitive load. Research indicates that these characteristics often contribute to persistent learning difficulties and misconceptions, especially when instruction relies heavily on abstract verbal explanations alone (Chakour et al., 2019). Consequently, Earth Science is particularly well-suited to multimodal instructional approaches that integrate visualizations, guided inquiry tasks, and digitally mediated feedback to support conceptual clarity and systems thinking.

Further, conceptual understanding in Earth Science requires more than the recall of facts (Scherer et al., 2017; Ribeiro & Orion, 2021). Students must reason with representations, explain dynamic systems, connect evidence to models, and imagine processes unfolding across scales of time and space (Scherer et al., 2017; Ribeiro & Orion, 2021). Topics such as plate tectonics are especially demanding because they depend on spatial reasoning, systems thinking, and interpretation of diagrams, maps, cross-sections, and simulations (McLaughlin & Bailey, 2023; Epler-Ruths et al., 2020). Research in geoscience education has repeatedly emphasized the importance of explicit support for spatial thinking in helping students understand Earth processes (McLaughlin & Bailey, 2023; Taylor et al., 2024).

Several recent reviews show that geoscience education has not consistently provided sufficient opportunities for students to develop spatial skills, despite the importance of these skills for understanding Earth systems (McLaughlin & Bailey, 2023). Coursework, mapping, modeling, sketching, gesture, and computer-based models have all been identified as useful supports for strengthening spatial reasoning in geoscience learning (McLaughlin & Bailey, 2023). Broader reviews in cognitive and STEM research likewise argue that visualizations are central to STEM learning because they connect spatial thinking to abstract disciplinary concepts (Taylor et al., 2024). When students engage with well-designed visual representations, they are better able to organize information, imagine transformations, and build explanatory models (Taylor et al., 2024; Epler-Ruths et al., 2020).

Earth Science learning also benefits from dynamic visualization tools. Three-dimensional and interactive geovisualization can help learners explore the spatial structure of Earth systems and understand processes that static images cannot easily communicate (Fitzpatrick & Hedley, 2024; Ribeiro & Orion, 2021). Reviews in geovisualization highlight the growing role of 3D and 4D tools in geosciences, emphasizing their potential for data exploration, interpretation, and communication (Fitzpatrick & Hedley, 2024). For school-level instruction, AI-supported simulations and visualizations may therefore help reduce cognitive barriers by making plate movement, boundary interactions, and Earth processes more visible and manipulable (Epler-Ruths et al., 2020; Fitzpatrick & Hedley, 2024)

2.5 Conceptual Framework

The TAIL model synthesizes four complementary components: TPACK-informed lesson design, AI-supported formative scaffolding, inquiry-based learning, and linear instructional sequencing. Figure 1 presents the operational flow of the model.

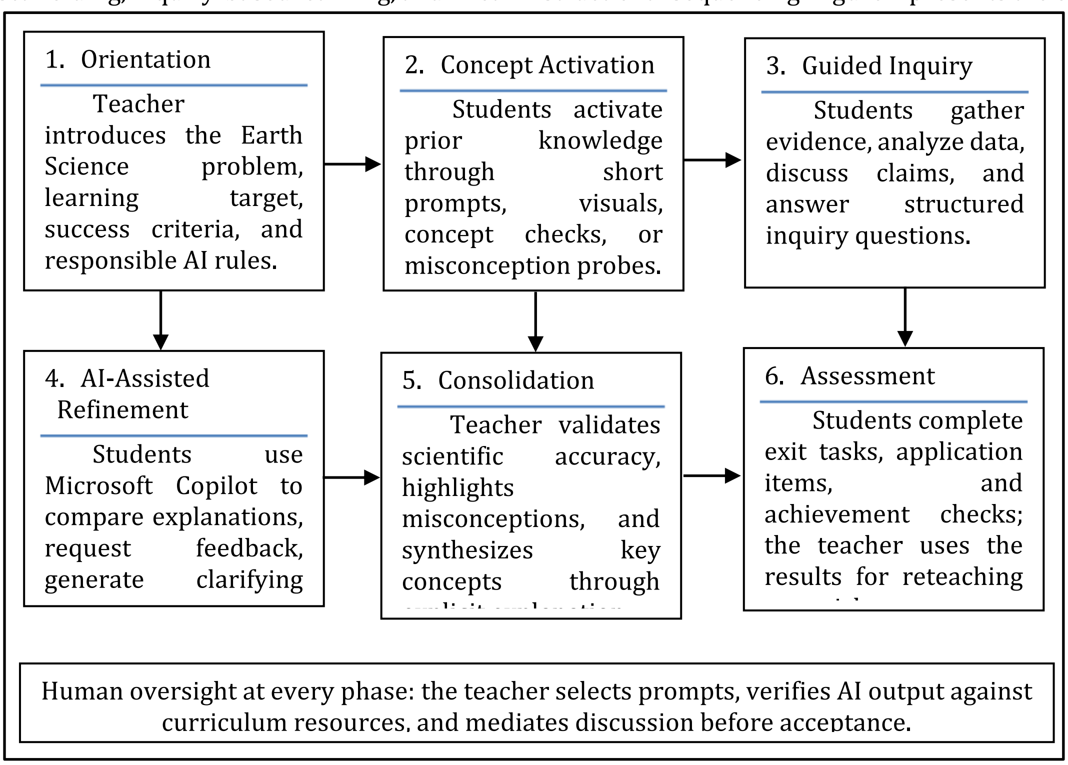

Figure 1. Operational flow of the TAIL-Earth model.

