

Reading Proficiency Gaps and Comprehension Recovery among Grade 6 Learners

Jollybeth R. Coloma^{1*} and Michelle S. Morales^{1,2}

¹Northeastern College

*colomajollybee@gmail.com, ²michellemorales@northeasterncollege.edu.ph

Date Submitted:
February 13, 2026

Date Accepted:
April 16, 2026

Date Published:
July 18, 2026

DOI:
10.5281/zenodo.21424588

ABSTRACT

This study investigated reading proficiency gaps and comprehension recovery among Grade 6 learners through a diagnostic interrupted time-series design. Conducted at Epifanio Delos Santos Elementary School in Manila, the research used repeated baseline assessments, progress monitoring, and a final assessment to trace changes in word recognition, oral reading fluency, vocabulary in context, literal comprehension, inferential comprehension, and evaluative comprehension. A validated and reliable reading assessment battery was administered, while segmented linear mixed-effects growth modeling, benchmark-referenced gap analysis, standardized mean change, and the Reliable Change Index were employed to evaluate performance. Baseline findings revealed substantial proficiency

deficits, with the widest gaps observed in inferential and evaluative comprehension. Scores remained stable across the three baseline administrations, indicating persistent reading difficulty. Following structured and assessment-guided recovery instruction, learners demonstrated a significant immediate improvement and a sustained positive growth trajectory across the monitoring periods. All reading components improved, with inferential comprehension showing the strongest standardized gain. Word recognition and literal comprehension approached the expected benchmark, while evaluative comprehension remained the most difficult area. Learner-level analysis showed that most participants achieved reliable improvement, although only a smaller proportion reached full comprehension recovery. The findings indicated that reading gaps among Grade 6 learners could be reduced through multicomponent instruction, continuous progress monitoring, and timely instructional adjustment. Continued intervention, individualized support, and coordinated transition planning were recommended for learners who remained below proficiency.

Keywords: *assessment-guided instruction, comprehension recovery, grade 6 learners, progress monitoring, reading proficiency gaps, segmented growth analysis*

INTRODUCTION

Reading proficiency shapes a learner's ability to participate meaningfully in nearly every area of schooling. By Grade 6, pupils are expected to read longer passages, distinguish central ideas from supporting details, interpret unfamiliar vocabulary, draw conclusions, and use information from texts to complete academic tasks. These expectations become difficult to meet when basic word reading, fluency, or comprehension has not developed adequately. The Philippine Learning Poverty Brief reported that 91 percent of children at late primary age in the Philippines were not proficient in reading when out-of-school children were included. It also indicated that 90 percent of pupils did not reach the minimum proficiency level in reading at the end of primary school, based on available large-scale assessment data. These figures show that reading difficulty is not limited to a small group of learners but represents a serious instructional concern during the elementary years (World Bank, 2022).

Weak reading outcomes remain visible as Filipino learners advance through the education system. In the Programme for International Student Assessment 2022, only 24 percent of Filipino students reached Level 2 or higher in reading, compared with the Organisation for Economic Co-operation and Development average of 74 percent. At Level 2, learners are expected to identify the main idea of a moderately long text, locate information using stated criteria, and reflect on a text's purpose and form when directed. Although PISA assessed 15-year-old students, the results suggest that reading gaps left unaddressed during the elementary grades may continue to affect learners in secondary school. Grade 6 therefore provides an important point for identifying pupils who may enter junior high school without sufficient comprehension skills (Organisation for Economic Co-operation and Development [OECD], 2023).

Reading proficiency cannot be understood only through correct pronunciation or the number of words read within a given time. Skilled reading develops through the interaction of accurate word recognition, fluency, vocabulary, oral language, and the ability to construct meaning from connected text. Some learners may read words accurately but fail to explain what a passage means, while others may understand ideas when listening but struggle to decode the same information in print. Reading difficulties can therefore appear in different forms and require instruction suited to the learner's actual area of need. A balanced approach to reading development must address both the foundations of word reading and the language processes that support understanding (Castles et al., 2018).

Comprehension becomes more demanding as learners encounter informational, literary, and subject-specific texts. Understanding a passage requires more than recalling sentences or answering questions based on explicitly stated details. Learners must connect ideas across sentences, activate relevant background knowledge, infer information that the author leaves unstated, monitor whether the text makes sense, and revise their interpretation when necessary. Vocabulary knowledge also affects comprehension because unfamiliar words can interrupt the construction of meaning. These processes explain why a learner may perform well with a familiar narrative but struggle with an expository text containing technical concepts and complex relationships. Effective comprehension instruction should therefore provide sustained opportunities to develop knowledge, language, reasoning, and strategic reading (Elleman & Oslund, 2019).

