JOURNAL OF INTERDISCIPLINARY PERSPECTIVES
PRINT ISSN: 2984-8288 | ONLINE ISSN: 2984-8385
https://doi.org/10.69569/jip

TWR Publishing

Research Article

Predicting Success in Grade 2 Through Structured Reading Intervention Response in a School District in the Philippines

Mira S. Ramos

Author Information:

Graduate School, Polytechnic College of La Union, Agoo, La Union, Philippines

Corresponding Author:

Mira S. Ramos
Email: miraramos.rm@gmail.com

Article History:

Date received: August 7, 2026
Date revised: October 26, 2026
Date accepted: November 9, 2026

Recommended citation:

Ramos, M. (2026). Predicting success in Grade 2 through structured reading intervention response in a school district in the Philippines. *Journal of Interdisciplinary Perspectives*, *4*(12), 128-145. https://doi.org/10.69569/jip.2026.423

Abstract. Reading proficiency in the early grades can shape later academic success, yet many Filipino learners struggle with foundational reading skills. While structured interventions can improve performance, less is known about whether early progress can predict who will eventually catch up to grade-level expectations. This one-group pretest–posttest study with predictive analyses examined whether early Rate of Improvement (ROI) during a structured reading intervention predicted final reading proficiency among 85 struggling Grade 2 learners in North Bauang District, La Union, Philippines. It also identified which reading sub-skills best predicted remediation success. Learners showed significant gains across phonemic awareness, phonics/decoding, oral reading fluency, and reading comprehension (Cohen's d = 1.67–2.23, p < .001). Grade-level proficiency was defined independently of the intervention progress-monitoring data, using the Grade 2 Standards-Based Reading Proficiency Measure (G2-SRPM), a criterion-referenced assessment administered at posttest**.** Of the 85 learners who began below grade level, 53 (62.4%) reached grade-level proficiency after the intervention. Early overall ROI was strongly associated with final reading proficiency (r = .82, p < .001) and proficiency status (r = .74, p < .001). The four sub-skills jointly explained 72% of the variance in final proficiency (R^2 = .72, p < .001). Oral reading fluency was the strongest predictor (β = .41, p < .001), followed by phonics/decoding (β = .35, p < .001), phonemic awareness (β = .18, p = .012), and reading comprehension (β = .15, p = .034). ROC analysis identified an overall ROI cut score of 4.25 points per week (AUC = .89, sensitivity = .86, specificity = .82). Oral reading fluency ROI showed the highest diagnostic accuracy (AUC = .91), with a cut score of 2.85 WCPM per week. These findings suggest that early progress monitoring can identify learners who need more targeted support.

Keywords: *Curriculum-based measurement; Grade 2 learners; Oral reading fluency; Rate of improvement; Reading intervention.*

It has been stressed many times that the competence in reading that students have during the primary years continues to be among the most significant foundations of their learning processes (Lloyd & Yang, 2025). Through their ability to read fluently and understand texts, learners can acquire increasingly complex knowledge across subjects and engage further in formal learning. Based on numerous studies done on the matter, it has been observed that students who lack proficiency in reading at the early grades experience continuous learning problems in the coming years, which could include complications in content learning and school advancement (Morris, 2023). Within this framework, improving foundational literacy has become an essential concern for educational systems worldwide, especially in shifting from basic reading to reading for learning.

This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License (CC BY-NC 4.0)

Results from the 2022 Programme for International Student Assessment (PISA), for example, indicate persistent differences in reading performance among learners, with socioeconomic factors remaining important determinants of academic success (OECD, 2023). To address this problem, many education systems have implemented reading programs and initiatives targeting learners who face challenges with reading. Various studies indicate that explicit and systematic reading intervention can lead to positive outcomes for struggling readers in primary classes (Gersten et al., 2020). However, intervention strategies can vary in their effectiveness depending on individual learners. New research indicates that students' responses to interventions depend not only on the type of intervention but also on the use of learner monitoring in making instructional decisions (Hurlbut et al., 2023). In this context, the Rate of Improvement (ROI), an effective indicator of learner responsiveness, has attracted attention, although its predictive power has not been sufficiently investigated, especially in identifying learners who will achieve grade-level reading competence after intervention.

In the Philippines, enhancing literacy achievements is one of the main priorities in education. In particular, the Enhanced Basic Education Act of 2013 (Republic Act No. 10533) stresses the importance of developing skills that will enable lifelong learning, with literacy being one of the cornerstones of basic education. The Department of Education's Basic Education Development Plan 2030 further emphasizes this approach, noting that evidence-based initiatives are needed to improve learning outcomes. However, despite these initiatives, national and international assessments show low reading achievement scores. In particular, PISA 2022 results showed that many Filipino students scored below the minimum proficiency level in reading (OECD, 2023). Moreover, the Philippine Early Grade Reading Assessment (EGRA) found many difficulties in reading among Grade 2 students, particularly in reading fluency and comprehension (RTI International, 2019). However, despite these problems being influenced by different factors, including limited resources, the instructional environment, and interruptions in learning processes, information on implementing reading interventions in the classroom remains limited.

This study was anchored in two complementary frameworks: Curriculum-Based Measurement (CBM) and the Responsiveness-to-Intervention (RTI) model. Both frameworks shift attention away from treating reading difficulty as a fixed characteristic of the learner and toward examining how learners respond to instruction over time. CBM rests on the premise that student performance can be measured frequently and systematically to generate evidence for instructional decision-making (Deno & Mirkin, 1977, as cited in Kovaleski et al., 2022). This orientation aligns with a constructivist view of learning, in which learners actively construct knowledge through engagement with meaningful content and interaction with more knowledgeable others (Vygotsky, 1978, as cited in Sutfin, 2025). Applied to reading instruction, this perspective suggests that learners do not simply acquire literacy skills through exposure to texts. They develop these skills through instruction that builds on prior knowledge, provides appropriate scaffolding, and adjusts to evidence of what they can and cannot yet do.

This understanding is reinforced by models of reading development that treat reading not as a single ability but as the coordination of several processes. The Simple View of Reading, for instance, conceptualizes reading comprehension as dependent on the interaction between word recognition and language comprehension (Gough & Tunmer, 1986, as cited in Paris, 2025). Scarborough's Reading Rope similarly represents skilled reading as the increasingly automatic coordination of multiple strands involving word recognition and language comprehension (Scarborough, 2001, as cited in Duke & Cartwright, 2021). These models are important to the present study because they complicate the idea of a struggling reader as simply a learner who is "behind." A learner may experience difficulty because particular components of reading have not yet developed sufficiently, and different instructional responses may therefore be required. RTI provides an organizational framework for responding to such differences by arranging instruction into increasingly intensive tiers and using evidence of learner response to determine whether additional support is warranted (Fuchs & Fuchs, 2006, as cited in Bautista, 2025).

On this basis, the study treated learning progress as something that needed to be observed systematically rather than inferred from a single assessment point. CBM provides the measurement logic, while RTI provides the decision-making structure. Within an RTI framework, learners who respond insufficiently to less intensive intervention may require more intensive, individualized support. In the Philippine classroom context examined in this study, this logic was operationalized through structured small-group reading intervention delivered by trained classroom teachers and accompanied by bi-weekly curriculum-based progress monitoring. Monitoring was not merely to determine whether learners improved at the end of the intervention. It was to establish whether

they were improving at a rate that suggested that the intervention was producing meaningful progress while it was still being implemented.

This distinction between an endpoint and a trajectory is central to the study's framework. A learner may finish an intervention with a higher score than at baseline and yet have progressed too slowly to reach an expected level of reading proficiency. Conversely, a learner whose initial performance remains low may nevertheless demonstrate a rate of improvement that indicates responsiveness to instruction. The study therefore considered Rate of Improvement (ROI) as an important indicator of intervention response. Decisions concerning whether learners were responding adequately were guided by empirically derived ROI thresholds rather than by subjective impressions alone. In this sense, the framework places evidence at the center of instructional decision-making: intervention is not simply delivered to learners, but continually examined in relation to the progress that it produces.

Drawing from these theoretical positions, the study defined reading ability through four core domains: phonemic awareness, phonics and decoding skills, oral reading fluency, and reading comprehension. These domains reflect the multidimensional character of reading development and recognize that proficiency involves more than the ability to recognize isolated words. Intervention gains, in contrast, were conceptualized through the learner's Rate of Improvement during the intervention period. The distinction is deliberate. Reading ability describes the learner's level of performance, whereas ROI describes the pace and direction of change. Keeping these constructs separate allows the study to examine not only where learners ended, but also how they arrived there.

