IGNATIAN
INTERNATIONAL JOURNAL FOR MULTIDISCIPLINARY RESEARCH
www.icceph.com
ISSN 2984-9942

FROM GUIDED TO INDEPENDENT: AI-ENHANCED GRADUAL RELEASE OF RESPONSIBILITY (GRR) MODEL TO ACCELERATE SCIENTIFIC LITERACY

John Carlo C. Tanzo

Department of Education, Schools Division Office – Muntinlupa City
Sucat Senior High School, Meralco Road, Sucat, Muntinlupa City, Philippines

https://doi.org/10.5281/zenodo.19383410

ABSTRACT

This study examined the effectiveness of the AI-enhanced Gradual Release of Responsibility (GRR) model in accelerating students' scientific literacy and fostering independent research skills with reduced AI dependence. Using a classroom-based sequential explanatory mixed-methods design, the study involved 72 Grade 12 STEM students in a public senior high school in Muntinlupa City, Philippines, distributed into an experimental group and a control group. The intervention was implemented in General Physics 2 across the phases of "I Do," "We Do," and "You Do," with ChatGPT integrated as a monitored cognitive scaffold only in the experimental condition. Quantitative data were gathered through a validated 50-item scientific literacy test and phase-based assessments, while qualitative data were drawn from reflection logs, learning journals, perception responses, rubric-scored outputs, and AI Declaration Tables. Descriptive statistics, Shapiro–Wilk tests, paired-samples t-tests, and an independent-samples t-test were used for quantitative analysis, while qualitative data were examined through thematic analysis and document analysis, followed by thematic–statistical triangulation. Findings showed that both groups improved across the intervention, but the experimental group demonstrated stronger phase-based performance, a significant pre- to post-test gain, and significantly higher post-test scores than the control group, with a moderate effect size. Qualitative results further showed that students progressed from curiosity and apprehension toward reflective, critical, and more autonomous use of AI, demonstrating increased scientific reasoning, improved research-related skills, and reduced dependence on technological support.

Keywords: *artificial intelligence, gradual release of responsibility, scientific literacy, self-regulated learning, student autonomy*

INTRODUCTION

Scientific literacy is widely recognized as a central outcome of contemporary science education because it enables learners not only to recall scientific facts but also to interpret evidence, evaluate claims, construct explanations, and apply scientific understanding to real-life and civic contexts. In this broader sense, scientific literacy is not limited to knowledge acquisition; rather, it includes the ability to use scientific language, reason with evidence, and make informed judgments in situations shaped by science and technology (Holbrook & Rannikmäe, 2009; National Research Council, 2012; Norris & Phillips, 2003). This orientation is especially important in present educational contexts, where students must learn to navigate increasingly complex information environments and distinguish between accurate, misleading, and unsupported claims. Thus, science instruction is now expected to cultivate not only conceptual mastery but also disciplined habits of inquiry, interpretation, and reflective judgment.

The urgency of this agenda is underscored by persistent concerns over student performance in science. In the Philippines, the OECD country note for PISA 2022 reported that only 23% of students attained at least Level 2 proficiency in science, the baseline level associated with recognizing correct explanations for familiar scientific phenomena and identifying whether conclusions are supported by the data provided, while almost no students reached Levels 5 or 6, where learners are expected to apply science knowledge creatively and autonomously in unfamiliar contexts (OECD, 2023). These results suggest that a substantial proportion of learners continue to experience difficulty with the very forms of reasoning and transfer that define scientific literacy. Such a pattern points to the need for instructional approaches that do more than cover content; they must deliberately support learners as they move from tentative understanding toward independent scientific thinking and performance.

One pedagogical framework that directly addresses this developmental transition is the AI-enhanced GRR model, grounded in the broader Gradual Release of Responsibility tradition. Fisher and Frey (2021) describe GRR as a structured instructional framework in which responsibility for learning shifts intentionally from teacher to student through interconnected phases of focused instruction, guided instruction, collaborative learning, and independent learning. The value of this framework lies in its explicit attention to how competence develops: students first observe expert thinking, then participate with support, then negotiate meaning collaboratively, and finally demonstrate understanding independently. In science education, such progression is especially relevant because students are not merely expected to remember definitions; they are expected to explain phenomena, justify conclusions, and use disciplinary reasoning with increasing autonomy. By organizing instruction through sequenced support, the AI-enhanced GRR model offers a strong pedagogical basis for accelerating scientific literacy while preserving the developmental logic of scaffolded learning.

The integration of artificial intelligence into education has intensified interest in new forms of support that may strengthen learning processes, engagement, and feedback. Recent reviews characterize artificial intelligence in education as a rapidly expanding field encompassing adaptive learning, intelligent tutoring, personalized feedback, assessment, and data-informed instructional support (Wang et al., 2024). Within STEM education, AI has been associated with more responsive learning pathways, improved feedback processes, and greater opportunities for personalized support, although its effective use depends on coherent instructional design rather than technological novelty alone (Xu & Ouyang, 2022). UNESCO likewise emphasizes that generative AI should be implemented through a human-centered educational vision, warning that its value lies in supporting human agency and learning rather than displacing judgment, reflection, or teacher mediation (Miao & Holmes, 2023). Together, these positions suggest that AI becomes educationally meaningful not when it simply produces answers, but when it is embedded in pedagogical structures that make learners think more carefully, not less.

This point is especially important in secondary and K–12 contexts, where learners are still developing foundational disciplinary habits and may be more vulnerable to overreliance on algorithmic outputs. Recent K–12 and high-school reviews indicate that generative AI offers promising opportunities for personalization, engagement, feedback, and instructional innovation, yet these same reviews also emphasize unresolved issues concerning teacher guidance, ethical use, reliability, and the relative scarcity of concrete classroom studies showing how AI can be integrated into day-to-day learning without weakening student agency (Akhmetova et al., 2025; Marzano, 2025). In other words, although the literature increasingly documents AI's potential in school settings, there remains a strong need for classroom-based studies that examine how AI can be used as a temporary and accountable support within a clearly sequenced teaching model. For science classrooms, this need is even more critical because the objective is not simply task completion but the development of students who can reason scientifically and work independently with evidence.