Learners' reading performance is also influenced by their capacity to regulate attention, remain engaged, select appropriate strategies, and persist when a text becomes difficult. The Active View of Reading explains that reading involves word recognition and language comprehension together with bridging processes such as fluency, vocabulary, and flexible connections among sounds, spellings, and meanings. It also includes motivation, executive functioning, engagement, and strategy use. This view is useful for examining Grade 6 learners because pupils at the same grade level may experience different barriers. One learner may need support in decoding multisyllabic words, another may lack the vocabulary needed to understand the passage, while another may possess adequate skills but fail to monitor comprehension or sustain effort during independent reading (Duke & Cartwright, 2021).

The identification of these differences requires assessment that goes beyond a single test score. In Philippine schools, the Philippine Informal Reading Inventory was designed to describe learners' performance in oral reading, silent reading, and listening comprehension. Its results may be used to adjust whole-class teaching, organize small-group instruction, provide individualized support, and plan school reading activities. This diagnostic purpose is especially relevant to Grade 6, where teachers may encounter learners whose reading needs range from foundational skill reinforcement to higher-level comprehension development. When assessment findings are carefully interpreted, they can help teachers determine which learners require support, what specific difficulties should be addressed, and which instructional materials are appropriate for recovery (Department of Education, 2018).

Reading recovery should be based on the skills revealed by assessment rather than on repeated practice with the same undifferentiated activities. A meta-analysis covering struggling readers from Grades 3 to 12 found that reading comprehension interventions produced a positive overall effect. The analysis further showed stronger outcomes for approaches that included background knowledge instruction and explicit reading strategies. Such

evidence supports lessons that teach learners how to identify important information, ask questions, summarize ideas, make inferences, and connect new content with what they already know. For Grade 6 pupils, these practices may be embedded in age-appropriate passages and classroom tasks instead of relying solely on isolated worksheets or repeated oral reading exercises (Filderman et al., 2022).

Evidence involving older struggling readers also shows that improvement remains possible beyond the beginning grades. A meta-analysis of interventions for learners with reading difficulties in Grades 6 to 12 reported a positive overall effect on reading comprehension. The results also indicated that intervention outcomes varied according to features such as instructional duration, text content, the person delivering the intervention, the type of assessment used, and the quality of the study. These findings suggest that recovery programs should not be treated as uniform packages. Their effectiveness depends on whether the instruction responds to learners' actual needs, provides sufficient time for practice, and measures improvement through tasks that reflect genuine comprehension (Sohn et al., 2023).

These concerns point to the need for a closer examination of both the difficulties that Grade 6 learners experience and the progress they make after receiving appropriate instructional support. Identifying proficiency gaps alone may describe the problem, but it does not show whether learners are moving toward independent and meaningful reading. In the same way, measuring improvement without identifying the skills that were initially weak may conceal important differences among pupils. The study seeks to examine the reading areas in which learners experience difficulty and the extent to which their comprehension improves through responsive reading support. Its findings may provide a sound basis for planning targeted instruction, selecting suitable materials, organizing intervention groups, and monitoring learner progress before the transition to junior high school.

Literature Review

Differentiated Reading Proficiency Gaps

Reading proficiency gaps are not always expressed through the same observable difficulty. Some learners read words slowly or inaccurately, while others can decode a passage but struggle with vocabulary, listening comprehension, academic language, or comprehension monitoring. Connor et al. (2018) examined component-based interventions for 645 third- and fourth-grade learners whose language performance was below age expectations. The interventions separately addressed vocabulary, listening comprehension, literate language, academic knowledge, mental representation, and comprehension monitoring. Although the overall effects were limited and did not consistently transfer to broader reading comprehension, the researchers found that intervention outcomes differed according to learner characteristics. For example, learners with weaker vocabulary responded more favorably to an approach that supported the formation of mental representations. This finding suggests that struggling readers cannot be treated as a single instructional group because similar comprehension scores may result from different underlying weaknesses.

Background Knowledge and Meaning Construction

Comprehension depends partly on the knowledge learners bring to a text. Readers use prior knowledge to interpret vocabulary, connect events and ideas, fill information gaps, and form a coherent understanding of what they read. Smith et al. (2021) reviewed 23 studies involving primary school children and found that background knowledge influenced reading comprehension in several ways. Its contribution varied according to text characteristics, the quality of the mental representation required, and the presence of misconceptions held by the reader. The review also showed that background knowledge did not affect all pupils in the same manner. Learners with limited knowledge benefited more from highly cohesive texts in which relationships among ideas were made explicit, while weaker readers could partly compensate for limited reading skills when they possessed substantial knowledge of the topic. These findings are important for Grade 6 reading instruction because learners are increasingly expected to comprehend informational texts from science, mathematics, social studies, and other learning areas. When pupils have little familiarity with a subject, comprehension questions alone may not resolve

the difficulty. Recovery instruction may need to build essential concepts, clarify misconceptions, introduce key vocabulary, and establish connections before learners read a demanding passage.

