International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

Evaluating the ChatGPT-Assisted Learning Method (CALM) for Struggling Readers: The Emergence of the Theory of Distributed Reading Gains

Romeo A. Abuhan Jr.
University of Perpetual Help Dalta System

Publication Date: 06-05-2026

ABSTRACT

This study aimed to determine the efficacy of ChatGPT-Assisted Learning Method (CALM) in enhancing reading proficiency among 9th grade students in a public secondary school situated in the Philippines. The study used a sequential explanatory mixed-methods, single-group pretest–posttest design. A total of $N = 50$ students were purposely selected who were identified as frustrated readers during the Phil-IRI assessment. The intervention period lasted eight weeks and was assessed using a pretest and posttest. To gain a deeper understanding of the participants' significant experiences, this was followed by a qualitative phase, in which five participants were selected as the saturated sub-sample to undergo one-on-one semi-structured interviews and focus groups. Thematic analysis was used to isolate the functional mechanisms of the AI-mediated scaffold. Results from the quantitative data revealed an increase in reading proficiency, with $M = 5.66$, and students reaching "Proficient" or "Advanced" levels rising from 54% to 92%. Cohen's $d = 1.45$ indicates a large effect, while $\eta^2 = .34$ suggests that 34% of the variance in reading proficiency is attributable to the CALM intervention. Hence, a reduction in the mean standard deviation from 4.13 to 2.82 indicates a significant improvement in reading performance. Furthermore, five themes emerged from the qualitative findings supporting the result from the quantitative phase which described as core functional mechanisms: (1) instructional responsiveness and mastery growth; (2) dialogic adaptability and strategic autonomy; (3) diagnostic feedback and self-regulated revision; (4) metacognitive prompting skill building; and (5) ownership and clarity through reconstructed expression, marking the internalization of the digital scaffold. This evidence establishes the Theory of Distributed Reading Gains (TDRG) through the CALM Feedback-Prompt-Scaffold (FPS) Instructional Framework cycle which operationalizes how distributed gains arise in resource-constrained high-density environments.

Keywords: *Educational Psychology, Reading Proficiency, Sequential Explanatory Mixed-Methods, Theory of Distributed Reading Gains (TDRG), ChatGPT-Assisted Learning Method (CALM), Philippines*

Received: 03-02-2026 **Revised:** 03-23-2026 **Accepted:** 05-30- 2026 **Published:** 06-05-2026

ISSN:3116-2991

Introduction

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.minduraresearch.com/journal/index

The growing global literacy gap is among the most concerning contemporary educational challenges. Despite being one of the best predictors of academic success, reading proficiency continues to decline (OECD, 2022; World Bank, 2024). In the context of the Philippines, this crisis is aggravated by systemic failures for instructional scaling (Acido & Caballes, 2024). According to a study by Basali (2024), students in grade 9 who find complex texts daunting receive insufficient support. If not targeted, these readers these readers remain mired in non-fluent decoding and miss out on higher-level comprehension.

Yet these systemic interventions can be misaligned with the unique linguistic needs of struggling readers (e.g., quality of instruction prior to reading interventions). In order to close these gaps, this study uses the ChatGPT-Assisted Learning Method (CALM) framework. Also, in accordance with global competence frameworks (UNESCO, 2020), CALM considers AI to be a 'pedagogical partner', a facilitator for the students' ZPD (the theoretical range of how much more learners can do with assistance from competent others), monitored in real-time (Cai et al., 2025; Vygotsky, 1978).

This requires a radical shift towards AI-mediated solutions that regulate cognitive load. There should be adequate scaling in reading tasks; otherwise, cognitive fatigue will occur according to Cognitive Load Theory (Sweller, 1988). The CALM framework (Trajkovski & Hayes, 2025) helps students address extraneous load by providing immediate feedback during complex decoding tasks.

Grounding the intervention in those concepts leads to the introduction of the Theory of Distributed Reading Gains (TDRG). In the high-density, traditional classroom model, teachers often face what is known as the MKO Paradox—the inability to scaffold 50 students simultaneously. CALM takes on this challenge by defining Generative AI as a "Digital More Knowledgeable Other" (MKO), thereby extending the teacher beyond her physical location. In TDRG, the teacher's role evolves into a "Calibration Expert," confirming student-AI interactions through the Feedback-Prompt-Scaffold (FPS) cycle. This system handles heavy tasks, freeing working memory for synthesis and deep understanding. Therefore, this study aims to determine the effectiveness of the CALM framework on positive reading proficiency outcomes among Grade 9 students and establishes the TDRG as a novel theoretical framework for literacy through artificial intelligence integration.