Operationally, TAIL unfolds through six phases. (1) Orientation introduces the target concept, success criteria, and responsible AI rules. (2) Concept activation surfaces prior knowledge and misconceptions. (3) Guided inquiry engages students in evidence-based tasks such as analyzing data, observing representations, or defending preliminary explanations. (4) AI-assisted refinement allows students to use Microsoft Copilot to generate clarifying questions, compare explanations, and receive feedback using teacher-approved prompts. (5) Consolidation provides explicit teacher validation, misconception correction, and synthesis of key concepts. (6) Assessment checks understanding through exit tasks, application items, and the EST. Human oversight operates across all phases; students are not asked to accept AI outputs without verification against class resources and teacher feedback.

Compared with related integrated frameworks, TAIL's distinctive contribution is not the presence of inquiry, technology, or sequencing alone, but the explicit orchestration of these elements into a repeatable classroom routine. The model therefore contributes a more operational account of how classroom AI can be inserted into inquiry-based science instruction without weakening disciplinary accuracy or teacher authority.

Moreover, the TAIL model is a synthesized instructional framework that integrates four complementary components designed to enhance conceptual understanding and achievement in Earth Science. First, the TPCK component functions as the teacher's design lens, guiding the deliberate alignment of content, pedagogy, and technology so that instructional decisions directly address learner misconceptions and competency targets. Second, AI-supported learning serves as an adaptive scaffold, enabling students to engage with structured questioning, receive immediate feedback on explanations, and practice concepts formatively within teacher-defined parameters and verification procedures that ensure accuracy and responsible use. Third, inquiry-based learning operates as the epistemic engine of the model, positioning learners as active constructors of knowledge who generate questions, analyze evidence, formulate claims, and justify conclusions through scientific reasoning. Fourth, the linear instructional flow provides sequenced coherence to instruction, ensuring that lessons unfold through clearly defined phases—orientation, concept activation, guided inquiry, AI-assisted refinement, consolidation, and assessment—thereby reducing cognitive overload and strengthening conceptual integration. Together, these components function synergistically to create a structured yet interactive learning environment intended to optimize engagement, deepen reasoning, and improve academic performance in Earth Science.

3. METHODOLOGY

3.1 Research Design

A quasi-experimental non-equivalent control group pretest–posttest design was employed in this study. This design was selected because it is well-suited for authentic classroom contexts in which intact classes must be used, and random assignment of students is not feasible due to administrative, scheduling, and ethical constraints. The design allowed for comparison of learning outcomes between an experimental group receiving the TAIL intervention and a control group receiving conventional instruction, while controlling for baseline differences using a pretest.

3.2 Participants and Sampling

Participants were Grade 10 students drawn from two intact class sections at one of the schools in Agusan del Sur, Philippines, during the School Year 2025–2026. One section ($n = 43$) was designated as the experimental group and received the TAIL instructional intervention, while the other section ($n = 43$) served as the control group and received conventional instruction, yielding a total sample of 86 participants. A purposive cluster sampling approach was employed, in which pre-existing classes were selected rather than randomly formed groups. Assignment of sections to experimental and control conditions was determined by practical scheduling considerations, including timetable alignment and teacher workload distribution, to ensure the feasibility of implementation while maintaining normal school operations.

3.3 AI Tools and Students' Interaction Protocol

The AI tool used in the experimental condition was Microsoft Copilot (Copilot Chat/web-based generative AI interface), accessed under teacher-supervised classroom procedures. Copilot was used only during designated formative checkpoints and was not used for autonomous grading or for replacing teacher explanation. Its functions in the intervention were limited to: (a) generating follow-up questions from a teacher-approved prompt, (b) checking the completeness of student explanations, (c) proposing alternative wording for scientific claims, and (d) providing short practice feedback on concept-related items.

To improve replicability, student interaction with Copilot followed a consistent routine. First, students completed an initial task individually or in pairs before seeing any AI output. Second, students entered a constrained prompt prepared or approved by the teacher (e.g., "Check whether this explanation of plate boundary movement is scientifically complete. Identify missing ideas but do not provide the final answer first."). Third, students compared the AI response with their textbook, class notes, and teacher guidance. Fourth, they revised their explanation and marked which changes were accepted, rejected, or flagged for teacher clarification. Finally, the teacher led a short debrief to validate scientific accuracy and discuss why some AI suggestions should not be accepted without evidence. This protocol positioned AI as a scaffold for metacognitive checking rather than as a source of unquestioned authority.