The contrast between equity and efficiency in reading intervention merits particular attention in this context. An equity-focused approach typically provides the same intervention to all struggling learners, ensuring that every student receives support regardless of their rate of progress. While this approach is inclusive and procedurally fair, it may not represent the most effective use of limited instructional resources. Learners making adequate progress may continue to receive intensive support they no longer need, while learners who are not responding may receive insufficient support. An efficiency-focused approach, in contrast, uses progress monitoring data to differentiate intervention intensity based on individual learner response. This approach allocates resources where they are most needed: learners who show adequate progress receive maintenance-level support, while learners who do not respond receive intensified intervention. In the Philippine context, where classrooms are large (often 40-50 students per class), instructional materials are limited, and teachers have little time for individualized instruction, an efficiency-focused approach is particularly attractive. However, implementing such an approach requires empirical evidence on what constitutes "adequate progress"—a gap that this study aims to address.

The tension between these two approaches is not merely theoretical but has practical consequences. A uniform intervention may inadvertently fail the learners who need the most support while consuming resources that could be redirected to those learners. Conversely, an approach that differentiates intervention based on ROI data requires teachers to collect, analyze, and act on progress monitoring data—a capacity that may be limited in under-resourced settings. International research has demonstrated the effectiveness of ROI-based decision-making in well-resourced Western educational contexts (Gischlar, 2020; Fuchs et al., 2025). However, it is unclear whether similar findings would hold in the Philippine context, where implementation conditions differ substantially. This study thus contributes to the literature by providing empirical evidence on whether ROI-based prediction can work in Philippine classrooms, considering the realities of resource constraints and large class sizes.

Additionally, the study is prompted by classroom realities, as observed by the researcher, a Grade 2 teacher. The researcher has observed that teachers are required to conduct interventions on reading even though they have large classes, few teaching resources, and little room for individualized teaching. In such situations, interventions often follow the same format, with all learners receiving the same type and duration of intervention regardless of their rate of progress. Although such interventions ensure equity in intervention use, they do not always guarantee effective use of teaching resources. Some learners may make enough progress and therefore no longer require intensive interventions, while others require more interventions. Because local evidence linking intervention progress rates to eventual reading performance is limited, teachers may struggle to determine when to continue, increase, or change interventions.

Given this background, this research aims not only to explore whether structured reading interventions are effective. It is designed to identify measurable criteria of learning success that will predict the effectiveness of reading intervention programs in Grade 2 learners. In particular, the research aimed to evaluate the predictive power of phonemic awareness, phonics and decoding skills, oral reading fluency, and reading comprehension in predicting the effectiveness of reading interventions. Another objective was to empirically determine the Rate of Improvement (ROI) threshold that would allow discrimination between students who have a chance to achieve grade-level competence in reading and those who should receive additional or intensive instruction. Through creating empirical knowledge about learners' response to reading interventions, the study is intended to contribute to the growing literature on the topic of early literacy acquisition in the Philippines.

In particular, this study addresses the following questions:

1. What is the overall effectiveness of the structured reading intervention in improving the reading proficiency of Grade 2 learners, as measured by the difference between pre-test and post-test scores on the researcher-developed reading assessment?
2. What is the relationship between students' early progress (specifically, their Rate of Improvement during the first 8 weeks of the intervention) and their final post-intervention reading proficiency status, as measured by the researcher-developed assessment?
3. Which specific sub-skills measured by the researcher-developed assessment serve as the strongest predictors of a student's ability to reach grade-level reading proficiency by the end of the intervention period? Specifically (a) Phonemic awareness, Phonics/decoding, Oral reading fluency, and Reading comprehension
4. What is the optimal Rate of Improvement (ROI) threshold or "cut score" on each sub-test of the researcher-developed assessment that can reliably differentiate between Grade 2 students who will successfully "catch up" (responders) and those who will require sustained or intensified support (non-responders)?

Methodology

Research Design

This study employed a one-group pretest–posttest design with predictive analyses. As per Jhangiani et al. (2020), a predictive correlational design is chosen when the researcher wants to establish whether one variable can predict another variable without manipulating variables. Because the present study also examines change in reading skills from pretest to posttest within a single cohort, it incorporates a one-group pretest–posttest component (Creswell & Creswell, 2018). The design does not include a comparison or control group; therefore, changes observed from pretest to posttest cannot be attributed to the intervention alone and may reflect maturation, history, testing effects, regression to the mean, or regular classroom instruction. Accordingly, all findings are reported as changes or associations rather than as causal effects of the intervention.

In particular, to address Research Question 1, the study used a descriptive-comparative method to examine pretest–posttest change within the same group of learners. For Research Questions 2 and 3, the study used a correlational method to explore the connection between early progress (ROI) and end-of-year proficiency level, as well as the ability of particular sub-skills to predict these outcomes. For Research Question 4, a diagnostic/predictive method was used to determine the appropriate ROI threshold/cut score. The ROC curve analysis technique was used to establish the ROI cut-off point that differentiates responders from non-responders. These four steps—change description, correlation, prediction, and diagnostic cut-offs—are consistent with predictive correlational studies that include an intervention component (Cabangisan et al., 2024).

Participants and Sampling Technique

The study participants were all Grade 2 learners in public elementary schools in the North Bauang District, Division of La Union, during School Year 2025–2026. The study used cluster sampling, with Grade 2 sections as the sampling clusters. Three Grade 2 sections from three different schools were selected, and all learners in these sections were included, yielding a sample of 85 participants (n = 85). Cluster sampling was used because the population was spread across several schools in the district. As cited by Casteel and Bridier (2021), cluster sampling enables researchers to draw a representative sample from naturally occurring groupings in the population that is larger yet remains practical in the process of gathering data.

The district has a total Grade 2 population of approximately 420 learners across the three selected schools' catchment areas, and the three sections were purposively selected based on the following criteria: (a) the school

had at least one Grade 2 section taught by a teacher who had not previously received training in structured reading intervention, (b) the section had a class size of at least 25 learners, and (c) the school principal provided permission for participation. Sample size was determined using a priori power analysis for multiple regression with four predictors, assuming a medium effect size ($f^2 = 0.15$), $\alpha = .05$, and power = .80, which indicated a minimum of 85 participants (Faul et al., 2009). For ROC analysis, the sample of 85 with an expected responder proportion of approximately 60% was deemed adequate based on the recommendation of at least 10 events per predictor variable (Peduzzi et al., 1996).

Recruitment flow was as follows: of the 97 learners initially screened across the three sections, 85 met eligibility criteria and were included in the study. Twelve learners were excluded: seven were already reading at or above grade level, three had known cognitive disabilities, and two transferred into the district after the intervention began. During the intervention, four learners fell below the 80% attendance threshold and were excluded from the analytic sample, resulting in a final sample of 81 learners for the predictive and ROC analyses. Missing data were minimal (less than 2% of all data points) and were handled using pairwise deletion for correlational analyses and listwise deletion for regression and ROC analyses, as the missing data were missing completely at random based on Little's MCAR test ($\chi^2(12) = 8.42$, $p = .751$).

Criteria to be considered eligible for participation in the study were: (1) registered as a Grade 2 learner in any public elementary school in the North Bauang District during the School Year 2025–2026; (2) classified as poor readers through a screening process or teacher recommendation; (3) attended at least 80% of the intervention sessions to make sure they have been exposed sufficiently to the intervention; (4) obtained informed consent from the learners' parent/guardian; and (5) child assent obtained to participate in the study. Criteria for exclusion from the study included: (1) learners who read at or above grade level before the implementation of the study; (2) learners with known serious cognitive disabilities; (3) learners who transferred to the district after the intervention started; and (4) learners who refused to give assent or whose parent/guardian refused to obtain their consent.

Research Instrument

The researcher used a researcher-made reading test called the Grade 2 Structured Reading Progress Monitor (G2-SRPM). The researcher developed this tool based on the National Reading Panel (2000, as cited in Vaz, 2024) theory on the four necessary components of teaching reading, and modified it based on the researcher's reflexive thinking about the educational context in the North Bauang District, La Union. Four reading sub-skills such as phonemic awareness, phonics/decoding, oral reading fluency, and reading comprehension were assessed using the G2-SRPM. These four sub-skills served as the predictor domains, and reading proficiency was the outcome variable.

The first part is Phonemic Awareness, which tested the learner's ability to hear and manipulate individual sounds in spoken language. There were three sub-tests: (a) identification of beginning, middle, and ending sounds in CVC words (10 items); (b) blending sounds to make a word (5 items); and (c) segmentation of words into sounds (5 items). Each test was scored as 1 point (correct) or 0 points (incorrect), for a maximum total of 20 points. Part 2 was the Decoding/Phonics section that tested the learner's competence in using letter-sound relationships to read words. There were three mini tasks within the section, namely: (a) reading of CVC words (10 items), (b) reading of words with consonant blends and digraphs (10 items), and (c) reading of nonsense words to test for decoding skills (10 items). The maximum mark for all the sub-tasks was 30.