Statement of the Problem

This study primarily examined and explained the effectiveness of a CALM in improving the reading proficiency of Grade 9 students at a public secondary school in the Philippines. Specifically, the study sought to answer the following questions:

1. What are the reading proficiency levels of the Grade 9 respondents before the CALM intervention?
2. What are the reading proficiency levels of the respondents after the CALM intervention?
3. What is the effect size of the significance of the intervention?
4. Is there a significant difference between the reading proficiency levels in the pretest and posttest?
5. What themes emerged from the students' narratives regarding the use of the CALM intervention?
6. Based on the findings of the study, what output may be proposed?

Methodology

Research Design

This study used an explanatory sequential mixed-methods design (Creswell & Creswell, 2018). Quantitative data were collected and analyzed first, followed by a qualitative phase that would shed light, elucidate, and deepen understanding of the earlier findings. Specifically, this study adopted a single-group pretest-posttest design to evaluate the effectiveness of the CALM on Grade 9 students' reading proficiency during an 8-week intervention period. Descriptive statistics (means and standard deviations) were used to describe literacy levels, and inferential statistics (paired t-tests) were used to assess the significance of the improvements relative to both time points.

The qualitative phase was conducted to explain and elaborate in depth on the quantitative results. Thematic Analysis (Braun & Clarke, 2006) was used by the researcher to identify and report the emerging themes from semi-structured interviews and

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

Focus Group Discussions (FGDs). In this phase, the researcher focused on layers of students' experience of CALM during FPS cycles. Objectively, the researcher proposes a TDRG that works because integrating these two datasets enabled robust empirical triangulation on subjective and objective performance across all domains in reading proficiency.

Population and Sampling Technique

Three hundred forty-nine (349) Grade 9 students who officially registered for School Year 2025–2026 composed the study population. Purposive sampling was used to obtain a target sample consisting of fifty (50) participants who: (1) were students of the English class, and (2) had mobile devices with internet connectivity. A non-probability testing strategy was employed to identify students for whom literacy profiles that fall under CALM evaluative conditions would be achieved. In recruiting participants, the researcher selected a narrower demographic with specific literacy deficits, allowing closer exposure to the intervention and sacrificing broad statistical generalizability for in-depth inquiry. To promote equitable participation and help close the digital divide, the school provided students lacking adequate access with additional tablets and stable internet access at school. This ensures that choices are based on educational need rather than socio-economic ability to purchase.

Respondents of the Study

This study's respondents were fifty (50) Grade 9 learners of a public secondary school in the Philippines, or "struggling readers" from among the "Frustration Level" based on the Philippine Informal Reading Inventory (Phil-IRI). These participants were purposefully selected based on their pedagogical needs for reading intervention, the digital tools necessary for a CALM intervention, and the schools' ability to provide equal access for all.

Research Instruments

This study utilized a twenty (20) item multiple-choice instrument, researcher-developed and aligned with the Phil-IRI but contextualized to measure the target proficiency levels: (1) Vocabulary building; 2) Questioning-and-answering; 3) Generating summary, and 4) Interactive reading. To gain richer insights, semi-structured interviews and FGDs were conducted during AI-mediated sessions and afterward.

Validity and Reliability of the Instrument

The instrument's content validation was conducted by experts: two professors from the University of Perpetual Help System DALTA and a master teacher in English at Molugan National High School. Using a 5-point Likert scale, the content validity score was $M = 4.82$, indicating high relevance and clarity (Lawshe, 1975). Refinement of these items was also carried out by 35 non-respondent student-participants in a pilot study, and their internal consistency, as measured by Cronbach's Alpha ($\alpha = .85$), indicated excellent psychometric reliability and excellent internal consistency in measuring the target proficiency domains. (Tabachnick & Fidell, 2019).