3.4 Intervention

The instructional intervention was conducted over eight weeks, with three sessions per week lasting 50 minutes each. The experimental condition (TAIL-Earth) implemented a structured instructional model characterized by TPACK-informed lesson planning by competency, guided inquiry tasks, Microsoft Copilot-supported formative checkpoints, and a linear progression from orientation to assessment. In this condition, lesson planning systematically aligned Earth Science learning competencies with appropriate pedagogical approaches and technological tools, thereby supporting coherent and purposeful integration of content, pedagogy, and technology (Mishra & Koehler, 2006; Wilson et al., 2020).

Across the intervention period, students in the experimental group engaged in guided inquiry tasks that required them to frame problems, examine evidence, interpret information, and defend scientific explanations. These tasks were scaffolded to support conceptual understanding and scientific reasoning (de Jong et al., 2024). In addition, Microsoft Copilot-supported formative checkpoints were incorporated at designated points in the lesson sequence. These checkpoints included question generation, explanation validation, and feedback-supported revision, and were intended to promote reflective thinking, provide immediate formative support, and reduce unnecessary cognitive load during learning (Wang et al., 2024).

Instruction in the experimental group followed a linear instructional sequence consisting of orientation, concept activation, guided inquiry, AI-assisted feedback and refinement, consolidation, and assessment. This sequencing was designed to provide instructional coherence, minimise fragmentation, and guide students through increasingly complex learning tasks in a structured manner (Rosenshine, 2012; Quintana et al., 2004).

By contrast, the control condition reflected prevailing classroom practice at the school and consisted primarily of conventional discussion, textbook-based exercises, and teacher-led recitation without structured AI or inquiry integration. This condition did not involve deliberate TPACK-informed design, Microsoft Copilot-supported formative processes, or inquiry-oriented learning tasks. Instead, instruction relied mainly on teacher exposition and routine seatwork, consistent with commonly reported patterns in secondary science classrooms (Sharma & Desai, 2023).

3.5 Instruments

3.5.1 Earth Science Test (EST)

Student achievement was measured using a researcher-developed 40-item multiple-choice test aligned with Grade 10 Earth Science competencies and corresponding curriculum standards. Test content validity was established through expert review by five specialists in science education and assessment, yielding a scale-level Content Validity Index (S-CVI) of .92, which exceeds the commonly accepted adequacy threshold of .90 for educational instruments (Polit & Beck, 2021). Pilot testing with a comparable cohort produced a Cronbach's alpha coefficient of .86, indicating high internal consistency reliability for group-level measurement (Watkins, 2021). Item analysis procedures were conducted to evaluate difficulty and discrimination indices; only items demonstrating moderate difficulty and acceptable discrimination were retained, consistent with recommended practices for classroom-based achievement tests (Haladyna & Rodriguez, 2019). Sample question is "Which geologic feature is most likely to form at an oceanic–continental convergent boundary? It has the following choices: A. Rift valley B. Volcanic arc C. Mid-ocean ridge D. Transform fault.

3.6 Validity, Reliability, and Fidelity Procedures

Parallel pretest and posttest forms were administered, with randomized item order to minimize recall and testing effects. Instructional fidelity was monitored using a 12-indicator lesson implementation checklist completed weekly by a trained peer observer. The experimental group demonstrated a mean fidelity score of 91.3%, indicating strong adherence to intervention procedures. Monitoring fidelity is widely recognized as essential in quasi-experimental classroom studies to ensure that outcome differences can be attributed to the intervention rather than implementation variability (O'Donnell, 2022).

3.7 Ethical Considerations

All procedures complied with institutional and school-division research requirements. Written informed consent from parents or guardians and assent from students were secured before data collection. Participant anonymity was protected through coded identifiers, and results were reported only in an aggregated form. The use of artificial intelligence tools followed school policy and international guidance on human-centered, transparent, and developmentally appropriate educational AI implementation (UNESCO, 2021; OECD, 2021). Ethical conduct throughout

ANALISTA: International Journal of Multidisciplinary Research (AIJMR)

Volume 4, Issue 1, June 2026 | ISSN: 3048-6130

the study adhered to established research standards for educational investigations involving minors (American Psychological Association, 2020).

4. DATA ANALYSIS & RESULTS

Descriptive statistics (mean and standard deviation) summarized baseline and outcome performance. To examine pre-intervention equivalence (RQ1), an independent-samples *t*-test compared pretest scores between groups. To evaluate post-intervention differences (RQ2), analysis of covariance (ANCOVA) was conducted using pretest scores as the covariate, a recommended approach for adjusting baseline differences in nonrandomized designs (Field, 2018). Effect sizes (RQ3) were calculated using partial eta squared (η^2_p) for ANCOVA and Cohen's d for adjusted mean differences, enabling interpretation of practical significance alongside statistical significance. All inferential tests used α = .05, and 95% confidence intervals were reported to convey estimate precision and support transparent interpretation (Cumming, 2021).