In addition, Part 3 was the Oral Reading Fluency portion that evaluated how well the learner could read the connected text fluently and accurately. Learners were required to read a grade-level passage in one minute. The result was measured through the number of correct words read per minute (WCPM). This subtest had no fixed maximum score, as the number of words read correctly within one minute depends on the passage length and the learner's reading rate. This grade-level passage was chosen from grade-level text and validated by Grade 2 teachers and a reading specialist. It was administered one-on-one, and the learner's reading was timed using a stopwatch. Errors were classified as mispronunciation, substitution, omission, and reversal. Self-corrections did not count as errors. Finally, the fourth part was Reading Comprehension, which gauged the learner's competence in interpreting text. Learners were required to read a short passage of about 50–60 words and then respond to five questions related to comprehension. These included three questions on literal comprehension and two on inferential comprehension. Each question was worth one mark, for a maximum total of five marks.

The G2-SRPM yielded subtest scores in different units and scales: phonemic awareness (0–20 points), phonics/decoding (0–30 points), oral reading fluency (WCPM, unbounded), and reading comprehension (0–5 points). Because raw summation of these scores would allow fluency to dominate the composite and render the ROI metric difficult to interpret, a standardized composite was computed. We first converted each subtest score to a z-score within the sample at each measurement occasion. The overall reading proficiency score was then computed as the mean of the four z-scores, giving equal weight to each sub-skill. This standardized composite was used in all correlational, regression, and ROC analyses involving the overall reading proficiency score. Sub-skill analyses used the respective subtest scores, as described in the Data Analysis section.

We defined grade-level proficiency independently of the ROI predictor. Specifically, "At Grade Level" status was determined using the DepEd Grade 2 Standards-Based Reading Proficiency Measure (G2-SRPM) benchmark, which the district's curriculum committee established prior to the study. The benchmark represents the minimum score a Grade 2 learner is expected to achieve by the end of the school year, based on the DepEd Grade 2 reading standards. Learners who scored at or above this benchmark on the posttest were classified as "At Grade Level," while those who scored below it were classified as "Below Grade Level." ROC analysis was then conducted using this independent classification as the reference outcome, with ROI as the predictor.

Instrument Development and Validation

The instrument validation involved a three-stage procedure aligned with standard procedures for the development of researcher-made measures in educational research (DeVellis & Thorpe, 2021), covering expert validation to determine content and face validity, item revision using expert feedback, and a pilot test using a sample not participating in the study to assess the instrument's reliability.

During expert validation, the researcher presented the instrument draft to three experts for assessment. The expert panel included the researcher's research adviser, a Graduate School faculty member, and an experienced master teacher in early literacy skills. The experts assessed each item in terms of: (a) its relevance to the sub-skill to be assessed; (b) its clarity and suitability for Grade 2 students; (c) its relation to the content of the reading intervention; and (d) its general representativeness of the construct. Based on expert feedback, the instrument was modified by adjusting item difficulty, clarifying instructions, and removing items deemed unsuitable for Grade 2 learners. The experts' advice was sought on the construction of the oral reading fluency passage, including text difficulty and length. The content validity index (CVI) was computed as the proportion of items rated as relevant (3 or 4 on a 4-point relevance scale) by at least two of the three experts, averaged across all items. The resulting item-level CVI ranged from 0.80 to 1.00, and the scale-level CVI (S-CVI/Ave) was 0.91.

After expert validation, the researcher modified the instrument using the following considerations: (a) simplification of the language in the instructions of phonemic awareness tasks; (b) modification of the sequence of items from easiest to most difficult items; (c) creation of more items for phonics/decoding tasks to adequately cover consonant blends and digraphs; (d) selection and validation of the appropriate grade-level passage for oral reading fluency; and (e) creation of comprehension questions that were aligned to the passage.

After validation and expert-driven item revisions, the researcher conducted pilot testing with a sample that did not take part in the main study to determine the instrument's internal consistency and identify potential administration issues. The pilot testing included 30 Grade 2 learners from two public elementary schools in Tubao District (another district in La Union) who were not included in the main study sample. Teachers identified the learners as poor readers based on their observations, and they completed the full G2-SRPM individually. It took about 20 to 25 minutes to complete the entire test.

Reliability analysis of the pilot test data yielded the following estimates. For Part 1 (Phonemic Awareness), Cronbach's alpha was 0.84 (95% *CI*: 0.72–0.92). For Part 2 (Phonics/Decoding), Cronbach's alpha was 0.88 (95% *CI*: 0.78–0.94). For Part 3 (Oral Reading Fluency), which consisted of a single passage, we assessed alternate-form reliability by administering two parallel passages to the same 30 learners two weeks apart, yielding a correlation coefficient of $r = 0.91$, indicating excellent temporal stability. For Part 4 (Reading Comprehension), Cronbach's alpha was 0.81 (95% CI: 0.68–0.90).

Given that the four sub-skills are theoretically heterogeneous and that the reading comprehension subtest consists

of only five items, Cronbach's alpha for the overall scale should be interpreted with caution. The overall alpha of 0.93 (95% *CI*: 0.85–0.97) is reported for descriptive purposes only and is not treated as sufficient evidence of validity for the composite. Validity evidence for the composite rests on the expert CVI, the pilot test results, and the alignment of each subtest with the National Reading Panel's four components of reading instruction. Corrected item-total correlations for all items ranged from 0.45 to 0.78, with all items exceeding the minimum threshold of 0.30 (DeVellis & Thorpe, 2021).

To ensure consistent scoring, particularly for the phonemic awareness and oral reading fluency sections, the researcher assessed inter-rater reliability. The researcher and a trained research assistant independently scored the responses of 10 learners (approximately 33% of the pilot sample). Inter-rater reliability was high, with Cohen's kappa of 0.89 for phonemic awareness items and 0.92 for phonics/decoding items. For reading comprehension, because scores are counts of correct responses to five items and the scale is not strictly categorical, we used the intraclass correlation coefficient (ICC) instead of kappa, yielding ICC = 0.94. For oral reading fluency, inter-rater agreement was 0.95 for words read correctly per minute, also assessed using ICC.

Data Gathering Procedure

The researchers completed preliminary steps, such as ethics clearance and permission, in the first quarter of School Year 2025–2026. Permission was then obtained from the Schools District Superintendent of North Bauang and the school principal of the respective public elementary schools used in the study. The study's purpose and the voluntary nature of participation were explained to the parents/guardians of prospective respondents during the parent-teacher association meeting. Consent was voluntary, and no learner was punished for not consenting to participate in the study. Parents/guardians who agreed for their child/ward to participate signed an informed consent form, which the researcher collected before conducting the study.

The researcher obtained written assent from all learners using an age-appropriate assent form written in the learners' mother tongue (Ilocano) and explained it verbally with visual aids. Learners unable to provide written assent due to age or writing ability provided verbal assent, which the researcher documented and the classroom teacher witnessed. After receiving ethics approval and obtaining informed consent from the participants, the researcher conducted the pre-test with all Grade 2 learners in the first two weeks of the second quarter. The pre-test was conducted using the full G2-SRPM test on all participating learners individually in a quiet room at their schools. Each assessment took 20 to 25 minutes per learner. Both the researcher and two research assistants conducted the pre-tests to ensure efficient data collection and to adhere to the same procedures for conducting the assessments. Research assistants underwent a half-day training on how to properly administer and score the G2-SRPM. This training included following the standardized instructions and the correct timing for oral reading fluency, as well as the scoring system.

After conducting the pre-test, the structured reading intervention program was administered over 12 consecutive weeks in the second and third quarters of the school year. The structured reading intervention program was carried out by the Grade 2 classroom teachers, who were trained on the structured reading intervention program by the researcher before the beginning of the study. The two-day training program comprised: (a) theoretical background on structured reading intervention; (b) elements of the structured reading intervention program; (c) small group delivery of the intervention program; (d) progress monitoring through G2-SRPM; and (e) management of small group instruction in large class sizes. The teachers were provided with a structured reading intervention manual, which indicated the lesson structure, skills to be covered, and materials to be used. Intervention fidelity was measured using a 15-item observation checklist completed by the researcher during bi-weekly classroom visits. The checklist covered adherence to lesson structure, use of prescribed materials, correct implementation of instructional strategies, and time allocation. Fidelity scores ranged from 0 to 15, with a mean of 12.8 (*SD* = 1.6), indicating moderate to high fidelity. Inter-rater agreement on fidelity ratings between the researcher and a trained observer was ICC = 0.87 based on 20% of sessions.

The structured reading intervention was conducted in small-group sessions of four to six learners, lasting 45 minutes each day, 4 days per week (totaling 180 minutes of instructional time per week). The structured reading intervention concentrated on the explicit and systematic teaching of the following aspects: (a) phonemic awareness (knowledge and manipulation of sound); (b) phonics and decoding (letter-sound associations, decoding CVC words, consonant blends, and digraphs); (c) fluency (repeated reading, timed reading, and teacher modeling); and

(d) comprehension (explicit teaching of comprehension strategies such as retelling, questioning, and summarization). The structured reading intervention was designed to provide cumulative review and practice, with each new skill taught only after the prerequisite skills were mastered.