Scoring Procedure

To simplify reporting of student reading proficiency, a four-level categorization system was utilized. Advanced, Proficient, Basic (Developing), and Below Basic (Beginning) were adapted from the National Assessment Governing Board (2021) framework. The actual numerical ranges (16.00–20.00, 11.00–15.99, 6.00–10.99 and 1.00–5.99) were defined by the researcher to fit with the study's 20-item assessment instrument

Data Gathering Procedure

The study secured formal approval from the school principal and division office to follow institutional protocols and ethical standards. Consent and assent forms were provided and obtained from parents or guardians and students, respectively. Data were kept confidential, and participants were treated respectfully. The data used in the quantitative phase were gathered through a pretest, followed by 8 weeks of CALM intervention. A posttest was given to measure the literacy gains. To further explore the significant experiences, semi-structured interviews and FGDs were conducted. All data were monitored, organized, computed, and analyzed using a mixed-methods approach that integrated t-tests with thematic coding to examine the study's central research questions.

Statistical Treatment

A sequential explanatory mixed-methods design was applied to provide an overview of the quantitative and qualitative results of this study. Descriptive statistics (mean, percentage distribution, and standard deviation) described central tendency,

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

frequencies by reading level, and score variability. A paired-samples t-test was used to determine the effect of the intervention. Cohen's (*d*) and Eta Squared (η^2) were utilized to calculate effect size and the variance explained by treatments, respectively. Cronbach's Alpha (0.85) confirmed the instrument's reliability, indicating internal consistency. Qualitative data were thematically analyzed to make sense of students' salient experiences with the CALM intervention.

Qualitative Data Analysis

Thematic analysis (Braun & Clarke, 2006) was used to contextualize quantitative findings through a systematic process that included verbatim transcription, initial coding of participants' significant experiences, and identification of defining patterns. Emerging themes were synthesized from pooled categories and validated through triangulation with quantitative trends. To strengthen the trustworthiness and credibility of the data, member checking and an audit trail were documented to confirm qualitative rigor.

Ethical Considerations

The study was performed in accordance with institutional ethical standards and with formal written approval, as well as informed consent from parents and assent from students. In line with the principles of beneficence, non-maleficence, autonomy, and justice, the protective aspects of delivery were proactive, respecting participants' welfare whilst respecting their right to withdraw without penalty. Alphanumeric codes and pseudonyms protected confidentiality, while digital data were stored in encrypted, password-protected files that would be destroyed after publication. Maintaining transparency regarding parent-teacher consultations and sharing in relatable formats ensured that the benefits of the intervention extended to the school community.

Results and Discussion

The Reading Proficiency Levels of the Grade 9 Respondents Before the CALM Intervention were Determined

Table 1

Frequency, Percentage, and Mean Distribution of Respondents' Reading Proficiency Before the Interventions

Range	Interpretation	Pretest	
		F	%
16.00 - 20.00	Advanced	2	4%
11.00 - 15.99	Proficient	13	26%
6.00 - 10.99	Developing	27	54%
1.00 - 5.99	Beginning	8	16%
Total		**50**	**100%**
Mean		**10.28**	
Interpretation		**Developing**	
SD		**4.13**	

Note. N = 50. F = Frequency; SD = Standard Deviation; M = Mean. The proficiency levels and scoring ranges are adapted from the UNESCO (2020) performance standards for secondary literacy.

Table 1 shows the frequencies, percentages, and mean scores of respondents' performance levels prior to the intervention. The largest group comprised those classified as "Developing" (*N* = 27, 54%), but there were also respondents who displayed proficiency (*N* = 13, 26%). The distribution of respondents was 8 (16%) in the "Beginning" category; 2 (4%) in the "Advanced" category. The computed *M* = 10.28 (total *SD* = 4.13). As the mean fell within the developing range (6.00-10.99), one could conclude that prior to intervention, the average reader was at a lower level. According to Abril et al. (2022), there is a significant reading proficiency gap among High School Grade 9 students in the Philippines, as indicated by the Phil-IRI assessment results.

The Reading Proficiency Levels of the Grade 9 Respondents After the CALM Intervention were Determined

After the CALM intervention, many respondents had scores above their baseline reading assessment levels, as shown in Table 2. Fifty percent (50%) of the participants advanced to the "Advanced" level. Roughly, two-fifths (42%) read at a "Proficient" reading level, and eight percent (8%) remained "Developing." None of the respondents was at the beginner level. Consequently, the CALM intervention resulted in a statistically significant improvement in reading performance, with an average posttest reading assessment score of 15.94 (*SD* = 2.82), indicating proficient reading. Such quantitative improvements align with previous CALM impacts observed in secondary settings, where increases in literacy performance have also been observed. (Alamo et al., 2024).