4.1 Participant Profile

Table 1 presents the demographic profile of participants in the experimental and control groups. Both groups were comparable in terms of sex distribution and age. The experimental group consisted of 46.5% males and 53.5% females, while the control group included 48.8% males and 51.2% females. The mean age of students was also similar across groups (Experimental: M = 15.8, SD = 0.6; Control: M = 15.7, SD = 0.6). Such similarity suggests demographic equivalence prior to intervention, which strengthens internal validity in quasi-experimental studies where random assignment is not feasible (Shadish et al., 2022). Because intact classes were used, non-random assignment was acknowledged as a design constraint but statistically addressed through baseline testing and covariate adjustment.

Table 1: Participant Profile

Variable	Experimental (n=43)	Control (n=43)	Total (N=86)
Male, n (%)	20 (46.5)	21 (48.8)	41 (47.7)
Female, n (%)	23 (53.5)	22 (51.2)	45 (52.3)
Age (years), Mean ± SD	15.8 ± 0.6	15.7 ± 0.6	15.8 ± 0.6

Note: Groups were intact classes; assignment was non-random.

4.2 Baseline Equivalence

Table 2 summarizes pretest descriptive statistics and the independent-samples *t*-test used to examine initial group equivalence. The experimental group (M = 15.37, SD = 4.21) and control group (M = 14.95, SD = 4.08) differed by only 0.42 points. This difference was not statistically significant, t (84) = 0.48, p = .633, with a 95% confidence interval of [−1.32, 2.16]. The effect size was negligible (d = 0.10), indicating a trivial practical difference (Cohen, 1988; Lakens, 2017). These results indicate that the two groups were statistically comparable before the intervention, satisfying the assumption of baseline equivalence and supporting the validity of subsequent outcome comparisons.

Table 2: Pretest descriptive statistics and baseline comparison (RQ1)

Group	n	Mean	SD
Experimental	43	15.37	4.21
Control	43	14.95	4.08
Baseline test			**Value**
Mean difference (Exp – Ctrl)			0.42
95% CI of difference			[-1.32, 2.16]
t(df=84)			0.48
p-value			.633
Cohen's d			0.10

Interpretation: No significant pretest difference between groups.

4.3 Post-Test Performance

As shown in Table 3, posttest scores increased in both groups; however, gains were notably larger in the experimental group. The experimental group achieved a mean posttest score of 30.21 (SD = 4.87), compared with 24.81

ANALISTA: International Journal of Multidisciplinary Research (AIJMR)

Volume 4, Issue 1, June 2026 | ISSN: 3048-6130

(SD = 5.12) for the control group. Mean gain scores further highlight this pattern, with the experimental group improving by 14.84 points versus 9.86 points for the control group. These descriptive findings suggest a substantial advantage associated with the intervention, although inferential testing is required to determine whether this difference remains significant after controlling for baseline performance.

Table 3: Posttest Descriptive Statistics by Group

Group	n	Mean	SD	Mean Gain (Post–Pre)
Experimental	43	30.21	4.87	14.84
Control	43	24.81	5.12	9.86

Note. Raw posttest means favor the experimental group.

4.4 Intervention Effects

Table 4 presents the results of the ANCOVA analysis using pretest scores as a covariate. The covariate significantly predicted posttest performance, $F(1,83) = 45.10, p < .001$, partial $\eta^2 = .352$, indicating that prior knowledge accounted for a substantial proportion of variance in outcomes. After adjusting for pretest differences, a statistically significant group effect was observed, $F(1,83) = 34.52, p < .001$, partial η^2 = .294. The adjusted mean difference between groups was 5.31 points (95% CI [3.43, 7.18]), demonstrating that students exposed to the TAIL model performed significantly better than those receiving conventional instruction.

The partial eta squared value of .294 indicates a large effect according to commonly accepted benchmarks ($.01 = small, .06 = medium, .14 = large$), suggesting that approximately 29.4% of the variance in posttest scores can be attributed to the instructional intervention after controlling for pretest scores (Richardson, 2011). The adjusted Cohen's d of 0.93 likewise indicates a large practical effect, meaning the magnitude of improvement is educationally meaningful, not merely statistically detectable (Hattie, 2018). Assumption testing confirmed model validity: the homogeneity of regression slopes was satisfied, $F(1,82) = 0.68, p$ .412, and Levene's test indicated homogeneity of variance, $F(1,84) = 1.37, p = .245$.

Table 4: Inferential Test Results (ANCOVA with Pretest Covariate)

Source	**F**	**df**	**p**	**Partial η^2**	**95% CI / Notes**
Pretest (covariate)	45.10	1,83	<.001	.352	Significant baseline predictor
Group (TAIL-Earth vs Control)	34.52	1,83	<.001	.294	Adjusted mean diff = 5.31, 95% CI [3.43, 7.18]
Homogeneity of slopes ($Pretest \times Group$)	0.68	1,82	.412	—	Assumption met
Levene's test	1.37	1,84	.245	—	Variance assumption met

Additional effect size: **Cohen's d (adjusted)** = **0.93** (large).