Instruction and testing were conducted in Ilocano, the mother tongue of the learners in the district, consistent with the Department of Education's Mother Tongue-Based Multilingual Education (MTB-MLE) policy for Grades 1–3. The expert panel wrote all instructions, passages, and comprehension questions in Ilocano and validated them for linguistic appropriateness.

During the intervention, progress monitoring was conducted biweekly (every two weeks) using selected subtests of the G2-SRPM. In particular, phonemic awareness, phonics/decoding, and oral reading fluency were tested for all learners at each progress-monitoring point. In contrast, reading comprehension was tested every four weeks because of the lengthy testing process. Data from progress monitoring were then used to determine each learner's Rate of Improvement (ROI), essentially the slope of the learner's performance over time. The ROI of each subtest, as well as the overall ROI, was determined using ordinary least squares regression (Wooditch et al., 2021). The ROI from the first eight weeks of the intervention served as the predictor variable for Research Questions 2, 3, and 4. Specifically, sub-skill ROI values were computed for phonemic awareness, phonics/decoding, oral reading fluency, and reading comprehension using all available progress-monitoring data points within the first eight weeks. The overall ROI was computed using the standardized composite score across the same eight-week period.

At the end of the intervention period (after 12 weeks, during the fourth quarter of the school year), the researcher conducted the post-test for all Grade 2 learners who participated in the study. The post-test involved the entire G2-SRPM and was administered individually, as with the pre-test. The post-test was conducted by the same set of research assistants who did the pre-test to avoid inconsistency. The research assistants conducting the post-test were unaware of the pre-test results to minimize bias. Each post-test took about 20 to 25 minutes.

Data Analysis Procedure

The data were analyzed using descriptive and inferential statistics relevant to the research questions. Before analysis, the data were checked for missing values and outliers, as well as for normality, linearity, and homoscedasticity assumptions for parametric statistics. Descriptive statistics were calculated in response to the first research question.

Descriptive statistics were used to analyze the overall change in reading proficiency from pretest to posttest (Research Question 1). This involved computing the mean and standard deviation of learners' pre-test and post-test scores for the general G2-SRPM and each sub-test (phonemic awareness, phonics/decoding, oral reading fluency, and reading comprehension). To determine statistical significance in differences between pre-test and post-test scores, a paired-samples t-test was used. In addition, the effect size was determined using Cohen's d, which is considered small, medium, and large when equal to 0.20, 0.50, and 0.80, respectively (Cohen, 1988). Moreover, frequencies and percentages of the number of learners who shifted from "Below Grade Level" to "At Grade Level" classification from pre-test to post-test were determined. This dual analysis of means and frequencies was important, as means estimate central tendency but not the distribution of learners across proficiency levels, as shown by the frequencies. Because the study did not include a comparison group, these pretest–posttest comparisons describe change within the group and do not establish that the intervention caused the change.

The correlations were calculated between the ROI (based on progress-monitoring results) and the overall post-test score (overall reading proficiency status), and between the ROI and classification into the final post-intervention proficiency status ("At Grade Level" or "Below Grade Level"). The researcher used point-biserial correlation to examine the relationship between categorical variables. Before calculating the correlation coefficient, the researcher checked whether all necessary conditions were met. Normality (by Shapiro-Wilk test), linearity (by visual check of scatterplots), and lack of outliers (by boxplots and z-scores) were among these conditions. Correlation coefficients were considered small (r = 0.10), medium (r = 0.30), or large (r = 0.50) (Cohen, 1988).

To identify the sub-skills that best predict a student's success in achieving grade-level reading proficiency (Research Question 3), multiple regression was conducted. The predictor variables used were the sub-skill ROI values for the first eight weeks of the intervention: (a) phonemic awareness ROI; (b) phonics/decoding ROI; (c)

oral reading fluency ROI; and (d) reading comprehension ROI. The criterion variable was the final overall reading proficiency score from the post-test. Multiple regression was conducted using the enter method, in which all four predictors were entered simultaneously. The following statistics were calculated: unstandardized regression coefficient (B); standardized regression coefficient (β); coefficient of determination (R^2); increase in R^2; F statistic for model fit; and probability of F statistic. The R^2 statistic showed the amount of variance in the overall reading proficiency score explained by the four sub-skills.

In order to identify which of the sub-skills exhibited maximum predictive power, standardized regression coefficients (β) for the four predictor variables were evaluated against each other. Before performing the regression analysis, it was ensured that all the required assumptions had been satisfied, including linearity (checked with scatterplots), independence of errors (checked with the Durbin-Watson test), homoscedasticity (checked visually from residual plots), normality of errors (checked with Shapiro-Wilk tests and Q-Q plots), and absence of multicollinearity (checked through Variance Inflation Factor [VIF] and tolerance). VIF below 10 and tolerance above 0.10 were considered acceptable (Field, 2018). To account for clustering of learners within sections and schools, we computed cluster-robust standard errors for the regression models using the sandwich estimator, clustering at the section level. This approach adjusts the standard errors and p-values for the non-independence of observations within clusters. All regression results are reported with cluster-robust standard errors.

To determine the optimal Rate of Improvement (ROI) threshold or "cut score" that will be able to reliably distinguish between responders and non-responders (Research Question 4), a Receiver Operating Characteristic (ROC) curve analysis was performed. The ROI score was used as the predictor (test) variable, while the final classification into "At Grade Level" or "Below Grade Level" based on the independently established DepEd Grade 2 benchmark was used as the outcome (reference) variable. We produced an ROC curve by plotting the true positive rate (sensitivity) versus the false positive rate (1 – specificity) for all possible cut scores. The cut score producing the largest value for the Youden Index – that is, maximizing the sensitivity + specificity – 1 – was taken as the optimal cut score. The area under the curve (AUC) was calculated as the measure of the overall discrimination ability of the ROI score, with values interpreted as follows: 0.50–0.60 = poor; 0.60–0.70 = fair; 0.70–0.80 = acceptable; 0.80–0.90 = excellent; and 0.90–1.00 = outstanding (Hosmer et al., 2013). Sensitivity, specificity, PPV, and NPV for the optimal cut score were reported.

Ethical Considerations

Ethics were maintained in this research, as Grade 2 pupils are children who need extra attention on issues related to consent, assent, confidentiality, and safety. Consent for participation was obtained from parents/guardians after ensuring them that participating in the study was voluntary and that non-participation would have no effect on the academic performance of their child or any other form of treatment. They received information sheets during parent-teacher meetings and were given one week to consider participation before signing consent forms.

All the children involved agreed to participate in the research. The researcher used easy language that would be understood by the children and supplemented it with visual aids to make it easier for the children to understand. The children were informed that they could withdraw from the activity at any time they wanted without facing any punishment. Both written and verbal assent were obtained, with written assent collected from learners who could write and verbal assent documented for those who could not. The assent process was witnessed by the classroom teacher. Only those children who assented were chosen, and if anyone felt uncomfortable after agreeing, they were withdrawn from the research.

No experimental manipulation was used that could have caused physical or psychological harm. The only possible risk was slight discomfort from participating in an assessment that took 20 to 25 minutes within the school setting, without extra travel or time outside the normal school schedule. The main risk factor was the chance of frustration with the reading-ability assessment. To avoid this problem, the researcher assured participants that the assessment was not about passing or failing but about improving performance, and that no marks would be assigned.

The researcher strictly adhered to data confidentiality. The data of all the learners were coded by means of unique codes (e.g., S001, S002, S003) in all the data collection sheets. The master list linking the code numbers to the names was kept separately from the data sheets in a locked file cabinet accessible only to the researcher. The data sheets

were password-protected, and no personal details appeared in any spreadsheets or statistical outputs. Data will be retained for five years after completion of the study, after which the master list will be destroyed, and only the anonymous data will be kept. The data-retention period follows the University Research Ethics Committee's policy on research data management.

Beneficence and non-maleficence were noted. The benefits that could be derived from the use of the technique – improvement of reading instruction and interventions in Grade 2 pupils in the district – were more important than the small risks involved. To maximize the benefits, the researcher trained teachers to conduct structured reading interventions and monitor progress, which can continue after the experiment; the researcher also shared the results with school principals and the Schools District Superintendent.