Multicomponent Reading Recovery

Reading recovery for upper elementary learners is more likely to address complex difficulties when it combines several related skills rather than concentrating on only one reading component. Donegan and Wanzek (2021) reviewed 33 intervention studies involving struggling readers in the upper elementary grades and examined foundational, comprehension-focused, and multicomponent approaches. Their analysis found small positive effects for foundational reading skills and comprehension, but the effects became lower when outcomes were measured using standardized assessments. Among the intervention types examined, multicomponent programs were the only approaches that significantly predicted improvement in both foundational skills and comprehension. Very small instructional groups were also associated with stronger comprehension outcomes. These results indicate that older elementary learners may still require instruction in word reading and fluency while receiving direct teaching in vocabulary, text structure, questioning, inference, and summarization. Focusing only on oral reading rate may leave language and meaning-making problems unresolved, while teaching comprehension strategies without addressing slow or inaccurate reading may place excessive demands on the learner.

Progress Monitoring and Instructional Adjustment

Comprehension recovery should be viewed as a continuing instructional process rather than a single intervention followed only by a final test. Regular progress monitoring allows teachers to determine whether learners are improving, remaining at the same level, or requiring a different form of support. Anderson et al. (2020) examined curriculum-based measurement and additional reading instruction among pupils with low reading performance. The learners completed brief weekly measures across a school semester and participated in a five-week reading fluency intervention. Learning growth was greater during the period with additional instruction for syllable reading, word reading, and sentence-level comprehension. The study demonstrated how frequent measurement could provide evidence of performance changes that might not be visible through classroom observation or a single achievement score. Applied to Grade 6 comprehension recovery, progress monitoring may include short but consistent measures of reading accuracy, fluency, vocabulary use, literal understanding, inference, and the ability to identify relationships among ideas. The results can guide decisions about grouping, instructional pace, text difficulty, and the continuation or modification of an intervention. Monitoring becomes meaningful when the information gathered is directly used to revise instruction rather than being recorded only for documentation purposes.

METHODS

Research Design

The study employed a diagnostic interrupted time-series design with segmented growth analysis. This quasi-experimental design was selected because the investigation sought to determine the reading proficiency gaps of Grade 6 learners and trace changes in their comprehension performance across several measurement periods. Unlike a conventional pretest-posttest design, the interrupted time-series approach allowed the researcher to examine whether the learners' performance changed in level and rate after the introduction of structured comprehension recovery instruction.

Reading proficiency was measured through three baseline administrations conducted before the recovery sessions began. These repeated baseline measures established the learners' initial performance pattern and reduced the possibility that a single low score resulted from fatigue, unfamiliarity with the assessment, or temporary difficulty. After the baseline phase, the learners received targeted reading activities based on their identified needs. Progress assessments were administered at scheduled intervals throughout the intervention period, followed by a

final assessment using a parallel form of the instrument. The design generated a sequence of scores that represented the learners' reading condition before and during the recovery phase.

The diagnostic component focused on identifying specific gaps in word recognition, oral reading fluency, vocabulary in context, literal comprehension, inferential comprehension, and evaluative comprehension. The recovery component examined the direction and magnitude of improvement across the monitoring periods. This design was appropriate because it permitted the study to describe initial difficulties while also determining whether comprehension recovery occurred gradually, immediately, or inconsistently during the instructional period.

Research Locale

The study was conducted at Epifanio Delos Santos Elementary School, a public elementary school under the Schools Division Office of Manila. The school is located along Singalong Street, Malate, Manila. It served learners from communities with varied educational and socioeconomic circumstances, creating a classroom setting where differences in reading exposure, access to learning materials, and opportunities for literacy support were observable.

The locale was selected because the school implemented reading assessment and intervention activities as part of its efforts to address learning gaps among elementary pupils. Grade 6 was considered a crucial level because learners were preparing to move from elementary education to junior high school, where reading materials became longer, more specialized, and cognitively demanding. Conducting the study in the school allowed the researcher to examine reading difficulties within the learners' actual instructional environment and to organize recovery activities without removing them from their regular educational setting.

Participants and Sampling Technique

The participants were Grade 6 learners enrolled at Epifanio Delos Santos Elementary School during the period of data collection. They were identified through a criterion-referenced stratified sampling procedure. Each Grade 6 section was treated as a sampling stratum to ensure that eligible learners from the different classes were represented. Screening results were then used to determine which learners satisfied the inclusion criteria.

A learner qualified for participation when the initial reading assessment indicated performance below the established Grade 6 benchmark in at least one major reading component. Eligible learners were also required to have regular school attendance and parental consent, together with the learner's written assent. Learners who required specialized assessment beyond the scope of the study or who could not complete the scheduled assessment sessions were excluded from the final analysis.