5. Summary of Findings

Table 5 synthesizes statistical decisions aligned with each research question. For RQ1, the null hypothesis was retained, confirming no significant baseline difference between groups. For RQ2, the null hypothesis was rejected, indicating that the experimental group significantly outperformed the control group after controlling for initial ability. For RQ3, effect size estimates demonstrated a large and practically meaningful impact of the TAIL-Earth intervention on Earth Science achievement. Taken together, these findings provide convergent statistical evidence that the integrated instructional model produced substantial learning gains beyond those achieved through conventional teaching approaches.

Table 5: Summary of findings aligned to research questions

Research Question	**Statistical Evidence**	**Decision**	**Summary Finding**
RQ1: Pretest difference between groups?	$t(84) = 0.48, p = .633, d = 0.10$	Fail to reject H0	Groups were statistically comparable at baseline.
RQ2: Posttest difference controlling baseline?	ANCOVA group effect $F(1,83) = 34.52, p < .001, partial\ \eta^2 = .294$	Reject H0	The experimental group significantly outperformed the control after adjustment.

RQ3: Effect size of TAIL?	partial $\eta^2 = .294;\ d = 0.93$	—	Large practical effect on Earth Science achievement.

6. Discussion

The findings indicate that the TAIL-Earth model improved Grade 10 Earth Science achievement beyond conventional instruction, with a large effect size. This pattern is theoretically coherent with scholarship showing that technology integration is most productive when it is aligned with content and pedagogy rather than introduced as tool novelty. It is also consistent with recent work in science education suggesting that AI can enhance learning when used as a scaffold for questioning, feedback, and explanation checking within teacher-mediated routines rather than as a replacement for instruction.

The inquiry component of TAIL likely contributed by positioning students as evidence users and explanation builders, while the linear phase structure may have reduced instructional fragmentation and cognitive overload. Recent discussions in science education support this blended stance. Inquiry is valuable, but it is most effective when balanced with explicit guidance, especially for learners who are still developing prerequisite knowledge. In this sense, TAIL's sequencing is not incidental; it is part of the mechanism through which the intervention may have improved achievement.

At the same time, the results should be interpreted cautiously because alternative explanations cannot be fully eliminated. Possible confounding variables include teacher enthusiasm for the intervention, novelty effects associated with AI use, variation in student digital familiarity, peer dynamics within intact classes, and differences in out-of-class study habits. Although baseline equivalence and ANCOVA adjustment reduce some threats, quasi-experimental designs cannot fully control for all unmeasured influences. The reported effect, therefore, supports strong practical promise, but not definitive proof that every observed gain was caused solely by the TAIL model.

The implications for teacher development are substantial. Effective classroom AI use requires more than technical access; teachers need support in prompt design, output verification, bias awareness, and the orchestration of AI checkpoints inside disciplinary instruction. Professional learning communities, LAC sessions, and INSET programs may therefore focus on designing teacher-approved prompt banks, building routines for checking AI accuracy, and aligning AI use with competency-based lesson objectives. The findings also have curriculum implications. Rather than treating inquiry, explanation, and digital support as separate strands, curriculum guides and lesson exemplars may be redesigned to include structured inquiry tasks followed by AI-supported refinement and teacher-led conceptual consolidation.

7. CONCLUSION

This quasi-experimental study found that the TAIL-Earth model significantly improved Earth Science achievement among Grade 10 students at Azpetia National High School. Baseline equivalence was established, and adjusted posttest comparisons showed a statistically significant and educationally large advantage for the experimental group. The findings support integrated, teacher-led innovation that combines TPACK-informed design, responsible AI use, guided inquiry, and coherent instructional sequencing. The study's contribution lies both in its positive achievement results and in its operational clarification of how generative AI can be inserted into a structured inquiry-based lesson without displacing teacher judgment.

7.1 Implications for Practice

The findings of this study carry important implications for classroom practice, teacher development, assessment, governance, and curriculum design. In Earth Science instruction, inquiry-based tasks should be deliberately integrated with explicit conceptual consolidation rather than treated as competing approaches, as this combination appears more responsive to the abstract and systems-oriented nature of the subject. The results also suggest the need for sustained school-based professional learning that equips teachers to design effective prompts, verify AI-generated outputs, and use AI strategically for formative feedback and misconception repair within TPACK-guided instruction. In evaluating instructional innovations in authentic school contexts where intact classes are commonly used, baseline-adjusted analyses such as ANCOVA may provide a more valid basis for interpretation than simple posttest comparisons alone. At the same time, schools should establish clear protocols for classroom AI use that address verification routines, transparency, privacy, and age-appropriate safeguards to ensure responsible and pedagogically sound implementation. Finally, curriculum and lesson design may be strengthened by positioning AI-supported checkpoints only after students have generated initial ideas or explanations, thereby preserving cognitive engagement, supporting disciplinary reasoning, and reinforcing the learner's active role in knowledge construction.