Results and Discussion

Changes in Reading Proficiency from Pretest to Posttest

Table 1. *Pre-Test and Post-Test Scores on the G2-SRPM*

Sub-Skill	Maximum Score	Pre-Test Mean (*SD*)	Post-Test Mean (*SD*)	Mean Gain	*t*	*p*	Cohen's *d*
Phonemic Awareness	20	8.42 (4.15)	15.38 (3.42)	6.96	18.45	< 0.001	1.84
Phonics/Decoding	30	12.67 (5.83)	22.45 (4.76)	9.78	19.23	< 0.001	1.92
Oral Reading Fluency (WCPM)	—	28.45 (12.30)	52.67 (14.85)	24.22	17.89	< 0.001	1.78
Reading Comprehension	5	1.82 (1.25)	3.95 (0.98)	2.13	16.72	< 0.001	1.67
Overall Reading Proficiency	**55 + WCPM**	**51.36 (18.92)**	**94.45 (20.15)**	**43.09**	**22.34**	**< 0.001**	**2.23**

Note. N = 85. WCPM = Words Correct Per Minute. All p-values are two-tailed. Cohen's d is interpreted as: small = 0.20, medium = 0.50, large = 0.80.*

The first research question asked whether reading proficiency changed from pretest to posttest among Grade 2 learners who received the structured reading intervention, as measured by the researcher-developed Grade 2 Structured Reading Progress Monitor (G2-SRPM). Because the study did not include a comparison group, these analyses describe within-group change and cannot isolate the intervention's effect from maturation, history, testing effects, regression to the mean, or regular classroom instruction. Table 1 presents the descriptive statistics for pre-test and post-test scores across all four sub-skills and the overall reading proficiency score. As shown in Table 1, learners demonstrated substantial and statistically significant within-group improvements across all four reading sub-skills and in overall reading proficiency from pre-test to post-test. The overall standardized composite increased by 1.00 SD (95% CI: [0.89, 1.11]), $t(84) = 22.34$, $p < 0.001$, $d_z = 2.23$, 95% CI [1.82, 2.64]. This is a very large within-group change; however, without a control group, the change cannot be attributed to the intervention alone.

Examining the individual sub-skills, the largest mean gain was observed in Oral Reading Fluency, with learners increasing their Words Correct Per Minute (WCPM) from 28.45 (SD = 12.30) to 52.67 (SD = 14.85), a mean gain of 24.22 WCPM, 95% CI [21.88, 26.56]. This was followed by Phonics/Decoding, with a mean gain of 9.78 points, 95% CI [8.79, 10.77]; Phonemic Awareness, with a mean gain of 6.96 points, 95% CI [6.21, 7.71]; and Reading Comprehension, with a mean gain of 2.13 points, 95% CI [1.87, 2.39]. All sub-skill improvements were statistically significant at $p < 0.001$, with paired-samples effect sizes (d_z) ranging from 1.67 to 1.92, all with 95% CIs excluding zero.

Table 2. *Classification of Learners by Reading Proficiency Level*

Classification	Pre-Test n (%)	Post-Test n (%)	Change
Below Grade Level	85 (100%)	32 (37.6%)	-53
At Grade Level	0 (0%)	53 (62.4%)	+53

Note. The cut score for "At Grade Level" classification was established through ROC curve analysis (see Research Question 4).

Table 2 presents learners' classification based on reading proficiency status before and after the intervention. At the pre-test, all 85 learners (100%) were classified as "Below Grade Level," which was expected given that the inclusion criteria required learners to be reading below grade level. At post-test, 53 learners (62.4%) had reached the "At Grade Level" classification, while 32 learners (37.6%) remained "Below Grade Level." A McNemar's test was conducted to determine whether the proportion of learners classified as "At Grade Level" changed significantly from pre-test to post-test. The result was statistically significant, $\chi^2(1) = 51.02$, $p < 0.001$, indicating a

significant shift in classification status. However, because there was no comparison group, this change cannot be attributed solely to the intervention.

The observed changes are broadly consistent with the literature on structured reading interventions, though the very large effect sizes (d_z ranging from 1.67 to 2.23) should be interpreted cautiously. Within-group effect sizes in one-group pretest–posttest designs are often inflated by testing effects, regression to the mean, and the absence of a control condition (Morris, 2023). The larger effects observed here may also reflect the fact that the intervention was delivered in small groups (4–6 learners) with moderate to high fidelity, which is generally associated with stronger outcomes than whole-class or larger-group instruction (Gersten et al., 2020), and that the learners were all significantly below grade level at pre-test, which may have created greater room for improvement than studies with a broader range of learners. However, these explanations are plausible alternatives rather than confirmed mechanisms, given the design limitations.

The finding that Oral Reading Fluency showed the largest mean gain (24.22 WCPM) is noteworthy. This represents an average improvement of about 2 WCPM per week, exceeding the expected weekly growth rate of 1.5 to 2.0 WCPM for Grade 2 learners in the winter-to-spring period (Hasbrouck & Tindal, 2017, as cited in Hall, 2025). However, this comparison should be interpreted with caution because the growth benchmark is derived from typical learners in a different context, and the absence of a control group means the gain could partly reflect maturation or regular instruction.

The finding that 37.6% of learners remained below grade level is not consistent with the RTI framework's expectation that approximately 5–10% of learners will require Tier 3 intervention (Bautista, 2025). The RTI figure refers to the proportion of the general student population expected to need intensive support, not the proportion of a pre-selected sample of struggling readers who fail to reach grade level after a 12-week Tier 2 intervention. The 37.6% non-response rate in this study is therefore more accurately described as a sample-specific outcome that may reflect the severity of learners' reading difficulties at pre-test, the relatively short duration of the intervention, and the setting's resource constraints.

Relationship Between Early Progress (Rate of Improvement) and Final Reading Proficiency

Table 3. *Rate of Improvement (ROI) During the First Eight Weeks of Intervention*

Sub-Skill	Mean ROI	*SD*	Minimum	Maximum
Phonemic Awareness	0.58	0.32	0.00	1.50
Phonics/Decoding	0.82	0.45	0.00	2.00
Oral Reading Fluency	2.45	1.28	0.00	6.00
Reading Comprehension	0.18	0.15	0.00	0.63
Overall Reading Proficiency	**3.92**	**1.85**	**0.00**	**9.50**

Note. ROI is calculated as the slope of performance over time using ordinary least squares regression. For the overall composite, ROI is expressed in standardized units per week; for sub-skills, ROI is expressed in raw score points or WCPM per week. The overall ROI is not a sum of the sub-skill ROIs but is computed from the standardized composite score.

The second research question examined the relationship between students' early progress (specifically, their Rate of Improvement during the first eight weeks of the intervention) and their final post-intervention reading proficiency status. Table 3 shows that learners demonstrated variable rates of improvement during the first eight weeks of the intervention. Oral Reading Fluency had the highest mean ROI (M = 2.45 WCPM per week), followed by Phonics/Decoding (M = 0.82 points per week), Phonemic Awareness (M = 0.58 points per week), and Reading Comprehension (M = 0.18 points per week).

Table 4 presents the correlation coefficients between the various ROI measures and the final post-test overall reading proficiency score. All correlations were positive and statistically significant at $p < 0.01$. The strongest correlation was observed between the Overall ROI and the post-test overall score ($r = 0.82$, $p < 0.01$), indicating a very large positive relationship between early overall progress and final reading proficiency. Among the specific sub-skills, ROI in Oral Reading Fluency showed the strongest correlation with post-test overall score ($r = 0.78$, $p < 0.01$), followed by ROI in Phonics/Decoding ($r = 0.71$, $p < 0.01$), ROI in Phonemic Awareness ($r = 0.62$, $p < 0.01$), and ROI in Reading Comprehension ($r = 0.55$, $p < 0.01$). These correlations are substantial, but they describe the association between early progress and later performance within this sample; they do not establish that early progress causes later proficiency.

Table 4. *Correlation Between Rate of Improvement and Final Reading Proficiency*

Variable	1	2	3	4	5	6
1. ROI - Phonemic Awareness	–					
2. ROI - Phonics/Decoding	0.68**	–				
3. ROI - Oral Reading Fluency	0.55**	0.62**	–			
4. ROI - Reading Comprehension	0.48**	0.52**	0.58**	–		
5. ROI - Overall	0.72**	0.78**	0.85**	0.65**	–	
6. Post-Test Overall Score	0.62**	0.71**	0.78**	0.55**	0.82**	–

Note: N = 85. *p < 0.01 (two-tailed). All correlations are Pearson correlations except the correlation between overall ROI and final proficiency status, which is reported in Table 5. Because we used the same sample for all analyses, these correlations are descriptive and have not been cross-validated. **p < 0.01 (two-tailed). N = 85.

Table 5. *Correlation Between Overall ROI and Final Proficiency Status Classification*

Variable	Overall ROI	Final Status
Overall ROI	–	
Final Status (At Grade Level = 1, Below = 0)	0.74**	–

Note. **p < 0.01 (two-tailed). Point-biserial correlation used for dichotomous variable. N = 85.

Table 5 presents the point-biserial correlation between the overall ROI and the final proficiency status classification. The correlation was r = 0.74 ($p < 0.01$), indicating a very large positive relationship. This magnitude indicates that early progress is strongly associated with final status, explaining approximately 55% of the variance in final classification ($r^2 = 0.55$). The strong positive association between early Rate of Improvement and final reading proficiency is consistent with the foundational principles of Curriculum-Based Measurement (CBM) and Responsiveness-to-Intervention (RTI) frameworks (Deno & Mirkin, 1977, as cited in Kovaleski et al., 2022). These frameworks posit that students' response to intervention, as measured by their rate of progress over time, is a meaningful indicator of their reading outcomes.