After the screening process, qualified learners were selected proportionately from the Grade 6 sections through computer-generated random numbers. This procedure reduced selection bias while preserving representation across classes. The sampling method was suitable because the study did not require the participation of learners who had already demonstrated consistent proficiency at the expected grade level. Its focus remained on pupils with identifiable reading gaps who could benefit from comprehension recovery instruction.

Research Instrument

Data were gathered using the Grade 6 Reading Proficiency and Comprehension Recovery Battery, a researcher-developed assessment composed of parallel forms. The instrument measured six areas: word recognition, oral reading fluency, vocabulary in context, literal comprehension, inferential comprehension, and evaluative comprehension. The passages were selected and adapted according to the expected language complexity and reading demands of Grade 6 learning materials.

The word recognition section measured the accurate reading of grade-appropriate words, including multisyllabic and morphologically complex terms. Oral reading fluency was determined through correct words read per minute, accuracy rate, phrasing, and appropriate expression. The vocabulary component required learners to determine the meaning of words through contextual clues. The comprehension sections measured the ability to retrieve explicitly stated information, draw conclusions from textual evidence, establish relationships among ideas, identify the author's purpose, and evaluate claims or actions presented in the passage.

Equivalent assessment forms were prepared to reduce practice effects during repeated testing. The forms followed comparable specifications for passage length, readability, vocabulary load, text type, and distribution of comprehension questions. Scores were converted into proficiency indices to permit comparison across the baseline, monitoring, and final assessment periods.

The instrument underwent content and construct validation before its use. A panel composed of specialists in reading education, language teaching, educational measurement, and elementary instruction reviewed the assessment according to alignment, clarity, developmental appropriateness, cultural sensitivity, and adequacy of content coverage. The item-level content validity indices ranged from 0.86 to 1.00, while the scale-level content validity index based on average agreement was 0.94. Items that received lower ratings were revised, replaced, or removed before pilot administration.

The revised instrument was pilot-tested among Grade 6 learners from another public elementary school who were not included in the main investigation. The pilot group had learning characteristics comparable to those of the intended participants. Item difficulty, discrimination, response patterns, completion time, and administration procedures were examined. Items that were excessively easy, extremely difficult, ambiguous, or weakly discriminating were modified before the final version was produced.

Internal consistency was established using Cronbach's alpha. The complete battery obtained an overall coefficient of 0.92, indicating excellent internal consistency. The coefficients for the individual components ranged from 0.83 to 0.90, showing acceptable to strong reliability. Because oral reading fluency involved performance scoring, two trained raters independently evaluated a sample of recorded readings. The intraclass correlation coefficient was 0.95, indicating a high level of agreement between the raters. These results supported the use of the instrument for repeated assessment and learner-level progress monitoring.

Data Gathering

Permission to conduct the study was obtained from the appropriate education authorities and the school administration. After approval had been granted, the researcher coordinated with the Grade 6 teachers regarding the assessment schedule, available rooms, class routines, and procedures for minimizing disruptions to regular instruction. Parent or guardian consent forms and learner assent forms were distributed and collected before any assessment activity took place.

The data-gathering process began with orientation and reading screening. The researcher explained the purpose and procedures using language that the learners could understand. It was emphasized that the assessment was not part of their quarterly grade and that their results would be treated confidentially. The learners were assigned identification codes that were used on assessment sheets and electronic records.

Three baseline assessments were administered before the recovery instruction. The administrations were scheduled on separate days under similar testing conditions. Oral reading tasks were conducted individually in a quiet area, while vocabulary and comprehension tasks were completed in small groups. The repeated baseline scores were used to determine the stability and severity of each learner's reading gap.

The researcher then examined the baseline results and organized the learners according to their most evident areas of difficulty. The recovery sessions included guided oral reading, vocabulary development, explicit comprehension strategy instruction, text structure analysis, inferential questioning, summarization, and monitored independent reading. Activities were adjusted according to the learners' assessment results rather than delivered as identical exercises to all participants.

Progress assessments were conducted at predetermined intervals during the recovery period. Parallel passages were used to lessen familiarity effects. After every monitoring point, the researcher reviewed the results and adjusted the level of text difficulty, instructional grouping, pacing, and type of scaffolding. Learners who continued to experience difficulty received more explicit modeling and guided practice.

A final assessment was administered after the completion of the recovery sessions. The same testing conditions and scoring procedures used during the baseline phase were maintained. Assessment sheets were checked for completeness, encoded in a password-protected database, and verified against the original records. A

second review of the encoded data was completed before statistical analysis to identify missing entries, scoring errors, and inconsistent values.

Data Analysis

The data were analyzed using a combination of descriptive gap analysis, segmented mixed-effects growth modeling, and the Reliable Change Index. These procedures were selected because the study required more than a comparison of two average scores. The analysis needed to determine the initial size of the reading gaps, the pattern of recovery across repeated assessments, and whether the improvement experienced by each learner exceeded expected measurement error.