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

Table 2

Frequency, Percentage, and Mean Distribution of Respondents' Reading Proficiency After the Intervention

Range	Interpretation	Pretest	
		F	%
16.00 - 20.00	Advanced	25	50%
11.00 - 15.99	Proficient	21	42%
6.00 - 10.99	Developing	4	8%
1.00 - 5.99	Beginning	0	0%
Total		**50**	**100%**
Mean		**15.94**	
Interpretation		**Proficient**	
SD		**2.82**	

Note. N = 50. F = Frequency; SD = Standard Deviation; M = Mean. *The proficiency levels and scoring ranges are adapted from the UNESCO (2020) performance standards for secondary literacy.*

The Effect Size of the Intervention was Computed to Determine its Significance

Table 3

Effect Size Significance of the Reading Intervention

Effect Size Measure	Value	Interpretation
Cohen's (d)	1.45	Large effect
Eta squared (η^2)	0.34	Large effect

Note. d interpretation: 0.20 = small, 0.50 = moderate, 0.80 = large (Cohen, 1988). η^2 was estimated from d using the conversion formula: $\eta^2 = d^2 / d + 4$.

Table 3 shows that the values calculated at d = 1.45, which was larger than 0.80, indicated a "large effect" size. Thus, the gain in reading proficiency between pretest and posttest is both statistically and educationally significant, paralleling evidence for the practical importance of large effect sizes when assessing educational interventions (Kraft, 2020). Thus, the intervention accounts for $\eta 2$ = 0.34 or approximately 34% of the variance in participants' reading performance scores. The CALM adopted scaffolding theories within a cyclic framework of FPS, allowing teachers to make on-the-fly modifications based on their perceptions of students' performance.

Measure	Mean	N	SD	SEM
Pretest	10.28	50	4.13	0.58
Posttest	15.94	50	2.82	0.40

These swift gains were rooted in alignment with the science of reading and in explicit instruction in decoding and language comprehension (Petscher et al., 2020). According to Dahl-Leonard et al. (2025), these effect sizes indicate that CALM is a replicable and scalable framework that far surpasses the small effects seen in most short-term literacy interventions.

The Significant Difference Between the Reading Proficiency Levels in the Pretest and Posttest was Determined

Table 4

Paired Samples Statistics

As seen in Table 4, there is a significant increase from pretest M = 10.28, SD = 4.13 to posttest M = 15.94, SD =2.82. The decrease in the SD and SEM from 0.58 to 0.40 indicates that post-intervention scores were more homogeneous and more accurate when compared. These 5.66-point gains demonstrate the effectiveness of structured scaffolding and AI-mediated feedback in closing literacy gaps. The literature suggests that CALM can support learning more effectively (El Hassan & Alsalwah, 2025) by aiding students in comprehending chapters from their reading materials, increasing vocabulary acquisition, and fostering greater engagement with the text.

Table 4.1

Paired Samples Correlation

N	Mean	Sig.
50	0.400	0.004

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

Table 4.1 shows a paired-samples correlation of 0.400 (N = 50), p (.004). This correlation (p < .05) suggests a moderately high positive relationship between the pretest and posttest scores. Hence, this further strengthens the fact that, although the CALM intervention did result in an increase in mean scores for students across all participants, their overall rank or standing remained stable, thus showcasing how effective an intervention it can be, irrespective of where students stood proficiency-wise at baseline.

Table 4.2

Paired Samples t-Test

Mean Diff	SD	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)
5.66	3.66	4.53	6.79	10.10	49	0.000

Table 4.2 presents the results of the paired-samples t-test, which assesses whether CALM had a statistically significant effect. Such an effect constitutes a large increase on average (M = 5.66, SD = 3.66) in the reading proficiency from pretest to posttest. This yields t = 10.10 (df = 49) and p =. 000 denoting statistically improvement. Additionally, with a 95% Confidence Interval (CI = 4.53 to 6.79), further validating that the improvement achieved is not due to chance alone but rather to structured scaffolding of learning and feedback via AI mediation. Bautista (2025), along with Relyea et al. (2024), demonstrated similar outcomes in reading intervention contexts, with moderate pretest–posttest correlations suggesting quite stable structures underlying the constructs. This means targeted skills-focused instruction leads to measurable gains without affecting students' relative ranking.