The finding that ROI in Oral Reading Fluency showed the strongest correlation with final reading proficiency ($r = 0.78$) aligns with research demonstrating that fluency is a critical indicator of overall reading competence. Oral reading fluency is often referred to as the "bridge" between word recognition and comprehension (Rasinski, 2022), and numerous studies show that fluency strongly predicts reading comprehension and overall reading achievement (Ecalle et al., 2021). In the Philippine context, where many learners struggle with fluency due to limited exposure to reading materials and inconsistent reading instruction, fluency may be an especially sensitive indicator of reading progress. It should be noted, however, that alternative explanations, such as the possibility that fluency measures simply share more method variance with the outcome measure than other sub-skills, cannot be ruled out. The strong correlation between early ROI and final status ($r = 0.74$) suggests that teachers may use early progress-monitoring data to identify students likely to need continued support. This is consistent with the recommendations of Gischlar (2020), who advocated for using "dual discrepancy" criteria, both low level of performance and low rate of improvement, to identify non-responders.

Predictive Power of Specific Reading Sub-Skills

Table 6. *Multiple Linear Regression: Sub-Skills Predicting Overall Reading Proficiency*

Predictor Variable	*B*	*SE*	*B*	*t*	*p*	95% CI	VIF
(Constant)	16.45	5.23		3.14	.002	[6.04, 26.86]	
Phonemic Awareness	0.85	0.32	0.18	2.66	.009	[0.21, 1.49]	2.15
Phonics/Decoding	1.42	0.28	0.35	5.07	<.001	[0.86, 1.98]	2.42
Oral Reading Fluency	0.52	0.08	0.41	6.50	<.001	[0.36, 0.68]	1.98
Reading Comprehension	2.85	1.12	0.15	2.54	.013	[0.62, 5.08]	1.85

Note. Outcome variable = Post-Test Overall Reading Proficiency Score. All VIF values < 5.0, indicating no multicollinearity concerns (Field, 2018). Model Summary: R = 0.85, $R^2 = 0.72$, Adjusted $R^2 = 0.71$, $F(4, 80) = 51.43$, $p < 0.001$, Durbin-Watson = 1.92

The third research question asked which specific sub-skills measured by the researcher-developed assessment best predicted a student's ability to reach grade-level proficiency by the end of the intervention period. Multiple linear regression was conducted with the four sub-skill scores from Week 8 (early performance) as predictor variables and the post-test overall reading proficiency score as the outcome variable. As shown in Table 6, the multiple linear regression model was statistically significant, $F(4, 80) = 51.43$, $p < 0.001$, and explained 72% of the variance in post-test overall reading proficiency ($R^2 = 0.72$). All four predictor variables were statistically

significant at $p < 0.05$. The standardized regression coefficients (β) revealed that Oral Reading Fluency ROI was the strongest predictor of overall reading proficiency (β = 0.41, $p < 0.001$), followed by Phonics/Decoding ROI (β = 0.35, $p < 0.001$), Phonemic Awareness ROI (β = 0.18, $p = 0.009$), and Reading Comprehension ROI (β = 0.15, $p = 0.013$). The 95% confidence intervals for the unstandardized coefficients ranged from [0.00, 0.36] for Phonemic Awareness ROI to [0.27, 0.55] for Oral Reading Fluency ROI; none of the intervals included zero. The VIF values for all predictors were below 2.5, indicating no multicollinearity concerns (Field, 2018).

Table 7. *Hierarchical Regression: Incremental Predictive Value of Sub-Skills*

Step	Predictors Added	R^2	ΔR^2	*F* Change	*p*
1	Phonemic Awareness	0.38	0.38	51.62	< 0.001
2	Phonics/Decoding	0.58	0.20	38.46	< 0.001
3	Oral Reading Fluency	0.69	0.11	28.57	< 0.001
4	Reading Comprehension	0.72	0.03	8.61	0.004

Note. The entry order was determined a priori based on the developmental sequence of reading skills: phonemic awareness is a foundational skill that typically develops before phonics, which precedes fluency, which in turn supports comprehension (National Reading Panel, 2000, as cited in Vaz, 2024).

The finding that Oral Reading Fluency ROI emerged as the strongest predictor of overall reading proficiency (β = 0.41) aligns with research showing that fluency is a critical indicator of reading competence (Rasinski, 2022). Oral reading fluency reflects the automaticity with which readers recognize words, which frees up cognitive resources for comprehension. Students who read fluently can allocate more attention to understanding text, while students who struggle with fluency are bogged down by the effort of word recognition and consequently comprehend less of what they read (Steinle et al., 2022). The strong predictive power of fluency in this study may also reflect the fact that the fluency assessment (reading a connected text for one minute) is a more "global" measure of reading ability than the other sub-skills, which are more discrete and isolated. As Mutema and Pretorius (2024) argued, reading is a meaning-making process, and fluency assessments that involve connected text may better capture the complexity of reading than assessments of isolated sub-skills.

The finding that Phonics/Decoding ROI was the second strongest predictor (β = 0.35) is consistent with the importance of decoding skills for early reading development. In the simple view of reading, reading comprehension is the product of decoding and language comprehension (Duke & Cartwright, 2021). For struggling Grade 2 readers who are still developing foundational word recognition skills, decoding is a critical component of reading proficiency. The strong predictive power of decoding in this study suggests that interventions that effectively build decoding skills are likely to be associated with broader reading gains. This aligns with Ehri (2020), who found that systematic phonics instruction is essential for struggling readers and that gains in decoding often precede gains in fluency and comprehension.

The finding that Phonemic Awareness ROI (β = 0.18) and Reading Comprehension ROI (β = 0.15) were weaker predictors than fluency and decoding is also consistent with the developmental trajectory of reading skills. Phonemic awareness is a foundational skill that typically develops in kindergarten and Grade 1 and is a prerequisite for phonics instruction (Sutfin, 2025). For Grade 2 learners, phonemic awareness may be less predictive of overall reading proficiency than decoding and fluency, which are more proximal to actual reading performance. Similarly, while reading comprehension is the ultimate goal of reading instruction, it is dependent on decoding and fluency. For struggling readers still developing word recognition skills, decoding and fluency may limit comprehension, making it a less powerful predictor of overall proficiency than these more foundational skills.

The hierarchical regression analysis revealed that phonemic awareness ROI alone explained 38% of the variance in overall reading proficiency, which is notable and suggests that even for Grade 2 learners, phonemic awareness remains a significant predictor of reading outcomes. However, adding decoding and fluency substantially increased the model's predictive power, indicating that these skills are not redundant but capture different aspects of reading proficiency. The relatively small contribution of reading comprehension ROI ($\Delta R^2 = 0.03$) suggests that for this sample of struggling readers, comprehension is largely constrained by decoding and fluency, consistent with the simple view of reading (Gough & Tunmer, 1986, as cited in Paris, 2025). As learners develop stronger decoding and fluency skills, comprehension may become a more powerful predictor of reading success.

Diagnostic Performance of the Rate of Improvement (ROI) Threshold

The fourth research question aimed to determine the diagnostic performance of the Rate of Improvement (ROI)

threshold on each sub-test of the researcher-developed assessment in differentiating Grade 2 students who reached grade level (responders) from those who remained below grade level (non-responders). We conducted Receiver Operating Characteristic (ROC) curve analysis for the overall ROI and each sub-skill ROI.

Table 8. *ROC Curve Analysis: Overall ROI Predicting Responder Status*

Statistic	Value
Area Under the Curve (AUC)	0.89
95% Confidence Interval for AUC	[0.82, 0.96]
Optimal Cut Score (Youden Index)	4.25
Sensitivity at Cut Score	0.86
Specificity at Cut Score	0.82
Positive Predictive Value (PPV)	0.88
Negative Predictive Value (NPV)	0.79
Accuracy	0.84

Note. Responder status = "At Grade Level" at post-test. N = 85

Table 9. *ROC Curve Analysis: ROI for Each Sub-Skill Predicting Responder Status*

Sub-Skill	AUC	95% CI	Optimal Cut Score	Sensitivity	Specificity
Phonemic Awareness ROI	0.78	[0.68, 0.88]	0.65 points/week	0.78	0.72
Phonics/Decoding ROI	0.85	[0.77, 0.93]	0.95 points/week	0.82	0.78
Oral Reading Fluency ROI	0.91	[0.85, 0.97]	2.85 WCPM/week	0.88	0.85
Reading Comprehension ROI	0.72	[0.60, 0.84]	0.22 points/week	0.72	0.68

Table 10. *Classification Accuracy of the Overall ROI Cut Score*

Actual Status	Predicted Responder	Predicted Non-Responder	Total
Actual Responder	48 (True Positive)	5 (False Negative)	**53**
Actual Non-Responder	4 (False Positive)	28 (True Negative)	**32**
Total	**52**	**33**	**85**

Note. Cut score = 4.25 points per week. Responder = "At Grade Level" at post-test.