7.2 Limitations and Future Research

Several limitations should be acknowledged more explicitly. First, the study involved only two intact classes from a single school, which constrains external validity. The observed effect may reflect contextual features of the school, teacher, timetable, or cohort that are not easily generalized to other settings.

Second, the quasi-experimental design limits causal inference because students were not randomly assigned, and intact classes may differ in unmeasured ways despite comparable pretest performance. Third, the intervention measured short-term achievement only; it did not assess retention, transfer, changes in inquiry quality, or the nature of students' AI interactions over time. Fourth, because AI use was supervised and tightly bound, the findings should not be generalized to unrestricted student use of generative AI.

Future studies should therefore include multiple schools, larger samples, and, where feasible, randomized or cluster-randomized designs. Mixed-methods studies could complement achievement data with classroom observation, interview data, prompt logs, and artifact analysis to examine how AI influences scientific reasoning. Additional work could also compare different levels of AI scaffolding, investigate retention and transfer, and test whether TAIL is equally effective across science domains beyond Earth Science.

REFERENCES

1) Allen, L. K., & Kendeou, P. (2023). ED-AI LIt: An Interdisciplinary Framework for AI Literacy in Education. *Policy Insights From the Behavioral and Brain Sciences*, *11*(1), 3–10. https://doi.org/10.1177/23727322231220339
2) Almasri, F. (2024). Exploring the Impact of Artificial intelligence in teaching and learning of Science: A Systematic Review of Empirical research. *Research in Science Education*, *54*(5), 977–997. https://doi.org/10.1007/s11165-024-10176-3
3) Anwar, M. N., Mushtaq, N., Mubeen, A., & Iqbal, M. (2024). The power of ZPD: Enhancing teaching and learning. *Journal of Education and Social Studies*, *5*(2), 396–405. https://doi.org/10.52223/jess.2024.5220
4) *APA PsycNet*. (n.d.). https://psycnet.apa.org/record/2024-19513-001
5) Artificial Intelligence: UNESCO launches Women4Ethical AI expert platform to advance gender equality. (2024, February 22). *UNESCO*. https://www.unesco.org/en/articles/artificial-intelligence-unesco-launches-women4ethical-ai-expert-platform-advance-gender-equality
6) Bearman, M., & Ajjawi, R. (2025). Artificial intelligence and gender equity: An integrated approach for health professional education. *Medical Education*, *59*(10), 1049–1057. https://doi.org/10.1111/medu.15657
7) Bergdahl, N., Bond, M., Sjöberg, J., Dougherty, M., & Oxley, E. (2024). Unpacking student engagement in higher education learning analytics: a systematic review. *International Journal of Educational Technology in Higher Education*, *21*(1). https://doi.org/10.1186/s41239-024-00493-y
8) Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational Technology: towards a conceptual framework. *Journal of Interactive Media in Education*, *2019*(1). https://doi.org/10.5334/jime.528
9) Busuttil, L. (2025). Constructing AI Literacy: A Hands-On Approach for School Children. *Informatics in Education*, 761–779. https://doi.org/10.15388/infedu.2025.26
10) Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, *10*(1). https://doi.org/10.1186/s40594-023-00418-7
11) Cheng, C., Wang, J., Zhai, X., & Yang, Y. C. (2025). AI literacy and gender equity in elementary education: A quasi-experimental study of a STEAM–PBL–AIoT course with questionnaire validation. *International Journal of STEM Education*, *12*(1). https://doi.org/10.1186/s40594-025-00574-y
12) Cheryan, S., Lombard, E. J., Hailu, F., Pham, L. N. H., & Weltzien, K. (2024). Global patterns of gender disparities in STEM and explanations for their persistence. *Nature Reviews Psychology*, *4*(1), 6–19. https://doi.org/10.1038/s44159-024-00380-3
13) Chiu, T. K. (2021). Digital support for student engagement in blended learning based on self-determination