Themes Emerged from the Students' Narratives Regarding Their Experiences with the CALM Intervention

Theme 1: Instructional Responsiveness and Mastery Growth

The first emergent theme is that the prompt feedback provided by CALM led to targeted responses and created a cycle of refinement where students were forced to evaluate mistakes and retest strategies almost instantaneously. Adaptive design achieved dramatic improvements in reading fluency, comprehension, and confidence. Results from the semi-structured and all FGD interviews corroborate the qualitative improvement in reading proficiency. One participant summarized, *"We only have a 6/10 score... we improved... to 8/10"*; and another confirmed: *"The best is the way ChatGPT gives advice... we can easily understand what part we should improve"* (P2). According to Röhl et al. (2025), CALM acted as an assessment of learning and offered criterion-referenced feedback on identified literacy gaps that could drive growth across iterations. Through triangulating these data, responsive growth in two key areas surfaced: improvement in instruction and mastery.

Theme 2: Dialogic Adaptability and Strategic Autonomy

Another emergent theme was the realization that CALM's prompt-based interaction framework features a fluid repertoire of exchanges that dynamically oscillate to support students' needs, facilitating continued cognitive momentum and avenues for critical thinking. And these improvements in reading-gain quantification have been accompanied by qualitative data to add productivity capabilities. As one participant noted in terms of the interactive process involved with ChatGPT, *"We just type a prompt, and ChatGPT gives the answer right away"* (P2); another added, *"...if my question is not clear, I rephrase it and ChatGPT answers again immediately"* (P3). Numerous participants described their involvement in prompt refinement, a type of metacognition and help-seeking behavior (Lee & Palmer, 2025; Xiao et al., 2025) that was also corroborated by observations during the interviews. As emphasized by Cai et al. (2025), who created a "dynamic" zone of proximal development (ZPD) aligned with the Vygotskian scaffolding theory, this reinvigorated them to provide more help and transfer strategies, as well as to learn strategically (Lineman et al., 2025). Data triangulated with learning experience data indicated that dialogic adaptability and strategic autonomy were prominent predictors of sequential success.

Theme 3: Diagnostic Feedback and Self-Regulated Revision

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.minduraresearch.com/journal/index

Moreover, the third theme showed that CALM also served as an adaptive diagnostic tool, explicitly pinpointing weaknesses and furnishing strategies for incremental revision of work output, thereby strengthening both cognitive and metacognitive skills. The quantitative gains in reading ability underpinned qualitative insights about the importance of corrective feedback. One respondent indicated: "*...sometimes ChatGPT gives the wrong answer, but it tells me why and how to fix*" (P4), and another said: "*...it points out mistakes and suggests how to make it better*" (P5). Students were seen revising through an iterative process that moved their thoughts around, and a second student who applied AI's suggestions to his reading activity all by himself, which should be a sign of having taken steps towards self-directed learning and setting internal standards for quality. From a pedagogical perspective (Kaldaras et al., 2024). These narratives are similar to formative assessment and scaffolding; personalized explanatory feedback promotes deep learning, self-regulation, and the transfer of skills.

Theme 4: Meta-cognitive Prompting Skill-Building

Building on this issue, the fourth theme illustrated how CALM enhanced students' meta-comprehension or awareness of what they already knew and what needed to be tackled with targeted prompts, thereby affording them much more agency over their thinking, cognition, and strategy (the strategic authority). The qualitative data showed how prompts guided students' understanding of content. As one participant reported, "*Whatever information that ChatGPT gives, we should ask the right prompts to get the right information*" (P3), and another explained: "*...we prompted 'What do these words in simple terms'*? (P4). These demonstrations characterized a "detect-act loop" in which students would slow down their reading, or paraphrase the prompts as self-help from confusion, indicating Self-Regulated Learning (SRL) through cognitive monitoring and strategic management (Shi et al., 2025). This confirms the Cognitive Load Theory (CLT) by Sweller (1988). In this case, students are encouraged to reflect on their own learning, fostering habits of metacognition that help them become more independent and successful. Consequently, a significant outcome of CALM was the development of students' metacognitive questioning skills, leading to increased independence and confidence in literacy. More than building prompts, this metacognitive maturity signals a natural transition into a sense of expressive agency, which leads us towards our fifth theme.