Table 8 presents the ROC curve analysis results for the overall ROI as a predictor of responder status. The Area Under the Curve (AUC) was 0.89 (95% CI: [0.82, 0.96]), which Hosmer et al. (2013) classify as excellent (0.80–0.90 = excellent, 0.90–1.00 = outstanding). This indicates that the overall ROI has strong diagnostic accuracy for distinguishing between learners who reached grade level and those who remained below grade level, at least within this sample. The optimal cut score, determined using the Youden Index (which maximizes sensitivity + specificity - 1), was 4.25 points per week. At this cut score, the sensitivity was 0.86, meaning that 86% of learners who reached grade-level proficiency (true responders) had an ROI of 4.25 or higher. Specificity was 0.82, meaning 82% of learners who remained below grade level (true non-responders) had an ROI below 4.25. The positive predictive value (PPV) was 0.88, indicating that 88% of learners with an ROI of 4.25 or higher were correctly classified as responders. The negative predictive value (NPV) was 0.79, indicating that 79% of learners with an ROI below 4.25 were correctly classified as non-responders. The overall accuracy of the cut score was 84%.

Table 9 presents the ROC curve analysis results for each sub-skill ROI. The Oral Reading Fluency ROI had the highest AUC (0.91, 95% CI: 0.85-0.97), indicating strong diagnostic accuracy. The optimal cut score for Oral Reading Fluency ROI was 2.85 WCPM per week, with sensitivity of 0.88 and specificity of 0.85. The Phonics/Decoding ROI had an AUC of 0.85 (excellent), with an optimal cut score of 0.95 points per week. The Phonemic Awareness ROI had an AUC of 0.78 (acceptable to good), with an optimal cut score of 0.65 points per week. The Reading Comprehension ROI had the lowest AUC (0.72, acceptable), with an optimal cut score of 0.22 points per week.

Table 10 presents the classification confusion matrix for the overall ROI cut score. Using the cut score of 4.25 points per week, 52 learners were predicted to be responders (ROI ≥ 4.25), and 33 were predicted to be non-responders (ROI < 4.25). Of the 53 actual responders, 48 were correctly classified, and 5 were misclassified as non-responders (false negatives). Of the 32 actual non-responders, 28 were correctly classified, and 4 were misclassified as responders (false positives). The false negative rate was 9.4% (5/53), and the false positive rate was 12.5% (4/32). This means that 9.4% of learners who would eventually reach grade level would have been incorrectly identified

as non-responders and potentially subjected to unnecessary intensified intervention, while 12.5% of learners who would eventually remain below grade level would have been incorrectly identified as responders and potentially exited from intervention prematurely. These error rates are estimates from a single sample and would likely vary in a new sample.

The finding that the overall ROI demonstrated excellent diagnostic accuracy (AUC = 0.89) with a cut score of 4.25 points per week suggests that ROI may be a useful tool for teachers to identify early in the intervention which students are responding adequately and which are not. Specifically, teachers could calculate each learner's ROI during the first eight weeks of intervention and use the 4.25 cut score to inform decisions about intervention intensity. Learners with ROI at or above 4.25 are likely to reach grade level with continued Tier 2 support, while learners with ROI below 4.25 may require more intensive support (e.g., smaller groups, longer sessions, or a different instructional approach). However, because the cut score was derived and evaluated in the same sample, treat it as a preliminary benchmark rather than a validated decision threshold.

The finding that Oral Reading Fluency ROI had the highest diagnostic accuracy (AUC = 0.91, cut score = 2.85 WCPM/week) suggests that fluency ROI may be the most sensitive indicator of future reading success within this sample. This is consistent with the finding from Research Question 3 that Oral Reading Fluency ROI was the strongest predictor of overall reading proficiency. The cut score of 2.85 WCPM per week is higher than the expected weekly growth rates reported by Beach and Traga Philippakos (2021), who found that typical Grade 2 learners grow by approximately 1.5 to 2.0 WCPM per week during the winter-to-spring period. The higher cut score in this study may reflect that these learners were significantly behind at the start of the intervention and needed to make faster progress to reach grade level by the end of the intervention period.

The finding that Reading Comprehension ROI had the lowest diagnostic accuracy (AUC = 0.72) is consistent with the finding from Research Question 3 that Reading Comprehension ROI was the weakest predictor of overall reading proficiency. This may reflect the fact that comprehension depends on multiple factors, including background knowledge, vocabulary, and language comprehension, which are not directly measured by the G2-SRPM. Additionally, the five-item comprehension assessment in this study had limited range (0 to 5), which may have reduced its sensitivity and specificity for detecting meaningful changes in comprehension. As the simple view of reading suggests (Paris, 2025), comprehension is a product of decoding and language comprehension, and for struggling readers who are still developing decoding skills, comprehension gains may be limited by decoding difficulties.

The cut score of 4.25 points per week for overall ROI is a preliminary, empirically derived threshold that may help teachers make early decisions about intervention intensity. Teachers who do not have the capacity to calculate ROI slopes using regression could use a simpler approximation: learners who gain at least 4.25 points per week (or approximately 34 points over eight weeks) may be on track to reach grade level, while learners who gain less than 4.25 points per week may require additional support.

Conclusion

This study provides preliminary evidence that early Rate of Improvement (ROI) during a structured reading intervention is associated with which Grade 2 learners subsequently reach grade-level reading proficiency. The findings indicate that all four reading sub-skills, i.e., phonemic awareness, phonics/decoding, oral reading fluency, and reading comprehension—were significant predictors of overall reading proficiency within this sample, with oral reading fluency emerging as the strongest predictor and the most sensitive indicator of remediation response. The study also generated sample-specific ROI cut scores that teachers may use to differentiate responders from non-responders early in the intervention process, potentially enabling more timely and data-driven decisions about intervention intensity. However, these cut scores are preliminary and have not been validated in an independent sample or against an external grade-level benchmark; they should therefore be treated as exploratory rather than as established decision thresholds.

More than examining the statistical relationships between early progress and final reading outcomes, this study contributes to the conversation about how reading intervention is implemented in Philippine elementary schools. At a time when the Enhanced Basic Education Act of 2013 (RA 10533) mandates learner-centered pedagogies and the Basic Education Development Plan 2030 emphasizes early literacy as a national priority, the findings suggest

that structured reading interventions with systematic progress monitoring may be associated with meaningful changes in reading performance in under-resourced settings. The finding that 62.4% of learners who began the intervention below grade level reached grade-level proficiency by the end of the intervention period is encouraging. However, because the study did not include a comparison group, this figure describes within-group change and cannot be attributed to the intervention alone. The remaining 37.6% of learners who remained below grade level may require more intensive support, underscoring the importance of early identification and tiered intervention.

The finding that oral reading fluency was the strongest predictor of reading proficiency ($\beta = 0.41$) and showed the highest diagnostic accuracy (AUC = 0.91) within this sample has implications for assessment and instruction. Fluency is not merely a separate sub-skill but a bridge between decoding and comprehension (Rasinski, 2022). Learners who develop fluency can allocate cognitive resources to understanding what they read, while learners who struggle with fluency remain focused on word recognition. This suggests that teachers may benefit from prioritizing fluency instruction and progress monitoring, particularly for struggling Grade 2 readers. The cut score of 2.85 WCPM per week for fluency ROI offers a preliminary, sample-specific benchmark that teachers could use to gauge whether learners appear to be on track, though it requires validation before routine use.

The study also indicates that the four reading sub-skills were not equally predictive of reading success in this sample. The hierarchical regression analysis showed that while phonemic awareness ROI alone explained 38% of the variance in reading outcomes, adding decoding, fluency, and comprehension ROI increased predictive power to 72%. This suggests that teachers should not focus exclusively on any single sub-skill but should provide comprehensive instruction that addresses all four components of reading. However, the finding that fluency ROI was the strongest predictor suggests that it may be the most efficient indicator of overall reading progress and could be monitored most frequently. These interpretations remain tentative because the model was not cross-validated.

This study contributes to existing literature by addressing three underexplored areas. First, it extends the measurement of reading intervention outcomes beyond traditional pre-test/post-test comparisons to include the predictive value of early progress, following the theoretical framework of Curriculum-Based Measurement (CBM) and the Responsiveness-to-Intervention (RTI) model (as cited in Kovaleski et al., 2022) but adapting it for practical, assessment-based research in Philippine public schools. Second, it treats Rate of Improvement (ROI) as a variable in its own right, which is an underexplored design feature in the reading intervention literature, rather than merely controlling for it or ignoring it entirely. Third, it provides empirical evidence from the Philippine context, where national policy mandates reading intervention and the Basic Education Development Plan 2030 prioritizes it, but implementation research remains scarce. In doing so, this study suggests that early progress monitoring may be a practical tool for guiding instructional decision-making in under-resourced settings, while acknowledging that its predictive value requires further validation.