theory. *Computers in Human Behavior*, *124*, 106909. https://doi.org/10.1016/j.chb.2021.106909
14) Chiu, T. K. F. (2021a). Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *Journal of Research on Technology in Education*, *54*(sup1). https://doi.org/10.1080/15391523.2021.1891998
15) Chiu, T. K. F. (2021b). Student engagement in K-12 online learning amid COVID-19: A qualitative approach from a self-determination theory perspective. *Interactive Learning Environments*, *31*(6), 3326–3339. https://doi.org/10.1080/10494820.2021.1926289
16) Cho, N., Liao, W., & Lalitphasan, U. (2025). Empowering teachers with artificial intelligence literacy: A systematic review on teacher professional development. *Research Square*. https://doi.org/10.21203/rs.3.rs-8267583/v1
17) Cobian, K. P., Hurtado, S., Romero, A. L., & Gutzwa, J. A. (2024). Enacting inclusive science: Culturally responsive higher education practices in science, technology, engineering, mathematics, and medicine (STEMM). *PLoS ONE*, *19*(1), e0293953. https://doi.org/10.1371/journal.pone.0293953
18) Demirbag, M., & Bahcivan, E. (2021). Comprehensive Exploration of Digital Literacy: Embedded with Self-Regulation and Epistemological Beliefs. *Journal of Science Education and Technology*, *30*(3), 448–459. https://doi.org/10.1007/s10956-020-09887-9
19) Dost, G. (2024). Students' perspectives on the 'STEM belonging' concept at A-level, undergraduate, and postgraduate levels: an examination of gender and ethnicity in student descriptions. *International Journal of STEM Education*, *11*(1). https://doi.org/10.1186/s40594-024-00472-9
20) Du, H., Sun, Y., Jiang, H., Islam, A. Y. M. A., & Gu, X. (2024). Exploring the effects of AI literacy in teacher learning: an empirical study. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-03101-6
21) Duncan, V. L., Holt, E. A., & Keenan, S. M. (2023). Creating an equitable and inclusive STEM classroom: a qualitative meta-synthesis of approaches and practices in higher education. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1154652
22) Epler-Ruths, C. M., McDonald, S., Pallant, A., & Lee, H. (2020). Focus on the notice: evidence of spatial skills' effect on middle school learning from a computer simulation. *Cognitive Research Principles and Implications*, *5*(1), 61. https://doi.org/10.1186/s41235-020-00263-0
23) Fadjarajani, S., As'ari, R., Ningrum, E., & Ahmad, M. Z. (2024). GeoSmart Application as a Tool for Interactive Learning: Literature Review Completed with Experiments for Improving Understanding of Geospheric Phenomena in Geography Education. *ASEAN Journal of Science and Engineering*, *4*(3), 405–428. https://doi.org/10.17509/ajse.v4i3.74409
24) Fitzpatrick, M., & Hedley, N. (2024). Review of the state of practice in geovisualization in the geosciences. *Frontiers in Earth Science*, *11*. https://doi.org/10.3389/feart.2023.1230973
25) Florian, L. (2015). Inclusive Pedagogy: A transformative approach to individual differences but can it help reduce educational inequalities? *Scottish Educational Review*, *47*(1), 5–14. https://doi.org/10.1163/27730840-04701003
26) Getenet, S., Cantle, R., Redmond, P., & Albion, P. (2024). Students' digital technology attitude, literacy and self-efficacy and their effect on online learning engagement. *International Journal of Educational Technology in Higher Education*, *21*(1). https://doi.org/10.1186/s41239-023-00437-y
27) Godsk, M., & Møller, K. L. (2024). Engaging students in higher education with educational technology. *Education and Information Technologies*, *30*(3), 2941–2976. https://doi.org/10.1007/s10639-024-12901-x
28) Gülsoy, A., Uyan, Y., Özcan, E., & Durmuş, M. (2025). The relationship between disaster literacy levels and disaster preparedness among adults. *Natural Hazards*, *121*(9), 10667–10681. https://doi.org/10.1007/s11069-025-07225-z
29) Halkiopoulos, C., & Gkintoni, E. (2024). Leveraging AI in E-Learning: Personalized Learning and Adaptive Assessment through Cognitive Neuropsychology—A Systematic Analysis. *Electronics*, *13*(18), 3762. https://doi.org/10.3390/electronics13183762
30) Handoyo, B., Soekamto, H., Putra, A. K., Kamil, P. A., & Wulandari, F. (2024, November 8). *Disaster preparedness: The role of spatial disaster learning using geospatial technology*. Handoyo | Jàmbá: Journal of Disaster Risk Studies. https://jamba.org.za/index.php/jamba/article/view/1576/3266