Theme 5: Ownership and Clarity through Reconstructed Expression

Therefore, the fifth theme indicated that CALM promotes active meaning-making by enabling students to articulate concepts in varied formats and revise AI-generated outputs to support clarity, originality, and self-confidence. Qualitative data indicated that participants saw CALM both as a way to write more clearly. As one informant said, "*...we put fixed grammar, we picked our words, spelling, and also we ordered our summary*" (P1). Another mentioned, "*...we requested ChatGPT to explain the meaning of words in simple ways... Then, we created our own sentences*" (P4). These narratives validate that the constructivist learning model identified by Vygotsky (1978) whereby knowledge is constructed through active involvement and continual iteration—was mobilized, leading students to reorganize their thinking and vocabulary after several rounds of feedback (Owen, 2025). From a pedagogical standpoint, this approach is similar to student-centered theory: students can customize the output while maintaining control over the product.

Conclusion

The findings from this study empirically suggest that ChatGPT-Assisted Learning Method (CALM) as an instructional intervention is statistically and educationally effective for frustrated Grade 9 readers. It is evident that the "Developing" baseline (M = 10.28) transitioned to a "Proficient" post-intervention level (M = 15.94), and with such robust support at a large effect size of d = 1.45, it can be concluded that this framework bridges the literacy gap. Moreover, this confirms that the 5.66-point gain from deliberate scaffolding is not a random occurrence but is equally distributed among all participants (p = .000).

Furthermore, the findings lay the groundwork for the Theory of Distributed Reading Gains (TDRG). The students subsequently redirect their scarce cognitive energy (which was primarily devoted to content production through linguistic decoding in a low-level, AI-mediated intervention) towards higher-order

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

thinking and evaluative understanding, which further supports the study by drawing on a high-density, limited-resources setting. The reduction in standard deviation from (*SD* = 4.13 to 2.82) suggests that post-intervention, over this population cohort, reading performance was homogenized and the achievement gap narrowed for those reporting opportunities for short activities or supported engagement with books through other means.

Qualitative evidence supports triangulating the quantitative gains by further showing that the CALM intervention fosters strategic autonomy. The "Feedback-Prompt-Scaffold" (FPS) cycles required learners to develop a "detect-act loop," which enabled them to shift from passive reading to performative, self-regulated learning (SRL). This allows them to become increasingly meta-cognitive, prompting a by-product of AI-mediated instruction that moves them from "Beginning" readers to "Advanced".

Finally, the study presents 'Calibration Expert' as a scalable approach to managing a resource-constrained environment. It effectively removes the teacher as a direct content provider and rather positions them as more of an AI-student interaction manager. Such an organizational pivot indeed allows "dialogic adaptability," and the cancellation of reading strategies so that "instructional responsiveness" can endure at high-pressure educational sites where enacting mastery in reading is paramount.

Recommendations

Based on the empirical evidence of the CALM-FPS Instructional Framework and the identified strengths from the collected cohort data for Grade 9, a series of strategic responses is recommended to address what is now considered a systemic literacy crisis:

For Educational Practitioners: The Transition to Calibration Expertise

Establishing the CALM-FPS Framework as the default remedial modality that school professionals have with students and remedies, and which requires leaving behind content as a central pedagogic consideration to Calibration Expertise, where qualitative auditing of student-AI exchanges takes priority over quantitative logging for reward-based output evaluation. This means that prompts must be in the learner's ZPD, and as metacognition becomes internalized, so too should digitally scaffolds.

For School Administrators: Setting up AI-Aware Literacy Hubs

It is advisable for administrators to make this their top priority as AI-Aware Literacy Hubs by pulling the data across the domains. It encompasses a targeted professional development program focused on "Pedagogical Prompt Engineering" for master teachers, as well as the procurement of necessary digital hardware. A key piece of these resources will be the centralization around equity, so directives to give access to high-density, low-resource learners are equalized and directed to combat and close the digital divide.

For Technical Directors: "Offline-First" AI Infrastructure

For such geographically challenged scenarios, executive technologists should assess the viability of Offline-First AI Infrastructure. This means running lightweight locally hosted LLMs on school servers. And so, this sort of local approach makes it possible for interactive reading sessions to continue without the delay or cost of a round-the-clock Internet connection.

For Curriculum Developers: Proficiency-Based AI-Mediated Integration

A secondary reading curriculum could help students work with an educator through a march toward mastery structured in increasingly high expectations. In this logic of stratification via systematized, gamified strata, cognitive load is managed through practice milestones. By aligning with these increasingly sophisticated cognitive demands, designers left open a roadmap towards that knowledge trajectory, showing how the designer can articulate and help delineate for learners the pathway they will need to take as it relates to fact gathering, understanding, and knowledge acquisition through evaluative synthesis.