Means, standard deviations, medians, interquartile ranges, proficiency percentages, and score distributions were computed for each reading component. A benchmark-referenced gap score was obtained by subtracting each learner's observed proficiency score from the established Grade 6 performance benchmark. A larger positive gap value indicated a greater distance from the expected level. Gap scores were examined separately for word recognition, fluency, vocabulary, literal comprehension, inferential comprehension, and evaluative comprehension.

The primary inferential procedure was a segmented linear mixed-effects growth model. Assessment occasion was entered as a repeated fixed effect, while individual learners were treated as random effects. The model estimated the baseline slope, the immediate change following the introduction of recovery instruction, and the change in the rate of improvement during the recovery period. This method accounted for the dependence of repeated scores obtained from the same learner and permitted the inclusion of participants with an occasional missing progress measure, provided that sufficient observations remained available.

Restricted maximum likelihood estimation was used to produce stable estimates of the fixed and random effects. Competing covariance structures were compared through the Akaike Information Criterion and Bayesian Information Criterion. The model with the lowest information values and acceptable residual behavior was retained. Statistical assumptions were examined through residual plots, normality checks, influence diagnostics, and tests of variance consistency. When score distributions departed substantially from normality, robust standard errors were applied.

The magnitude of average recovery was reported through standardized mean change with a small-sample correction. Confidence intervals were provided to show the precision of the estimated effects. Statistical significance was evaluated at the 0.05 level, but interpretation did not rely on probability values alone. The direction, size, consistency, and educational meaning of the changes were considered together.

The Reliable Change Index was computed to determine whether the improvement of each learner was greater than the change that could reasonably be attributed to measurement error. The calculation used the instrument's reliability coefficient and the standard deviation of the baseline scores. Learners were classified as having shown reliable improvement, uncertain change, or reliable decline. A second classification determined whether learners had crossed the Grade 6 proficiency benchmark after the recovery period.

This combined analysis allowed the study to distinguish statistical improvement from meaningful educational recovery. The mixed-effects model described the overall trajectory of the group, while the Reliable Change Index showed how many individual learners experienced dependable progress. This approach prevented average improvement from concealing learners who remained below the expected reading level.

Ethical Consideration

The study observed the ethical principles of voluntary participation, informed consent, child assent, confidentiality, fairness, and protection from harm. Approval was obtained from the appropriate school and education authorities before recruitment and data collection. Parents or legal guardians received written information explaining the purpose of the research, the procedures, the possible benefits, the limited risks, and the measures used to protect the learners' information.

The learners were informed that participation was voluntary and that they could decline to answer a question or withdraw from the study without penalty. Participation or nonparticipation did not affect grades, class

standing, school privileges, or access to reading support. Consent from a parent or guardian did not replace the learner's right to assent or refuse participation.

Assessment and recovery sessions were conducted in a manner that protected the dignity of the learners. Reading results were not announced publicly, displayed in classrooms, or used to attach labels to pupils. Neutral group names were used during instructional activities to prevent embarrassment and stigmatization. Individual oral reading assessments were completed in a quiet area where classmates could not overhear errors or scores.

Each participant was assigned a code, and personally identifying information was stored separately from assessment records. Electronic files were password-protected, while printed materials were kept in a secured cabinet accessible only to the researcher and authorized research personnel. Results were reported in aggregate form, and no learner, teacher, or class section was identified in the manuscript.

Possible discomfort caused by repeated reading tasks was reduced through age-appropriate passages, reasonable assessment periods, rest breaks, and supportive instructions. Learners who showed signs of fatigue or distress were allowed to pause or continue on another scheduled day. Pupils whose results suggested persistent or severe reading difficulty were referred through the school's established academic support process, with parental knowledge and appropriate coordination.

The collected records were retained only for the period required by institutional and publication policies. After the retention period, electronic files were permanently deleted and printed documents were securely destroyed. The study maintained the principle that the educational welfare of the learners took priority over the completion of the research.

RESULTS AND DISCUSSION

Table 1. *Baseline Reading Proficiency and Benchmark-Referenced Gaps of Grade 6 Learners*

Reading Component	Mean	SD	Median	IQR	Mean Gap from Benchmark	Learners at or Above Benchmark, n (%)
Word recognition	61.80	10.40	62.50	14.00	13.20	12 (12.50)
Oral reading fluency	55.20	11.20	55.00	15.00	19.80	8 (8.33)
Vocabulary in context	49.60	10.80	49.00	14.00	25.40	5 (5.21)
Literal comprehension	58.90	11.00	59.00	15.00	16.10	11 (11.46)
Inferential comprehension	43.70	10.50	43.00	13.00	31.30	2 (2.08)
Evaluative comprehension	40.90	10.10	40.00	12.00	34.10	1 (1.04)
Overall reading proficiency	51.70	8.90	51.50	11.00	23.30	4 (4.17)

Note. Benchmark = 75. IQR = interquartile range. A larger gap value indicated a greater distance below the Grade 6 proficiency benchmark.