From this perspective, the AI-enhanced GRR model is pedagogically compelling because it offers a principled way to integrate AI without normalizing dependence on it. In the early instructional phases, AI may function as a scaffold that helps clarify concepts, generate examples, support questioning, and provide immediate feedback; however, as the lesson sequence progresses, responsibility must shift back to the learner. This logic aligns closely with emerging research on self-regulated learning in AI-supported environments. Chiu (2024) argues that generative AI can foster phases of self-regulated learning by supporting planning, performance, and self-reflection, while Banihashem et al. (2025) found that AI–self-regulated learning research has grown rapidly but remains concentrated primarily on cognitive and metacognitive outcomes, with comparatively limited work on motivation, school-aged learners, and pedagogically integrated classroom designs. The implication is clear: AI has the potential to support self-regulation, but its effects depend on whether learners are guided to monitor, evaluate, and ultimately internalize the thinking processes initially supported by the tool.

This issue becomes even more consequential when higher-order thinking is considered. A recent meta-analysis by Wang and Fan (2025) found that ChatGPT can positively influence learning performance, learning perception, and higher-order thinking, but the authors also stressed that such gains are strongest when AI use is accompanied by appropriate learning scaffolds and educational frameworks. Their conclusion is particularly relevant to science education, where higher-order learning involves explanation-building, evidence evaluation, concept transfer, and problem solving rather than mere answer retrieval. Thus, the educational question is no longer whether AI can assist learners, but under what pedagogical conditions AI support leads to deeper understanding, greater metacognitive regulation, and stronger independence. The AI-enhanced GRR model addresses this question directly by positioning AI within a structure that begins with support but is designed to end in autonomous learner performance.

At the same time, the conceptual literature on scientific literacy reinforces the importance of protecting the learner's interpretive role. Norris and Phillips (2003) argued that scientific literacy necessarily involves the ability to read, interpret, critique, and use scientific representations meaningfully, while Holbrook and Rannikmäe (2009) emphasized that the goal of science education should extend toward functional and socially relevant scientific literacy. These perspectives suggest that scientific literacy cannot be reduced to access to information alone; rather, it requires learners to construct meaning, judge validity, and apply knowledge responsibly. This becomes especially significant in AI-supported settings, where students may be tempted to treat AI-generated text as finished knowledge. If scientific literacy is genuinely the intended outcome, then instructional design must ensure that students remain active interpreters, validators, and users of knowledge. In this regard, the AI-enhanced GRR model provides a stronger foundation than unstructured AI exposure because it ties support to a deliberate process of release, evaluation, and independent application.

Despite the growing literature on AI in education, STEM learning, K–12 implementation, and self-regulated learning, an important gap remains. Existing studies have often examined AI as a tool in isolation, focused on general academic performance rather than scientific literacy as the principal outcome, or emphasized perceptions and adoption without tracing how learners change across instructional phases. Likewise, many AI studies document benefits in feedback, engagement, or personalization, yet fewer studies investigate how these benefits unfold within a phased instructional model that explicitly transitions learners from supported participation to independent performance. The novelty of the present study, therefore, is not simply the use of AI, and not merely the use of GRR. Rather, its contribution lies in the combination of four elements: scientific literacy as the main outcome, phase-based measurement of learning before, during, and after implementation, qualitative tracking of students' perceptions and experiences, and evidence of reduced AI dependence alongside the development of independent research skills. Taken together, these dimensions respond to the current literature's call for more pedagogically grounded, classroom-based, and developmentally sensitive investigations of AI-supported learning.

Anchored in these considerations, the present study investigated the AI-enhanced GRR model to accelerate scientific literacy among students in science learning. Specifically, it examined the students' level of scientific literacy before, during, and after the use of the AI-enhanced GRR model; determined whether there was a significant difference in the pre-test and post-test results of the experimental and control groups; analyzed whether the experimental group differed significantly from the control group in improving scientific literacy after the intervention; explored students' perceptions and experiences toward the AI-enhanced GRR model before, during, and after implementation; and examined how students demonstrated independent research skills and reduced AI dependence. By integrating achievement data with qualitative evidence of learner experience and autonomy, the study aims to contribute a more comprehensive understanding of how AI can be used not as a substitute for thinking, but as a temporary scaffold within an instructional design explicitly intended to move learners from guidance to independence.

Research Questions

1. What is the students' level of scientific literacy before, during, and after the intervention?
2. Is there a significant difference in scientific literacy in the pre-test and post-test of the experimental and control groups?
3. Is there a significant difference between the control and experimental groups in terms of improving students' scientific literacy after the intervention?
4. What are the students' perceptions and experiences toward the GRR model with AI before, during, and after implementation?
5. How do students demonstrate independent research skills and reduced AI dependence?

METHODOLOGY

This study employed a sequential explanatory mixed-methods design, with a quasi-experimental nonequivalent-groups quantitative strand and a qualitative explanatory strand. This design was appropriate because the study sought not only to determine whether the AI-enhanced GRR model improved students' scientific literacy, but also to explain how students experienced the intervention, how they interacted with AI-supported scaffolds, and how they demonstrated reduced AI dependence and increasing research independence over time. In the quantitative strand, the study followed a quasi-experimental nonequivalent-groups design using two intact classes, one assigned as the experimental group and the other as the control group. Because the classes already existed and could not be randomly reorganized, the use of intact groups was considered methodologically appropriate and educationally feasible for a real classroom setting. The subsequent qualitative strand was used to explain and deepen the interpretation of the statistical results, consistent with the logic of explanatory sequential mixed-methods research and classroom-based action inquiry (Creswell & Plano Clark, 2017; Mertler, 2024; Shadish et al., 2002).

The study was conducted in a public senior high school in Muntinlupa City, Philippines during School Year 2025–2026. The research context was a Grade 12 STEM General Physics 2 class, a subject area that requires conceptual understanding, application of scientific principles, interpretation of evidence, and reasoned explanation. The participants consisted of 72 Grade 12 STEM students distributed across two intact classes of 36 students each. One class served as the experimental group exposed to the AI-enhanced GRR model, while the other served as the control group, which received the same lesson objectives, time allotment, content coverage, and teacher guidance but without AI integration. The classes were selected through purposive sampling on the basis of their heterogeneous composition in terms of academic ability, classroom performance, and learning profiles, making them more suitable for comparison in a typical public-school setting. A third section was excluded because it was academically homogeneous and consistently higher performing, which could have reduced comparability between groups. In this way, the selected classes more closely reflected the diversity of learner profiles commonly found in Philippine public senior high school classrooms, while still permitting a structured classroom comparison within an action research design (Mertler, 2024; Shadish et al., 2002).