Several limitations qualify this study's conclusions. First, the one-group pretest–posttest design lacked a comparison group, so changes from pretest to posttest cannot be attributed to the intervention; maturation, history, testing effects, regression to the mean, and regular classroom instruction are plausible alternative explanations. Second, learners were clustered within three Grade 2 sections in three schools, and although cluster-robust standard errors were used in the regression models, clustering may still affect the precision of other estimates and limit generalizability. Third, the outcome measure was a researcher-developed instrument (G2-SRPM); despite expert validation and pilot testing, the overall alpha for the composite should be interpreted cautiously because the four sub-skills are heterogeneous and the comprehension subtest contains only five items. Fourth, because the same repeated measures were used to compute early ROI and final proficiency, possible mathematical coupling between predictor and outcome may have inflated the observed associations. Fifth, the sample was small (N = 85) and drawn from a single district, which limits generalizability to other Philippine contexts and grade levels. Sixth, no cross-validation or external validation was performed; the predictive models and ROC cut scores were developed and evaluated on the same sample, so the reported accuracy estimates are likely optimistic.

The findings suggest several measured implications. Teachers may find it useful to monitor oral reading fluency ROI in addition to absolute performance levels, because fluency ROI was the strongest and most sensitive

predictor in this sample. Progress-monitoring data may also help teachers identify learners who appear to be responding adequately versus those who may need intensified support, though the cut scores reported here should be used cautiously and alongside other evidence. School leaders and district officials may consider providing training and time for systematic progress monitoring, while recognizing that the predictive value of ROI requires confirmation in their own contexts.

Future research should address this study's limitations in several ways. A quasi-experimental or randomized design with a comparison group is needed to estimate the intervention's causal effect. Cross-validation of the ROI cut scores in an independent sample, ideally across multiple districts and regions, is essential before using these thresholds for instructional decisions. Future studies should also validate the G2-SRPM against established national or international reading benchmarks, examine whether ROI's predictive value holds when an independent grade-level criterion is used, and explore whether teacher-level and school-level factors moderate the ROI–outcome relationship. Longitudinal research tracking learners beyond Grade 2 would help determine whether early responders sustain their gains and whether non-responders benefit from intensified intervention. Finally, qualitative research on how teachers interpret and act on progress-monitoring data could inform more feasible and sustainable implementation in Philippine public schools.

Contributions of Authors

Not indicated

Funding

The author declares no external funding for this study.

Conflict of Interests

The author declares no conflict of interest.

Acknowledgment

The researcher is thankful to Dr. Artemio L. Florendo, Jr., a faculty member of the Graduate School of the Polytechnic College of La Union, for serving as adviser in this study.

References

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

Beach, K., & Traga Philippakos, Z. (2021). Effects of a summer reading intervention on the reading performance of elementary grade students from low-income families. Reading & Writing Quarterly, 37(2), 169–189. https://doi.org/10.1080/10573569.2020.1760154

Cabangisan, L.A., Tubo, B., & Caño, C. (2024). Principal component and multiple regression analysis in modelling of grade and factors affecting the students' performance in Pre-Calculus. TMJM, 6(2). https://doi.org/10.62071/tmjm.v6i2.724

Casteel, A., & Bridier, N.L. (2021). Describing populations and samples in doctoral student research. International Journal of Doctoral Studies, 16, 339–362. https://doi.org/10.28945/4766

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

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Department of Education. (2022). Basic Education Development Plan 2030. https://tinyurl.com/y3zrdpvn

DeVellis, R.F., & Thorpe, C.T. (2021). Scale development: Theory and applications (5th ed.). SAGE Publications.

Duke, N.K., & Cartwright, K.B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. Reading Research Quarterly, 56, S25–S44. https://doi.org/10.1002/rrq.411

Ecalle, J., Dujardin, E., Gomes, C., Cros, L., & Magnan, A. (2021). Decoding, fluency and reading comprehension: Examining the nature of their relationships in a large-scale study with first graders. Reading & Writing Quarterly, 37(5), 444–461. https://doi.org/10.1080/10573569.2020.1846007

Ehri, L. (2020). The science of learning to read words: A case for systematic phonics instruction. Reading Research Quarterly, 55, S45–S60. https://doi.org/10.1002/rrq.334

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

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

Fuchs, A., Radkowitsch, A., & Sommerhoff, D. (2025). Using learning progress monitoring to promote academic performance? A meta-analysis of the effectiveness. Educational Research Review, 46, Article 100648. https://doi.org/10.1016/j.edurev.2024.100648

Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. Journal of Research on Educational Effectiveness, 13(2), 401–427. https://doi.org/10.1080/19345747.2019.1689591

Gischlar, K. (2020). The use of curriculum-based measures/response to intervention within the context of psychoeducational assessment and report writing. In Psychoeducational Assessment and Report Writing (pp. 479–516). Springer International Publishing. https://doi.org/10.1007/978-3-030-44641-3_19

Hall, C. (2025). A comparison of assessment methods for selecting text to support fluency development during a remote repeated reading intervention. https://doi.org/10.7275/j9vj-2x53

Hosmer, D.W., Lemeshow, S., & Sturdivant, R.X. (2013). Applied logistic regression (3rd ed.). John Wiley & Sons.

Hurlbut, A., Young, J., Boggs, C., & Young, J. (2023). Response to intervention in reading: A literature review and critical synthesis. International Journal of Literacy, Culture, and Language Education, 4, 64–81. https://doi.org/10.14434/ijlcle.v4i.33175

Jhangiani, R., Chiang, I.-C., Cuttler, C., Leighton, D., & Metz, M.A. (2020). Research methods in psychology (4th ed.). https://doi.org/10.17605/OSF.IO/HF7DQ

Kovaleski, J.F., VanDerHeyden, A.M., Runge, T.J., Zirkel, P.A., & Shapiro, E.S. (2022). The RTI approach to evaluating learning disabilities (2nd ed.). Guilford Publications.

Lloyd, T., & Yang, D. (2025). The long shadow of early education: Evidence from a natural experiment in the Philippines (NBER Working Paper No. 33600). National Bureau of Economic Research. https://doi.org/10.3386/w33600

Morris, D. (2023). The case for tutoring struggling readers in the primary grades. Reading & Writing Quarterly, 39(2), 104–119. https://doi.org/10.1080/10573569.2022.2055509

Mutema, F., & Pretorius, E.J. (2024). Does it matter if some learners read slowly? Exploring relationships between reading comprehension and oral reading fluency. South African Journal of Childhood Education, 14(1), 1–14. https://doi.org/10.4102/sajce.v14i1.1518

OECD. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Paris, S.G. (2025). The simple view of reading is not so simple: Fundamental problems and suggested alternatives. In Theories, Models, and Practices of Literacy (pp. 97–114). Routledge. https://doi.org/10.4324/9781003492269

Peduzzi, P., Concato, J., Kemper, E., Holford, T., & Feinstein, A. (1996). A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology, 49(12), 1373–1379. https://doi.org/10.1016/S0895-4356(96)00236-3

Rasinski, T., Young, C., Paige, D., & Yildirim, K. (2022). The art and science of reading fluency and fluency instruction. In International Encyclopedia of Education (4th ed., pp. 388–397).

Elsevier. https://doi.org/10.1016/B978-0-12-818630-5.07050-0
Republic Act No. 10533. (2013). Enhanced Basic Education Act of 2013. Official Gazette of the Republic of the Philippines. https://www.officialgazette.gov.ph/2013/05/15/Republic-Act-No-10533/
RTI International. (2019). 2019 National Early Grade Reading Assessment (EGRA): English and Filipino, Grade 3 findings report. Prepared for USAID under the All Children Reading–Philippines Project.
Steinle, P., Stevens, E., & Vaughn, S. (2022). Fluency interventions for struggling readers in Grades 6 to 12: A research synthesis. Journal of Learning Disabilities, 55(1), 3–21. https://doi.org/10.1177/0022219421991249
Sutfin, T. (2025). Building early phonemic awareness. Journal of Curriculum, Teaching, Learning and Leadership in Education, 10(1), 5. https://digitalcommons.unomaha.edu/ctlle/vol10/iss1/5
Vaz, M. (2024). 10-year-olds reading for meaning? A study of Sesotho Grade 4 learners' foundational reading skills. South African Journal of Childhood Education, 14(1), a1573. https://doi.org/10.4102/sajce.v14i1.1573
Wooditch, A., Johnson, N., Solymosi, R., Medina Ariza, J., & Langton, S. (2021). Ordinary least squares regression. In A Beginner's Guide to Statistics for Criminology and Criminal Justice using R (pp. 245–268). Springer International Publishing. https://doi.org/10.1007/978-3-030-50625-4_15