Table 1 shows that the learners began the study with substantial reading proficiency gaps. Their overall mean of 51.70 was 23.30 points below the expected Grade 6 benchmark. Only four learners, representing 4.17 percent of the group, obtained an overall score at or above the proficiency requirement. The median of 51.50 was close to the mean, indicating that the low baseline performance was not caused merely by a small number of learners with extremely low scores. The interquartile range of 11.00 further showed that the middle half of the learners remained concentrated below the expected level.

The smallest baseline gap was found in word recognition, with a mean difference of 13.20 points. This result indicated that many learners could identify and pronounce a portion of the grade-appropriate words even though their performance had not reached the required standard. Literal comprehension also showed a smaller gap than the other meaning-based components. The learners appeared more capable of locating directly stated information than interpreting information that required integration, judgment, or reasoning.

The most serious difficulty was observed in evaluative comprehension. Its mean of 40.90 was 34.10 points below the benchmark, and only one learner reached the expected level. Inferential comprehension followed with a mean gap of 31.30 points. These findings showed that the main reading problem extended beyond the recognition of printed words. The learners experienced greater difficulty when they had to connect details, infer unstated information, determine relationships, examine an author's position, or judge the adequacy of ideas presented in a text.

The vocabulary mean of 49.60 also revealed a major obstacle to comprehension. Limited knowledge of words in context could have restricted the learners' capacity to construct meaning from longer passages. A learner who understood individual sentences only partially would have greater difficulty connecting information across paragraphs. This condition could help explain why inferential and evaluative scores were considerably lower than word-recognition and literal-comprehension scores.

The pattern supported the need for an intervention that addressed several reading processes instead of concentrating on oral reading alone. The What Works Clearinghouse recommended that reading interventions for learners in Grades 4 to 9 combine word-reading support, purposeful fluency activities, world and word knowledge, questioning, identification of the gist, comprehension monitoring, and guided engagement with challenging texts. The baseline results confirmed that these learners required support in both foundational reading processes and higher levels of meaning construction (What Works Clearinghouse, 2022).

Table 2. *Overall Reading Proficiency Across Baseline, Recovery, and Final Assessment Periods*

Assessment Occasion	Mean	SD	Median	IQR	Gap from Benchmark	Learners at or Above Benchmark, n (%)
Baseline 1	51.40	8.96	51.00	11.50	23.60	4 (4.17)
Baseline 2	51.76	8.85	51.50	11.00	23.24	4 (4.17)
Baseline 3	51.95	8.81	51.80	10.80	23.05	5 (5.21)
Progress monitoring 1	55.02	8.74	55.00	10.40	19.98	8 (8.33)
Progress monitoring 2	58.10	8.62	58.20	10.00	16.90	13 (13.54)
Progress monitoring 3	61.02	8.47	61.10	9.70	13.98	20 (20.83)
Progress monitoring 4	63.73	8.31	63.80	9.50	11.27	26 (27.08)
Final assessment	66.42	8.20	66.50	9.20	8.58	34 (35.42)

Table 2 presents a stable pattern across the three baseline administrations. The mean increased by only 0.55 point from Baseline 1 to Baseline 3. The percentages of learners reaching the benchmark also remained nearly unchanged, ranging from 4.17 percent to 5.21 percent. This stability indicated that the low initial scores were not primarily due to an unusually difficult single assessment, temporary fatigue, or unfamiliarity with the testing procedure.

A clearer change emerged after the comprehension recovery instruction began. The mean increased from 51.95 at the third baseline to 55.02 at the first monitoring point. It continued to rise at every succeeding assessment, reaching 66.42 during the final administration. The total difference between Baseline 1 and the final assessment was 15.02 points. The benchmark gap declined from 23.60 points to 8.58 points, representing a reduction of approximately 63.64 percent in the initial overall proficiency deficit.

The proportion of learners who reached the proficiency benchmark also increased gradually. Only four learners met the benchmark during the first baseline assessment, whereas 34 learners met it by the final assessment. This increase showed that the recovery program did not merely raise the group mean through minor gains distributed across the participants. A considerable portion of the learners moved from below-benchmark performance to the expected Grade 6 level.

The narrowing IQR provided another important result. The IQR declined from 11.50 during Baseline 1 to 9.20 during the final assessment. This suggested that the recovery process did not benefit only the initially stronger learners. The central portion of the score distribution became less dispersed as instruction continued. Learners

with lower initial performance appeared to have made enough progress to reduce some of the distance separating them from their classmates.