The intervention was anchored on the AI-enhanced GRR model, which combined the instructional logic of the Gradual Release of Responsibility framework with the supervised use of ChatGPT as a cognitive scaffold. The intervention proceeded through the three major pedagogical phases of “I Do,” “We Do,” and “You Do.” In the “I Do” phase, the teacher modeled key physics concepts, procedures, and scientific reasoning through direct instruction. In the “We Do” phase, students participated in guided and collaborative learning tasks, during which the experimental group used ChatGPT under teacher supervision for clarification, elaboration, and feedback, while the control group completed parallel activities without AI support. In the “You Do” phase, students independently completed tasks without AI assistance, allowing the researcher to examine whether previously scaffolded skills had been internalized and transferred to autonomous performance. This intervention structure enabled the study to examine not only learning gains, but also how support was gradually withdrawn as students moved toward independent scientific reasoning, which is central to both the GRR framework and mixed-methods investigation of process and outcome (Creswell & Plano Clark, 2017; Fetters et al., 2013).

Multiple instruments were used to generate quantitative and qualitative evidence. The main quantitative instrument was a 50-item multiple-choice scientific literacy test designed to assess conceptual understanding, application of scientific concepts, and scientific reasoning in physics. The instrument was reviewed for content alignment, clarity, and appropriateness by Mrs. Leticia Sabala, Master Teacher II, and was pilot-tested in a non-participating Grade 12 STEM section excluded from the main sample. Item analysis was conducted to examine clarity, difficulty level, and discriminating power, and internal consistency was estimated through the Kuder–Richardson Formula 20, yielding a reliability coefficient of 0.83, which indicated good reliability for research use. The same instrument was administered as both pre-test and post-test. In addition, phase-based assessments were employed across the GRR cycle. During the “I Do” phase, students completed short multiple-choice exit assessments to capture immediate conceptual learning. During the “We Do” phase, students completed group-based experimental performance tasks scored through a standardized rubric focusing on procedural accuracy, conceptual reasoning, data interpretation, and quality of scientific explanation. During the “You Do” phase, students completed individual written and computational tasks, scored through a rubric emphasizing conceptual accuracy, coherence of reasoning, evidence-based explanation, and independence of thought. For the qualitative strand, the study used reflection logs, learning journals, perception responses, rubric-scored outputs, and AI Declaration Tables. The AI Declaration Table required each group in the experimental condition to document its original draft, the prompt entered into ChatGPT, the AI-generated response, and the final revised output. This served both as an ethical transparency tool and as a documentary source for examining prompt refinement, evaluative judgment, revision practices, and changing levels of AI reliance (Kuder & Richardson, 1937; Bowen, 2009).

Data gathering was conducted over approximately eight instructional weeks, equivalent to a two-month operational period, with an average of four contact sessions per week. In the first week, administrative approval and ethical requirements were completed, parental consent and student assent were secured, and students were oriented on the nature of the study, confidentiality safeguards, and responsible AI use. Following the orientation, all participants took the pre-test to establish baseline scientific literacy levels. During Weeks 2 and 3, the “I Do” phase was implemented through teacher-led instruction and concept modeling, followed by short exit assessments. During Weeks 4 and 5, the “We Do” phase was conducted through guided collaborative performance tasks. At this stage, the experimental group used ChatGPT as a monitored support tool, while the control group completed comparable guided tasks without AI. Students in the experimental group were explicitly trained in prompt writing, cross-checking of AI-generated content, and verification against lesson materials, and they documented all AI use through the AI Declaration Table. During Weeks 6 and 7, the “You Do” phase required students to complete independent tasks involving circuit design, computation, and written scientific explanation without AI access. In Week 8, all participants completed the post-test, after which reflection logs, journals, and relevant classroom artifacts were collected. Although several instructional adjustments were necessary because of class suspensions related to natural calamities, holidays, and school activities, the researcher preserved the sequence and integrity of the intervention to the extent possible. Ethical procedures were aligned with the DepEd Research Management Guidelines, which govern research conduct, voluntary participation, confidentiality, and responsible implementation in Philippine basic education settings (Department of Education, 2017).

Quantitative data were analyzed using jamovi version 2.4, an open statistical software environment. Descriptive statistics, particularly means and mean percentage scores (MPS), were used to describe students’ scientific literacy before, during, and after the intervention. To assess whether the distribution of scores met the assumptions for parametric testing, the Shapiro–Wilk test was used as the normality test. Within-group differences between pre-test and post-test scores were examined through paired-samples t-tests, while post-test differences between the experimental and control groups were tested using an independent-samples t-test. The level of significance was set at $\alpha = 0.05$, and effect sizes were reported where appropriate in order to estimate the magnitude of observed differences beyond statistical significance alone. This analytical sequence was appropriate for a quasi-experimental explanatory design in which the first objective was to determine whether measurable gains occurred and whether those gains differed by condition before drawing on qualitative evidence for deeper interpretation (Shapiro & Wilk, 1965; The jamovi project, 2024; Creswell & Plano Clark, 2017).

Qualitative data from reflection logs, learning journals, perception responses, rubric-scored outputs, and AI Declaration Tables were analyzed using thematic analysis following the six-phase approach of Braun and Clarke (2006), which involves familiarization with the data, generation of initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the analytic report. To complement students' written reflections, the study also used document analysis of classroom artifacts in order to identify concrete evidence of prompt refinement, evaluation of AI outputs, revision behavior, conceptual correction, and the gradual reduction of AI dependence across phases. This was important because the study aimed not only to document what students said about the intervention but also to examine how their outputs reflected increasing independence and scientific reasoning. To enhance trustworthiness, the coding and thematic interpretation were peer-reviewed and counterchecked by an external reviewer with psychometric and higher education research expertise. This strategy is consistent with recommendations for strengthening the credibility, traceability, and trustworthiness of thematic analysis and qualitative interpretation (Braun & Clarke, 2006; Bowen, 2009; Lincoln & Guba, 1985; Nowell et al., 2017).

Finally, the quantitative and qualitative strands were integrated through thematic–statistical triangulation, such that the qualitative themes and documentary evidence were used to contextualize and explain the statistical results obtained across the GRR phases. In this way, numerical indicators of scientific literacy development were interpreted alongside students' perceptions, learning experiences, and artifact-based evidence of increasing autonomy. The mixed-methods integration was therefore not treated as a mere presentation of parallel results, but as an interpretive strategy for understanding both the extent and the process of learning under the AI-enhanced GRR model. The scope of the study was delimited to Grade 12 STEM students enrolled in General Physics 2 in one public senior high school during one grading period, and the intervention focused only on the competencies included in the selected lessons taught over approximately eight weeks. Accordingly, the findings should be interpreted within the limits of a single-school setting, one subject area, and intact-group comparison without random assignment. In addition, because the study examined AI use specifically through ChatGPT under teacher-regulated conditions, the findings should not be generalized automatically to other AI platforms or to unrestricted classroom AI use. These delimitations are consistent with the practical and contextual nature of action research and quasi-experimental educational inquiry (Mertler, 2024; Shadish et al., 2002; Fetters et al., 2013).