For National Policymakers: Adoption of Distributed Reading Gains (TDRG)

The TDRG as a conceptual basis for literacy strategies at the national level. These advances deserve a budgetary spending bonanza for AI-augmented pedagogy. Utilize AI- mediated intervention to establish scalable national programs: Divide cognitive labor between learning agents and AI-mediated; scale national programs to bridge achievement gaps in high-density public education systems.

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.mindumaresearch.com/journal/index

Ethical Statement

Ethical clearance for this study was obtained from the University of Perpetual Help System–DALTA Research Ethics Committee, in accordance with the Declaration of Helsinki. As previously described, prior to the introduction of the CALM-FPS Framework, formal informed consent was obtained from the legal guardians of all participants, and assent was obtained from the $N = 50$ cohort. To preserve data privacy and confidentiality, all personal identifiers were anonymized using alphanumeric coding (i.e., P1–P50). Participants and their parents were informed of their voluntary participation in the study and that they may withdraw at any time without prejudice.

Declaration of Competing Interest

The researcher declares no conflicts of interest.

Declaration of Generative AI and AI-Assisted Technologies

The researcher applied generative AI only to improve the legibility of the documented study and the clinical tone. The author reviewed and edited the content and is fully responsible for this publication. This is in addition to which the researcher states that the conceptual framework under revision (TDRG), including the current CALM-FPS Learning Package and the subsequent empirical data analysis ($N = 50$), was devised and conducted solely by human researchers. The principal data, analytical results, or final pedagogical conclusions were neither generated nor adapted using generative AI.

Acknowledgments

The researcher would like to extend his deepest gratitude to Dr. Dennis M. Marquez for his expert mentorship and technical guidance in developing the CALM-FPS Framework. A special thank you to the University of Perpetual Help System–DALTA. The Schools Division Superintendent, Dr. Randolph B. Tortola, and Dr. Liberty G. Puertos also deserve special attention for their administrative clearance and technical convergence with the national literacy standards. The researcher acknowledges the administration of the participating secondary school's support for the pilot program. Finally, the researcher would like to extend thanks also to the review panelists, Dr. Joemar Furigay, Dr. Anthony Gusto, and John Alfred L. Bajuyo, for methodological validation, while statistical analysis was assisted by Dr. Jolly D. Puertos. Lastly, the authors thank Dr. Therese Angeli M. Romasanta for her professional support.

References

Abril, J. G., Acerbo, C. T., & Abocejo, F. T. (2022). The Philippine informal reading inventory (Phil-IRI) program: A critical analysis. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *5*(4), 432-441. https://doi.org/10.33258/birle.v5i4.7270

Acido, J. V., & Caballes, D. G. (2024). Assessing educational progress: A comparative analysis of PISA results (2018 vs. 2022) and HDI correlation in the Philippines. *World Journal of Advanced Research and Reviews*, *21*(1), 462–474. https://doi.org/10.30574/wjarr.2024.21.1.0020

Alamo, Gian Bryce, Faith Nicole Astillo, Angela Marie Capillo, Charles Martin De Los Santos, Frances Nicole Esmeña, Robert Javier, and Andrea Denisse. "Examining the Impact of ChatGPT on the Reading Comprehension of Senior High School Students: A Sequetial Explanatory Approach." https://10.13140/RG.2.2.21826.70082

Basali, L. P. (2024). Reading comprehension levels and skills of Filipino Grade 8 students: A basis for catch-up Friday activities. *International Journal of Multidisciplinary Research*, *6*(3). https://www.ijfmr.com/papers/2024/3/23232.pdf

Bautista, J. C. (2025, March). Response to intervention within a multi-tiered systems of support: a missing element in the literacy curricula, policies, and practices in the Philippines. In *Frontiers in Education* (Vol. 10, p. 1484566). Frontiers Media SA. https://doi.org/10.3389/feduc.2025.1484566

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

Cai, L., Msafiri, M. M., Kangwa, D., et al. (2025)

Exploring the impact of integrating AI tools in higher education using the Zone of Proximal Development. *Education and Information Technologies*, *30*(6), 7191–7264. https://doi.org/10.1007/s10639-024-13112-0