31) Hariyanto, Kristianingsih, F. X. D., & Maharani, R. (2025). Artificial intelligence in adaptive education: a systematic review of techniques for personalized learning. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00908-6
32) Holder, L. N., Scherer, H. H., & Herbert, B. E. (n.d.). *Student Learning of Complex Earth Systems: A model to guide development of student expertise in Problem-Solving.* https://eric.ed.gov/?id=EJ1161321
33) Jackson-Summers, A. G., Mrakovcich, K. L., Gray, J. P., Fleischmann, C. M., Emami, T., & Page, E. J. (2024). A systematic review of inclusive pedagogical research using the CIRTL inclusive pedagogy framework: multi-disciplinary and STEM perspectives, current trends and a research agenda. *Discover Education*, *3*(1). https://doi.org/10.1007/s44217-024-00093-y
34) Javier-Aliaga, D., Neyra, O. R. S., Calizaya-Milla, Y. E., & Saintila, J. (2024). Academic self-efficacy and digital competence in a sample of university students. *Contemporary Educational Technology*, *16*(4), ep540. https://doi.org/10.30935/cedtech/15601
35) Jia, F., Sun, D., & Looi, C. (2023a). Artificial Intelligence in Science Education (2013–2023): Research trends in ten years. *Journal of Science Education and Technology*, *33*(1), 94–117. https://doi.org/10.1007/s10956-023-10077-6
36) Jia, F., Sun, D., & Looi, C. (2023b). Artificial Intelligence in Science Education (2013–2023): Research trends in ten years. *Journal of Science Education and Technology*, *33*(1), 94–117. https://doi.org/10.1007/s10956-023-10077-6
37) *K-12 AI curricula: A mapping of government-endorsed AI curricula*. (2025, April 15). https://www.unesco.org/en/articles/k-12-ai-curricula-mapping-government-endorsed-ai-curricula
38) Karim, R. A. (2025). The role of Artificial intelligence (AI) in Advancing gender equality and Women's Empowerment in Academia: A Conceptual framework. *European Scientific Journal ESJ*, *21*(43), 73. https://doi.org/10.19044/esj.2025.v21n43p73
39) Kavitha, K., & Joshith, V. P. (2024). Pedagogical incorporation of artificial intelligence in K-12 science education: A decadal bibliometric mapping and systematic literature review (2013-2023). *Journal of Pedagogical Research*. https://doi.org/10.33902/jpr.202429218
40) Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *ACM Digital Library*, 1–16. https://doi.org/10.1145/3313831.3376727
41) Master, A., Alexander, T., Thompson, J., Fan, W., Meltzoff, A. N., & Cheryan, S. (2024). Causes and consequences of stereotypes: interest stereotypes reduce adolescent girls' motivation to enroll in computer science classes. *Journal of Research on Technology in Education*, *57*(1), 56–83. https://doi.org/10.1080/15391523.2024.2402355
42) Master, A., Meltzoff, A. N., & Cheryan, S. (2021a). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48). https://doi.org/10.1073/pnas.2100030118
43) Master, A., Meltzoff, A. N., & Cheryan, S. (2021b). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48). https://doi.org/10.1073/pnas.2100030118
44) McLaughlin, J. A., & Bailey, J. M. (2022). Students need more practice with spatial thinking in geoscience education: a systematic review of the literature. *Studies in Science Education*, *59*(2), 147–204. https://doi.org/10.1080/03057267.2022.2029305
45) Msambwa, M. M., Daniel, K., Lianyu, C., & Antony, F. (2024). A systematic review using feminist perspectives on the factors affecting girls' participation in STEM subjects. *Science & Education*, *34*(3), 1619–1650. https://doi.org/10.1007/s11191-024-00524-0
46) Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, López-Pernas, S., & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. *Smart Learning Environments*, *11*(1). https://doi.org/10.1186/s40561-024-00350-5
47) Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI Literacy: An Exploratory Review. *Computers and Education Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041
48) Nshimiyimana, E., & Ndayambaje, I. (2025). Enhancing students' understanding of Earth's spatial relationships

using virtual reality: a case study of secondary schools in Nyamasheke District, Rwanda. *African Journal of Empirical Research*, *6*(1), 773–791. https://doi.org/10.51867/ajernet.6.1.66

49) Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003
50) Ribeiro, T., & Orion, N. (2021). Educating for a Holistic View of the Earth System: A review. *Geosciences*, *11*(12), 485. https://doi.org/10.3390/geosciences11120485
51) Rogers, T., & Carbonaro, M. (2025). From Understanding to Creating: Bridging AI literacy and AI fluency in K-12 education. *Journal of Teaching and Learning*, *19*(4). https://doi.org/10.22329/jtl.v19i4.9420
52) Roozafzai, Z. S., & Zaeri, P. (n.d.). *Enhancing Comprehension of Earth Science Concepts through Digital Animation and Inclusive Discourse.* https://eric.ed.gov/?id=EJ1467650
53) Simply Psychology. (2025, March 19). *Kolb's Learning Styles & Experiential Learning Cycle*. https://www.simplypsychology.org/learning-kolb.htm
54) Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education Artificial Intelligence*, *4*, 100127. https://doi.org/10.1016/j.caeai.2023.100127
55) Sperling, K., Stenberg, C., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, *6*, 100169. https://doi.org/10.1016/j.caeo.2024.100169
56) Tal, M., Lavi, R., Reiss, S., & Dori, Y. J. (2024). Gender Perspectives on Role Models: Insights from STEM Students and Professionals. *Journal of Science Education and Technology*, *33*(5), 699–717. https://doi.org/10.1007/s10956-024-10114-y
57) Taylor, H. A., Burte, H., & Renshaw, K. T. (2023). Connecting spatial thinking to STEM learning through visualizations. *Nature Reviews Psychology*, *2*(10), 637–653. https://doi.org/10.1038/s44159-023-00224-6
58) Wang, N., & Lester, J. (2023). K-12 Education in the Age of AI: A Call to Action for K-12 AI Literacy. *International Journal of Artificial Intelligence in Education*, *33*(2), 228–232. https://doi.org/10.1007/s40593-023-00358-x
59) Willis, S., & Stern, R. J. (2025). Plate Tectonics and Metamorphism: Teaching complex systems using videos and animations. *Geosciences*, *15*(3), 79. https://doi.org/10.3390/geosciences1503007
60) Yim, I. H. Y., & Su, J. (2024). Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education*. https://doi.org/10.1007/s40692-023-00304-90
61) Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, *2021*(1). https://doi.org/10.1155/2021/8812542