RESULTS

Table 1. Mean Percentage Scores and Mean Across Teaching and Learning Phases

Group	Pre-test (MPS)	I Do (MPS)	We Do (Mean)	You Do (Mean)	Post-Test (MPS)
Experimental	44.61	89.67	93.63	85.96	83.16
Control	44.27	83.56	88.62	83.40	57.17

Table 2. Paired-Samples t-Test for Pre-test and Post-test Scores of the Experimental Group

Group	Pre-test MPS	Post-test MPS	t-value	df	p-value	Decision
Experimental	44.61	83.16	-7.59	35	<0.001	Reject H_0

Shapiro-Wilk (w) = 0.95, (p) = 0.089

Table 3. Paired-Samples t-Test for Pre-test and Post-test Scores of the Control Group

Group	Pre-test MPS	Post-test MPS	t-value	df	p-value	Decision
Control	44.27	57.17	-3.36	35	0.002	Reject H_0

Shapiro-Wilk (w) = 0.96, (p) = 0.27

Table 4. Significant difference between Control and Experimental in the Post-test

Group	Experimental MPS	Control MPS	t-value	p-value	Cohen's d	Decision
Post-test scientific literacy score	83.16	57.17	2.17	0.038	0.51	Reject H_0

Shapiro-Wilk (w) = 0.95, (p) = 0.086

Table 5. Thematic Analysis of the Students' Perceptions and Experiences Toward the Intervention

Phase	Codes	Subthemes	Major Themes
Before Implementation	Curiosity, excitement, fear of failure, self-doubt, uncertainty about AI accuracy	Mixed expectations and emotional ambivalence	Curiosity and apprehension about AI with openness to learning tools.
	Expectation of ease and innovation	Anticipation of modernized learning	
During Implementation	AI as guide and source of feedback, confusion from complex outputs, searching for reliable information, self-correction	Use of AI as scaffolding and support for problem-solving	AI as a Scaffold for Learning, Engagement, and Active Participation
	Active engagement with technology, exploration, practice with AI support	Development of self-directed learning behaviors	
After Implementation	Increased independence, self-confidence, realization of own learning capacity, reduced reliance on AI	Transition from dependence to autonomy	From Dependence to Autonomy: Developing Reflective, Adaptive Learners with AI Support
	Appreciation of AI's supportive role, recognition of limits, reflection on learning growth	Integration of AI as learning partner	

Table 6. Thematic analysis on students' independent research skills and AI dependence

GRR Phase	Codes	Subthemes	Major Themes
"We Do" Phase (AI Supportive)	AI prompt refinement, evaluation of AI output, synthesis of ideas, application of conceptual	Guided critical engagement with AI responses	AI-Mediated Collaborative Learning for Cognitive and

	knowledge, improvement of sentence structure		Research Skill Development
	Recognition of inaccuracies, revision of AI's overly complex terms, contextualization of scientific concepts	Metacognitive regulation and discernment in AI use	
"You Do" Phase (No AI Support)	Independent rewriting, scientific reasoning, conceptual accuracy, evidence-based explanation	Transfer of learned skills without external aid	From Supported Engagement to Independent Research Performance
	Self-directed validation of concepts, use of prior AI-guided learning, critical evaluation of content	Independent problem-solving and application	
Across Phases	Reduced AI reliance, confidence in constructing final outputs, internalization of learned formats and scientific tone	Sustained learning and autonomy	Internalization of Research and Writing Skills

DISCUSSION

Table 1. Mean Percentage Scores and Mean Across Teaching and Learning Phases

Table 1 indicates a clear phase-based pattern in scientific literacy development. Although both groups began from comparable baseline levels, improved during instruction, and showed some decline under more independent conditions, the experimental group demonstrated a stronger overall trajectory, particularly during guided practice and in the post-test. This pattern suggests that while structured instruction benefited both groups, the AI-enhanced GRR model provided an additional instructional advantage in strengthening learning retention and transfer.

The near-equivalent pre-test scores of the experimental (44.61) and control (44.27) groups indicate comparable starting levels of scientific literacy, supporting the interpretation that later differences were more plausibly associated with the intervention than with baseline disparities. At the same time, the low initial scores reflect the broader challenge of underdeveloped scientific literacy, particularly in students' capacity to explain scientific phenomena, evaluate evidence, and apply knowledge independently (OECD, 2023). This reinforces the instructional relevance of scaffolded models such as GRR, given that scientific literacy is a complex outcome that extends beyond factual recall and requires deliberate pedagogical support (Holbrook & Rannikmäe, 2009; National Research Council, 2012; Norris & Phillips, 2003).

Both groups performed strongly during the “I Do” phase, underscoring the value of teacher modeling in establishing conceptual foundations and making disciplinary reasoning visible, as proposed in the GRR framework (Fisher & Frey, 2021). However, the more meaningful divergence emerged in the “We Do” phase, where the experimental group attained the highest MPS. This suggests that AI was most beneficial when embedded within guided and collaborative learning rather than treated as an independent instructional substitute. One plausible interpretation is that AI functioned as a cognitive scaffold that enabled learners to clarify ideas, test explanations, and refine responses while remaining under teacher mediation. This reading is consistent with recent studies showing that AI improves learning performance most effectively when integrated into coherent instructional structures rather than used in isolation (Wang et al., 2024; Xu & Ouyang, 2022), and with UNESCO’s position that generative AI becomes educationally valuable when it strengthens, rather than replaces, human learning processes (Miao & Holmes, 2023).

The experimental group’s stronger performance during guided practice also aligns with research on AI-supported self-regulated learning. Generative AI has been shown to support planning, monitoring, and reflection when pedagogically structured (Chiu, 2024), and its cognitive and metacognitive benefits appear strongest when learners are required to evaluate and refine AI-generated outputs rather than accept them uncritically (Banihashem et al., 2025). Similarly, meta-analytic evidence suggests that the positive effects of ChatGPT on learning and higher-order thinking are most pronounced when paired with appropriate scaffolds and educational frameworks (Wang & Fan, 2025). The present findings are consistent with these claims, suggesting that AI enhanced learning not by supplying answers, but by deepening guided meaning-making within the GRR sequence.