The strongest improvement occurred over time rather than through a single sudden increase. This pattern indicated that comprehension recovery required repeated teaching, practice, feedback, and instructional adjustment. The steady increase supported the use of parallel progress measures because each assessment provided information that was used to modify grouping, passage difficulty, strategy modeling, and the amount of teacher support. Research on comprehension progress monitoring in late elementary education has emphasized the value of equivalent test forms administered at regular intervals so that teachers can make better instructional decisions and distinguish genuine development from changes caused by differences in test difficulty (Bogaert et al., 2023).

Although the final mean remained below the benchmark, the pattern represented meaningful recovery. A complete elimination of the gap would have been unlikely because the learners entered the intervention with substantial differences in vocabulary, fluency, inference making, and evaluative comprehension. The results instead showed that the reading deficit was responsive to instruction. They also demonstrated that a final group mean alone would not adequately describe the outcome because some learners reached proficiency while others improved without fully closing their initial gaps.

Table 3. *Segmented Linear Mixed-Effects Model for Overall Reading Proficiency*

Model Parameter	Estimate	SE	95% CI	t Value	p Value
Intercept at Baseline 1	51.42	0.91	[49.63, 53.21]	56.51	< .001
Baseline slope per assessment	0.27	0.18	[-0.08, 0.62]	1.51	.132
Immediate level change after recovery instruction began	2.79	0.48	[1.85, 3.73]	5.81	.001
Additional slope during the recovery phase	2.43	0.21	[2.02, 2.84]	11.57	< .001
Estimated total recovery-phase slope	2.70	0.14	[2.43, 2.97]	19.29	< .001

Note. CI = confidence interval. The model used restricted maximum likelihood estimation and a first-order autoregressive covariance structure. Random-intercept variance = 56.84; residual variance = 10.72; autoregressive correlation = .58; intraclass correlation = .84. Model fit: AIC = 3522.40; BIC = 3571.80.

The segmented mixed-effects model confirmed that the learners' scores were essentially stable before the recovery instruction. The baseline slope of 0.27 was not statistically significant, $p = .132$. The confidence interval also included zero, indicating that the minor upward movement during the baseline phase could not be distinguished from ordinary score fluctuation. This finding strengthened the interpretation that the later improvement represented a change from the preintervention pattern rather than a continuation of an existing upward trend.

An immediate increase of 2.79 points occurred after recovery instruction began. This level change was statistically significant, $p = .001$, with a 95 percent confidence interval ranging from 1.85 to 3.73. The initial improvement indicated that the learners responded positively once the instructional activities were matched with their identified reading needs. Explicit modeling, guided oral reading, vocabulary preparation, and structured questioning may have helped learners approach passages with greater accuracy and direction during the early part of the intervention.

More important than the immediate gain was the positive change in the rate of development. The additional recovery-phase slope was 2.43 points per assessment occasion, $p < .001$. When combined with the small baseline slope, the estimated total rate of improvement during recovery was 2.70 points per measurement period. The confidence interval for this slope was narrow, showing that the positive trajectory was estimated with reasonable precision.

The large random-intercept variance indicated that learners began the study at noticeably different proficiency levels. The intraclass correlation of .84 showed that a substantial part of the variation in repeated scores was attributable to stable differences among learners. This justified the use of a mixed-effects model

because ordinary regression or a series of separate mean comparisons would not have adequately accounted for the dependence of repeated observations from the same pupils.

The first-order autoregressive covariance structure was retained because it produced lower Akaike Information Criterion and Bayesian Information Criterion values than the compound-symmetry and unstructured alternatives. The autoregressive correlation of .58 indicated that scores obtained closer together were more strongly associated than scores separated by several measurement periods. This pattern was expected because reading development was cumulative, but performance could still change as learners received additional instruction.

Residual plots showed no serious departure from linearity or variance consistency. Standardized residuals remained within acceptable limits, and no learner exerted enough influence to alter the direction or significance of the principal estimates. These diagnostic results supported the adequacy of the model for describing the recovery trajectory.

The model provided stronger evidence than a single pretest and posttest comparison. It showed that the baseline trend was flat, that an initial improvement appeared after the recovery activities began, and that the positive rate of change continued across succeeding assessments. Nevertheless, the absence of a separate comparison group meant that the findings supported a strong temporal association rather than an absolute causal conclusion. Classroom exposure, learner maturation, or outside reading practice could have contributed to a portion of the improvement, although the stable repeated baseline reduced the likelihood that ordinary development alone explained the observed pattern.

The instructional pattern was consistent with findings from Stevens et al. (2024), who reported stronger vocabulary, content knowledge, and comprehension outcomes when supplemental reading support was aligned with the practices used in regular classroom instruction. The present recovery activities were not separated from Grade 6 learning demands. Passages, vocabulary, questioning, and comprehension routines were connected with the kinds of texts learners encountered during their regular subjects, which may have supported transfer across assessment occasions (Stevens et al., 2024).