Creswell, J. W., & Creswell, J. D. (2018). Research design: *Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dahl-Leonard, K., Hall, C., Cho, E., Capin, P., Roberts, G. J., Kehoe, K. F., ... & Demchak, A. (2025). Examining the Effects of Family-Implemented Literacy Interventions for School-Aged Children: A Meta-Analysis. *Educational Psychology Review*, *37*(1), 10. https://doi.org/10.1007/s10648-025-09985-3

El Hassan, F. A. M., & Alsalwah, A. F. (2025). Exploring the impact of ChatGPT on EFL reading practices: Opportunities and challenges. *International Journal of English Language Teaching*, *13*(1), 85-93. https://doi.org/10.37745/ijelt.13/vol13n18593

Kaldaras, L., Wang, K. D., Nardo, J. E., Price, A., Perkins, K., Wieman, C., & Salehi, S. (2024). Employing technology-enhanced feedback and scaffolding to support the development of deep science understanding using computer simulations. International Journal of STEM Education, 11(1), 30. https://doi.org/10.1186/s40594-024-00490-7

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: a systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, *22*(1), 7. https://doi.org/10.1186/s41239-025-00503-7

Lineman, J., Sweet, R., & Sutton, M. (2025). Beyond content: Leveraging AI and metacognitive strategies for transformative learning in higher education. The Transnational Journal of Business, 18(2), 45–62. https://doi.org/10.64010/IWMI3821

McGuire, A., Qureshi, W., Saad, M., et al. (2024). A constructivist model for leveraging GenAI tools for individualized, peer-simulated feedback on student writing. *International Journal of Technology in Education*, *7*(2), 326–352. https://doi.org/10.46328/ijte.639

National Assessment Governing Board. (2021). Reading framework for the 2021 National Assessment of Educational Progress. U.S. Department of Education. https://www.nagb.gov/naep-frameworks/reading.html

OECD. (2022). *PISA 2022 Results: Country performance in reading, math, and science.* Organisation for Economic Co-operation and Development. https://www.oecd.org/pisa/publications/

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., ... & Wagner, R. K. (2020). How the science of reading informs 21st-century education. Reading research quarterly, 55, S267-S282. https://doi.org/10.1002/rrq.352

Relyea, J. E., Kim, J. S., Rich, P., & Fitzgerald, J. (2024). Effects of tier 1 content literacy intervention on early-grade English learners' reading and writing: Exploring the mediating roles of domain-specific vocabulary and oral language proficiency. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000882

Röhl, S., Bijlsma, H., & Schwichow, M. (2025). Can feedback from students to teachers improve different dimensions of teaching quality in primary and secondary education? A hierarchical meta-analysis. Educational Assessment, Evaluation and Accountability, 37(1), 35-71. https://doi.org/10.1007/s11092-024-09450-9

Sardar, A. (2023). Constructivism in education: A

comprehensive review of theory and research. *International Research Journal of Modernization in Engineering Technology and Science, 5*(7). https://doi.org/10.56726/IRJMETS43553

Shi, J., Liu, W., & Hu, K. (2025). Exploring how AI literacy and self-regulated learning relate to student writing performance and well-being in generative AI-supported higher education. *Behavioral Sciences, 15*(5), 705. https://doi.org/10.3390/bs15050705

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Trajkovski, G., & Hayes, H. (2025). AI-assisted formative assessment and feedback. In *AI-Assisted Assessment in Education: Transforming Assessment and Measuring Learning* (pp. 283–312). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-88252-4_7

UNESCO Institute for Statistics. (2020). *Global proficiency framework for reading: Grades 1 to 9.* UNESCO.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 487662. https://doi.org/10.3389/fpsyg.2019.03087

Xiao, R., Hou, X., Ye, R., Kazemitabaar, M., Diana, N., Liut, M., & Stamper, J. (2025). Improving Student-AI Interaction Through Pedagogical Prompting: An Example in Computer Science Education. https://doi.org/10.48550/arXiv.2506.19107

World Bank. (2024). *Philippines Learning Poverty Brief.* World Bank Group. http://documents1.worldbank.org/curated/en/099090524085540633/pdf/P1792091b0057b06a1a1c61271b276117ff.pdf

International Journal of Education, Literacies, and Curriculum Studies
ISSN: 3116-2991
https://ijelcs.minduraresearch.com/journal/index