The slight decline observed in both groups during the “You Do” phase is pedagogically unsurprising, as this stage requires students to perform without the immediate support available during modeling and collaboration. Within the GRR framework, such decline may reflect the normal challenge of transferring understanding from supported to independent application. Nevertheless, the experimental group maintained a modest advantage, suggesting that AI-supported guided practice may have better prepared learners for autonomous performance. This interpretation is consistent with the principle that effective scaffolding should culminate in gradual release rather than prolonged dependence (Fisher & Frey, 2021). At the same time, alternative explanations must be acknowledged, including variation in student confidence, motivation, or familiarity with task demands during independent work.

The post-test results provide the strongest evidence of differential instructional impact. Although both groups scored lower than in some phase-specific assessments, the decline was substantially less pronounced in the experimental group, indicating stronger retention and transfer of learning. Because post-tests typically demand broader and more independent application of knowledge, the superior performance of the experimental group suggests that the AI-enhanced GRR model supported more durable scientific understanding rather than merely improving immediate task performance. This interpretation is consistent with literature defining scientific literacy as the capacity to apply knowledge across contexts, rather than simply reproduce recently taught content (National Research Council, 2012; Norris & Phillips, 2003). Still, this result should be interpreted with caution, as the lower post-test performance of the control group may also reflect broader assessment demands, delayed retrieval difficulty, or task complexity rather than instructional insufficiency alone.

Overall, the findings indicate that both groups benefited from structured teaching, but the AI-enhanced GRR model produced a more robust and sustained pattern of scientific literacy development. Importantly, the results do not suggest that AI replaces effective instruction; rather, they indicate that AI can strengthen guided learning when positioned as a monitored scaffold within a disciplined pedagogical sequence. In this respect, the study contributes to the literature in two ways. First, it extends GRR research by demonstrating that AI can be productively integrated into the guided phase without disrupting the progression from support to independence. Second, it adds to AI-in-education scholarship by showing that the value of AI is best understood not as a stand-alone tool effect, but as a pedagogically mediated contribution to retention, transfer, and disciplinary literacy. For Objective 1, the evidence therefore supports the conclusion that the AI-enhanced GRR model was associated with a stronger pattern of scientific literacy development than instruction without AI support.

Table 2. Paired-Samples t-Test for Pre-test and Post-test Scores of the Experimental Group

The experimental group demonstrated a statistically significant improvement in scientific literacy following exposure to the AI-enhanced GRR model, as indicated by the paired-samples t-test, $t(35) = -7.59$, $p < .001$. This finding suggests that the intervention was associated with meaningful gains in students' ability not only to acquire scientific knowledge, but also to interpret information, apply concepts, construct explanations, and reason with evidence—dimensions central to scientific literacy as conceptualized by Holbrook and Rannikmäe (2009), Norris and Phillips (2003), and the National Research Council (2012). Substantively, the result supports the pedagogical logic of GRR, in which learning is strengthened through a sequenced movement from modeling to guided practice and eventually to more independent performance (Fisher & Frey, 2021). Within this structure, the observed gain suggests that students benefited from sustained scaffolding rather than fragmented or transmission-oriented instruction.

The result also supports the view that AI contributes most effectively to learning when it is pedagogically mediated rather than treated as a stand-alone tool. Consistent with Miao and Holmes (2023), Wang et al. (2024), and Xu and Ouyang (2022), the present finding suggests that ChatGPT functioned as a cognitive scaffold for clarification, elaboration, and feedback within a teacher-guided instructional sequence, rather than as a substitute for student thinking. This interpretation is further aligned with research on AI-supported self-regulated learning, which shows that generative AI can strengthen planning, monitoring, and reflection when learners are guided to engage with it critically and reflectively (Banihashem et al., 2025; Chiu, 2024). It is likewise consistent with Wang and Fan's (2025) meta-analytic conclusion that ChatGPT yields stronger effects on learning and higher-order thinking when embedded in appropriate pedagogical frameworks.

At the same time, the result should be interpreted with caution. The significant pre–post gain does not isolate the effect of AI alone, since improvement likely reflects the combined contribution of teacher modeling, guided practice, collaborative engagement, and independent application. Yet this does not weaken the study's central claim; rather, it reinforces it. The educational value of AI appears to lie not in technological exposure per se, but in its disciplined integration within a scaffolded instructional design. Overall, the finding provides evidence that the AI-enhanced GRR model was associated with meaningful improvement in scientific literacy and supports the broader argument that AI can strengthen science learning when used to extend, rather than replace, structured pedagogical support and learner autonomy.

Table 3. Paired-Samples t-Test for Pre-test and Post-test Scores of the Control Group

The control group likewise demonstrated a statistically significant pretest–posttest improvement in scientific literacy, $t(35) = -3.36$, $p = 0.001$, indicating that structured instruction without AI was sufficient to produce meaningful learning gains. This finding is theoretically important because it reinforces a central position in the scientific literacy literature: development in science learning does not depend on technology alone, but can be advanced through coherent teaching that supports conceptual understanding, evidence-based reasoning, explanation, and contextual application (Holbrook & Rannikmäe, 2009; National Research Council, 2012; Norris & Phillips, 2003). The result is also consistent with the pedagogical assumptions of the Gradual Release of Responsibility framework, in which explicit modeling, guided engagement, and progressive movement toward independence are themselves powerful supports for learning (Fisher & Frey, 2021). In this respect, the control-group gain affirms the continuing instructional value of scaffolded, non-AI-supported teaching, particularly in public-school settings where access to advanced digital tools may be uneven.

At the same time, this finding should be interpreted as evidence of baseline instructional effectiveness rather than as a challenge to the value of the AI-enhanced GRR model. The significance of the control-group gain indicates that both conditions operated within a genuinely functional instructional environment, which strengthens the credibility of the study by showing that student improvement was not confined to the technology-supported condition alone. When considered alongside the stronger pattern observed in the experimental group, however, the result suggests that conventional scaffolding may be beneficial but comparatively less potent in supporting broader retention and transfer. This interpretation is consistent with current scholarship arguing that AI is not inherently transformative, but becomes educationally valuable when it is positioned to augment rather than replace, sound pedagogy (Miao & Holmes, 2023; Wang et al., 2024; Xu & Ouyang, 2022). Thus, the control-group result supports a balanced conclusion: effective science learning remains possible without AI, yet the added scaffolding available in the AI-enhanced GRR model may deepen learning beyond what conventional structured instruction can achieve on its own, particularly for higher-order outcomes such as scientific literacy (Wang & Fan, 2025).

Table 4. Significant difference between Control and Experimental in the Post-test

The significant posttest difference between the experimental and control groups indicates that the AI-enhanced GRR model was associated with stronger scientific literacy outcomes than structured instruction without AI, $t = 2.17$, $p = .038$, with a moderate effect size ($d = 0.51$). Because the posttest functioned as a more integrative measure of retained understanding, conceptual transfer, and independent scientific reasoning, this finding suggests that the advantage of the experimental group extended beyond immediate task performance to a broader and more consolidated form of learning. Such an interpretation is consistent with established conceptions of scientific literacy as the capacity to interpret, evaluate, explain, and apply scientific knowledge meaningfully rather than merely reproduce content (Holbrook & Rannikmäe, 2009; National Research Council, 2012; Norris & Phillips, 2003). The result is likewise aligned with the pedagogical logic of GRR, which positions learning as a sequenced movement from modeling to guided participation and ultimately to independent performance (Fisher & Frey, 2021). In this study, the stronger posttest performance of the experimental group suggests that AI, when embedded within guided learning, may have enhanced students' opportunities to clarify misconceptions, elaborate explanations, and strengthen reasoning before the withdrawal of support. This reading is consistent with scholarship arguing that AI becomes educationally valuable not by its mere presence, but by its pedagogical positioning as a scaffold for human thinking and disciplinary engagement (Miao & Holmes, 2023; Wang et al., 2024; Xu & Ouyang, 2022). It also accords with work on AI-supported self-regulated learning, which suggests that generative AI can strengthen planning, monitoring, and reflective judgment when learners are guided to evaluate and regulate its use critically rather than depend on it uncritically (Banihashem et al., 2025; Chiu, 2024). The moderate effect size further suggests that the observed difference was not only statistically detectable but instructionally meaningful, particularly in relation to higher-order outcomes such as scientific literacy, where gains in reasoning and transfer are often difficult to secure. At the same time, the finding should be interpreted with appropriate caution. The control group also improved significantly, indicating that the comparison was made against an already functional instructional condition rather than ineffective teaching. Accordingly, the result should not be read as evidence that AI replaced sound pedagogy, but rather that AI provided added value when integrated into a disciplined instructional model. In this respect, the finding is consistent with meta-analytic evidence that ChatGPT yields stronger effects on learning and higher-order thinking when used within appropriate scaffolds (Wang & Fan, 2025), while also extending that literature by showing that such benefits may be observed specifically in scientific literacy. Overall, the result supports the conclusion that the value of the AI-enhanced GRR model lies not in the mere inclusion of

AI, but in its bounded and pedagogically regulated use as a temporary scaffold that strengthens the transition from guided participation to independent scientific thinking.

Table 5. Thematic Analysis on the students' perception and experience towards the intervention.

The thematic findings for Table 5 suggest a clear developmental shift in students' perceptions of the intervention, moving from initial curiosity and apprehension to strategic engagement with AI and, ultimately, to greater confidence, reflective awareness, and reduced reliance on technological support. This progression is pedagogically significant because it indicates that students did not experience AI as a static novelty or an unquestioned source of answers, but as a temporary scaffold within a structured movement toward learner autonomy. In this respect, the thematic trajectory is closely aligned with the logic of the Gradual Release of Responsibility (GRR) framework, which positions support as purposeful but diminishing over time in order to strengthen independent performance (Fisher & Frey, 2021).

Prior to implementation, students expressed both optimism and hesitation, reflected in themes of curiosity, excitement, self-doubt, fear of failure, and concern about AI accuracy. Such ambivalence is consistent with current scholarship showing that while students often perceive generative AI as accessible and potentially useful, they also remain concerned about its reliability, ethical use, and potential for overdependence (Akhmetova et al., 2025; Marzano, 2025; Miao & Holmes, 2023). Importantly, this initial caution should not be read merely as resistance. From the perspective of scientific literacy, skepticism toward AI-generated information may itself be educationally productive, since scientific literacy requires learners not only to access information but also to interpret, evaluate, and judge it critically (Holbrook & Rannikmäe, 2009; Norris & Phillips, 2003). The pre-implementation themes therefore suggest that students entered the intervention with an emerging awareness that technologically produced information still required scrutiny.

During implementation, students increasingly described AI as a guide, feedback source, and support for problem solving, while also acknowledging the need to verify outputs, correct inaccuracies, and reinterpret overly complex responses. This pattern suggests that AI was not experienced as a replacement for reasoning, but as a cognitively mediating tool within guided learning. Such a finding is consistent with UNESCO's human-centered view that AI becomes educationally meaningful when it supports agency, judgment, and reflection rather than displacing them (Miao & Holmes, 2023). It also aligns with research on AI-supported self-regulated learning, which indicates that generative AI can enhance planning, monitoring, and reflective thinking when learners are guided to use it critically and strategically (Banihashem et al., 2025; Chiu, 2024). In the present study, students' accounts of checking, correcting, and refining AI outputs suggest that the intervention fostered evaluative and metacognitive engagement rather than passive dependence. This is likewise consistent with findings that AI has stronger educational value when embedded within coherent instructional pathways linking feedback, participation, and knowledge construction (Wang et al., 2024; Xu & Ouyang, 2022).

After implementation, the themes shifted toward autonomy, confidence, and more reflective use of AI, with students reporting greater independence, increased trust in their own reasoning, and reduced reliance on technological assistance. This is particularly important in light of persistent concerns that generative AI may weaken learner agency through overdependence. Instead, the findings suggest that when AI is bounded, monitored, and embedded within a gradual-release structure, students may become more—rather than less—capable of independent thinking. This interpretation is consistent with the view that scientific literacy requires learners to remain the primary agents of interpretation and judgment (Holbrook & Rannikmäe, 2009; Norris & Phillips, 2003), and with evidence that AI supports higher-order learning most effectively when paired with appropriate pedagogical scaffolds (Wang & Fan, 2025). The qualitative progression identified here therefore contributes to the literature by showing that students' perceptions of AI are not fixed, but can evolve from uncertainty to critical, adaptive, and autonomous engagement when the technology is used within a disciplined instructional design. Overall, the findings support the study's central claim that AI contributes most meaningfully to scientific literacy when it functions as a temporary and accountable scaffold for learning, rather than as a substitute for learner thinking.

Table 6. Thematic analysis on students' independent research skills and AI dependence

The thematic findings for Table 6 indicate that students' research-related skills developed progressively across the instructional phases, while their dependence on AI became increasingly limited, critical, and regulated. Rather than using AI as a convenience tool, students appeared to learn how to interrogate, refine, and eventually move beyond AI support in producing scientifically grounded outputs. This developmental pattern is especially significant because it suggests that the AI-enhanced GRR model did not normalize technological dependence; rather, it supported the internalization of research-oriented thinking, disciplinary language, and evidence-based reasoning. In the "We Do" phase, students engaged with AI through collaborative critique, prompt refinement, evaluation of output quality, recognition of inaccuracies, and contextual revision of scientific explanations. These themes suggest that AI functioned as a provisional scaffold rather than an authoritative source, consistent with UNESCO's human-centered view that AI should support learner agency and reflective judgment rather than displace intellectual responsibility (Miao & Holmes, 2023). The pattern also aligns with research on AI-supported self-regulated learning, which emphasizes the value of generative AI for planning, monitoring, and reflection when learners are guided to regulate its use critically (Banihashem et al., 2025; Chiu, 2024). In this respect, students were not passively consuming output; they were evaluating validity, adjusting language, and deciding what to retain, revise, or reject—behaviors that signal emerging research literacy.

The "You Do" phase revealed a more consequential shift toward independent rewriting, conceptual validation, scientific reasoning, and evidence-based explanation without external aid. This suggests that the habits cultivated during guided AI use were carried forward into autonomous performance rather than remaining dependent on technological support. Such a trajectory is closely aligned with the logic of the GRR framework, in which scaffolding is intended to culminate in independent learner performance rather than sustained support (Fisher & Frey, 2021). It is also consistent with broader conceptions of scientific literacy as requiring learners to interpret, explain, and judge knowledge independently rather than rely on prefabricated answers (Holbrook & Rannikmäe, 2009; National Research Council, 2012; Norris & Phillips, 2003). In qualitative terms, the findings further support Wang and Fan's (2025) argument that AI contributes most meaningfully to higher-order learning when embedded within appropriate pedagogical scaffolds. The cross-phase movement from guided critique to self-directed validation therefore suggests that the value of AI in this study lay not in answer generation itself, but in its earlier role in cultivating durable habits of reasoning that remained functional after support was withdrawn.

The findings address an important gap in the literature by tracing how learners move from AI-supported engagement to independent disciplinary performance. Much of the existing scholarship has focused on performance gains, user perceptions, or technological affordances; fewer studies have examined how AI use can be pedagogically bounded so that it ultimately reduces, rather than reinforces, uncritical dependence. The present results suggest that students did not reject AI after the intervention, but repositioned it as a limited learning partner—useful for clarification and refinement, yet subordinate to their own judgment. This is a more educationally meaningful form of AI use, as it aligns with current concerns that students should be taught not only how to use AI, but how to use it selectively, responsibly, and critically. Overall, the thematic findings provide strong qualitative support for the study's central claim: the AI-enhanced GRR model appears to strengthen scientific literacy while also fostering independent research-related skills and a more disciplined, less dependent relationship with AI.

Conclusions

This study concludes that the AI-enhanced GRR model was effective in strengthening students' scientific literacy across the teaching and learning phases. In response to Research Question 1, both the experimental and control groups began with low and comparable pre-test performance, improved substantially during the "I Do," "We Do," and "You Do" phases, and then displayed different levels of retention in the post-test. The experimental group consistently achieved higher mean percentage scores than the control group after baseline and maintained a notably stronger post-test performance. This indicates that the AI-enhanced GRR model supported a more robust pattern of scientific literacy development, particularly during guided learning and in the later transfer of learning to more independent assessment conditions.

In response to Research Question 2, the findings confirm that there was a significant pre-test to post-test improvement in the experimental group, showing that students exposed to the AI-enhanced GRR model developed stronger scientific literacy after the intervention. At the same time, there was also a significant pre-test to post-test improvement in the control group, indicating that structured instruction without AI can still meaningfully enhance scientific literacy. However, the magnitude of improvement was stronger in the experimental group, suggesting that the inclusion of AI within a disciplined gradual-release structure provided an additional instructional advantage beyond conventional scaffolded teaching alone.

For Research Question 3, the study concludes that there was a significant difference between the experimental and control groups in the post-test, with the experimental group outperforming the control group after the intervention. This demonstrates that the AI-enhanced GRR model was more effective than the non-AI condition in improving students' scientific literacy. The difference was not only statistically significant but also educationally meaningful, indicating that AI, when embedded within a carefully sequenced pedagogical framework, can contribute to stronger retention, transfer, and independent use of scientific reasoning.

In relation to Research Question 4, the qualitative findings show that students' perceptions and experiences toward the intervention developed progressively across phases. Before implementation, students expressed curiosity, excitement, hesitation, and concern about AI accuracy. During implementation, they viewed AI as a scaffold for clarification, feedback, and guided participation, while also recognizing the need to verify and refine its outputs. After implementation, students reported greater self-confidence, clearer awareness of AI's limits, and reduced reliance on it. These findings indicate that students did not experience AI merely as a technological novelty, but as a temporary and accountable learning support that became less central as their own competence increased.

Finally, in response to Research Question 5, the study concludes that students demonstrated growing independent research skills and reduced AI dependence over the course of the intervention. During the "We Do" phase, students refined prompts, evaluated AI outputs, identified inaccuracies, revised explanations, and contextualized scientific ideas, showing guided critical engagement with AI. During the "You Do" phase, they were able to rewrite independently, validate concepts on their own, construct evidence-based explanations, and apply scientific reasoning without external aid. Across phases, students internalized research-oriented thinking, scientific tone, and evaluative judgment, indicating that the AI-enhanced GRR model fostered not only better performance but also greater autonomy, discernment, and responsible use of AI.

Overall, the study affirms that the AI-enhanced GRR model can serve as an effective pedagogical approach for accelerating scientific literacy while preserving the core educational goal of learner independence. The findings show that AI is most valuable not when it replaces student thinking, but when it is used as a temporary scaffold within a structured instructional design that guides learners from supported participation toward confident and independent scientific reasoning.

Recommendations

Based on the findings of the study, it is recommended that science teachers consider adopting the AI-enhanced GRR model as a structured instructional approach for improving students' scientific literacy. The results indicate that the combination of explicit teacher modeling, guided collaborative work, monitored AI use, and independent application can strengthen students' conceptual understanding, scientific reasoning, and transfer of learning. However, AI integration should remain pedagogically bounded and teacher-regulated. Its use should be framed as a temporary cognitive scaffold for clarification, feedback, and refinement rather than as a substitute for student thinking or scientific explanation.

School leaders and curriculum implementers may also consider supporting the responsible classroom integration of AI through teacher training, instructional guidelines, and monitoring mechanisms. Since the study showed that students benefited most when AI was used within a disciplined pedagogical sequence, professional development programs may focus on prompt design, verification strategies, critical evaluation of AI outputs, and gradual withdrawal of support. In this way, schools can maximize the learning benefits of AI while minimizing the risk of overdependence, misinformation, or passive use.

For classroom practice, it is further recommended that teachers integrate reflection tasks, AI declaration tools, and independent performance activities whenever AI is used in science instruction. The qualitative findings suggest that students became more critical, reflective, and autonomous when they were required to document, examine, and revise AI-supported outputs. Thus, future instructional applications should preserve opportunities for students to validate concepts independently, justify claims with evidence, and demonstrate scientific reasoning without technological assistance.

Future research is strongly recommended to test the applicability of the AI-enhanced GRR model in other contexts. Since the present study was limited to Grade 12 STEM students in one public senior high school and focused on General Physics 2, subsequent studies may examine the model using larger samples, multiple schools, other grade levels, and different subject areas, particularly in biology, chemistry, earth science, and non-STEM contexts. Replication in varied settings would help determine the generalizability of the findings and clarify whether the observed benefits remain consistent across disciplines and learner populations.

Further studies may also employ more rigorous or alternative research designs, such as randomized or matched-group designs, longitudinal follow-up, or multi-site mixed-methods investigations. Such approaches may help determine whether the gains observed in scientific literacy are sustained over time and whether the same pattern of reduced AI dependence emerges across longer instructional periods. In particular, the relatively sharper decline of the control group in the post-test, compared with the stronger retention of the experimental group, suggests the need for further investigation into the role of AI-supported scaffolding in long-term retention, transfer of learning, and independent reasoning.

In addition, future researchers may explore other variables related to the present study, such as metacognitive regulation, motivation, cognitive load, research self-efficacy, ethical AI use, and quality of prompt construction. Since the current findings suggest that students' perceptions and research-related behaviors changed over time, further inquiry may examine which specific aspects of AI-supported instruction most strongly contribute to learner autonomy. Comparative studies using different generative AI tools may also be conducted to determine whether the observed effects are specific to ChatGPT or extend to other AI systems.

The findings of the study support the continued but cautious use of AI in education. Future applications and research should therefore focus not merely on whether AI is present in instruction, but on how it is pedagogically designed, ethically regulated, and gradually withdrawn so that students become more scientifically literate, more critical, and more independent learners.

Compliance with Ethical Standards

This study was conducted in accordance with accepted ethical standards for educational research. Informed consent from parents or guardians and student assent were obtained prior to data collection. Participation was voluntary, and respondents were informed of their right to withdraw at any stage without penalty. Anonymity and confidentiality were maintained throughout the study, and all data were handled in accordance with Data Privacy principles. The well-being of the respondents was safeguarded, and no participant was exposed to coercion, harm, or academic disadvantage. The author declares that no conflict of interest influenced the conduct, analysis, or reporting of the study. Plagiarism was strictly avoided, all sources were properly acknowledged, and the findings were interpreted and reported honestly, without fabrication, falsification, or intentional bias. The results were used solely for research purposes. For full disclosure, ChatGPT was used only as part of the AI-enhanced GRR model under clearly defined ethical and pedagogical controls. Students were oriented on responsible AI use, required to verify and revise AI-supported outputs, and monitored through structured documentation. No personally identifiable participant data were entered into any AI system. Full responsibility for the accuracy, originality, and integrity of the manuscript remains with the author.

REFERENCES

Akhmetova, A. I., Sovetkanova, D. M., Komekbayeva, L. K., Abdrakhmanov, A. E., Yessenuly, D., & Serikova, O. S. (2025). A systematic review of artificial intelligence in high school STEM education research. Eurasia Journal of Mathematics, Science and Technology Education, 21(4), em2623. https://doi.org/10.29333/ejmste/16222

Banihashem, S. K., Bond, M., Bergdahl, N., Khosravi, H., & Noroozi, O. (2025). A systematic mapping review at the intersection of artificial intelligence and self-regulated learning. International Journal of Educational Technology in Higher Education, 22, 50. https://doi.org/10.1186/s41239-025-00548-8

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40. http://doi/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. http://doi/10.1191/1478088706qp063oa

Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. Educational Technology Research and Development, 72, 2401–2416. https://doi.org/10.1007/s11423-024-10366-w

Creswell, J. W., & Plano Clark, V. L. (2017). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

Department of Education. (2017). DepEd Order No. 16, s. 2017: Research management guidelines.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. Health Services Research, 48(6 Pt 2), 2134–2156. http://doi/10.1111/1475-6773.12117

Fisher, D., & Frey, N. (2021). Better learning through structured teaching: A framework for the gradual release of responsibility (3rd ed.). ASCD.

Holbrook, J., & Rannikmäe, M. (2009). The meaning of scientific literacy. International Journal of Environmental & Science Education, 4(3), 275–288.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. Psychometrika, 2(3), 151–160. http://doi/10.1007/BF02288391

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Marzano, D. (2025). Generative artificial intelligence (GAI) in teaching and learning processes at the K-12 level: A systematic review. Technology, Knowledge and Learning. Advance online publication. https://doi.org/10.1007/s10758-025-09853-7

Mertler, C. A. (2024). Action research: Improving schools and empowering educators (7th ed.). SAGE Publications.

Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO.

National Research Council. (2012). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. The National Academies Press.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. Science Education, 87(2), 224–240. https://doi.org/10.1002/sce.10066

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16, 1–13. http://doi/10.1177/1609406917733847

OECD. (2023). PISA 2022 results (Volume I and II) country note: Philippines. OECD Publishing.

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

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. doi:10.1093/biomet/52.3-4.591.

The jamovi project. (2024). *jamovi* (Version 2.4).

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. Humanities and Social Sciences Communications, 12, Article 621. https://doi.org/10.1057/s41599-025-04787-y

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. Expert Systems with Applications, 252, Article 124167. https://doi.org/10.1016/j.eswa.2024.124167

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. International Journal of STEM Education, 9, Article 59. https://doi.org/10.1186/s40594-022-00377-5

APA Citation:

Tanzo, J. C. C. (2026). FROM GUIDED TO INDEPENDENT: AI-ENHANCED GRADUAL RELEASE OF RESPONSIBILITY (GRR) MODEL TO ACCELERATE SCIENTIFIC LITERACY. Ignatian International Journal for Multidisciplinary Research, 4(3), 1744–1766. https://doi.org/10.5281/zenodo.19383410

johncarlo.tanzo@gmail.com