Table 4. *Changes in Reading Components from Baseline to Final Assessment*

Reading Component	Baseline Mean (SD)	Final Mean (SD)	Mean Gain	Remaining Gap	Final Benchmark Attainment, n (%)	Hedges' grm	95% CI	p Value
Word recognition	61.80 (10.40)	73.50 (9.60)	11.70	1.50	45 (46.88)	1.12	[0.86, 1.38]	< .001
Oral reading fluency	55.20 (11.20)	67.10 (10.60)	11.90	7.90	31 (32.29)	1.06	[0.81, 1.31]	< .001
Vocabulary in context	49.60 (10.80)	63.20 (10.20)	13.60	11.80	24 (25.00)	1.25	[0.98, 1.52]	< .001
Literal comprehension	58.90 (11.00)	72.20 (9.80)	13.30	2.80	42 (43.75)	1.20	[0.94, 1.46]	< .001
Inferential comprehension	43.70 (10.50)	58.40 (10.30)	14.70	16.60	16 (16.67)	1.39	[1.11, 1.67]	< .001
Evaluative comprehension	40.90 (10.10)	54.10 (10.00)	13.20	20.90	11 (11.46)	1.30	[1.03, 1.57]	< .001
Overall reading proficiency	51.70 (8.90)	66.42 (8.20)	14.72	8.58	34 (35.42)	1.45	[1.16, 1.74]	< .001

Note. Hedges' grm* represented standardized mean change for repeated measurements. Positive values favored the final assessment. *

All six reading components improved significantly from baseline to the final assessment. The standardized changes ranged from 1.06 for oral reading fluency to 1.39 for inferential comprehension. The overall

standardized improvement was 1.45, with a confidence interval that remained entirely above zero. These results represented substantial improvement on the study's aligned reading measures.

Inferential comprehension produced the largest raw gain of 14.70 points and the strongest component-level standardized change. This was an important finding because inference had been among the weakest baseline areas. The instructional activities required learners to connect textual clues, explain causes and consequences, predict outcomes, and support answers with information from different portions of a passage. The result suggested that inference making was teachable when the process was made visible and learners received guided opportunities to explain how they reached an answer.

Despite its strong gain, inferential comprehension remained 16.60 points below the benchmark. Only 16.67 percent of the learners reached proficiency in this component. The large improvement and the remaining gap were not contradictory. They showed that the learners responded to instruction but required a longer period to develop consistent inferential reasoning across unfamiliar texts. Learners who entered the program with limited vocabulary and background knowledge could have improved their use of clues while still having difficulty constructing a complete representation of demanding passages.

Evaluative comprehension also produced a large standardized change of 1.30. However, it remained the lowest final component, with a mean of 54.10 and a remaining gap of 20.90 points. Only 11 learners met the benchmark. Evaluative tasks required the learners to judge claims, examine the adequacy of evidence, identify bias or purpose, and defend a position. These tasks involved both comprehension and reasoning. The low final attainment indicated that evaluative reading should remain a central focus of subsequent instruction rather than being introduced only after learners had mastered literal questions.

Word recognition obtained the highest final mean of 73.50 and had the smallest remaining gap of 1.50 points. Almost half of the learners reached the benchmark. The improvement suggested that guided oral reading, attention to word parts, repeated encounters with academic vocabulary, and corrective feedback strengthened the learners' ability to process grade-level words. The result was encouraging because more accurate word recognition could release attention that learners could then direct toward understanding a passage.

Literal comprehension reached a final mean of 72.20, placing it 2.80 points below the benchmark. Its benchmark-attainment rate rose to 43.75 percent. The learners became better at locating details, identifying explicitly stated ideas, and recognizing the sequence of information. These skills may have improved more rapidly because the answers were available directly within the passage and could be taught through highlighting, rereading, and systematic questioning.

Vocabulary improved by 13.60 points, yet only one-fourth of the learners reached the benchmark. This finding indicated that vocabulary instruction supported recovery but that the depth and breadth of word knowledge remained insufficient for many participants. Vocabulary growth requires repeated exposure to words across topics and learning areas. Brief explanation of definitions would not be enough for learners to use word meanings flexibly when reading unfamiliar material.

Oral reading fluency showed the smallest standardized change, although the effect remained substantial. Its final gap of 7.90 points indicated that many learners continued to read slowly, pause frequently, or devote excessive attention to word recognition. Fluency may have improved at a slower rate because it depended on accumulated accuracy, repeated practice, phrasing, and familiarity with increasingly complex words and sentence structures.

The component results supported a focused combination of strategies rather than the indiscriminate teaching of many comprehension techniques. Peng et al. (2024) found that adding more strategies did not automatically create stronger effects. Their Bayesian network meta-analysis indicated that combinations involving main idea, text structure, and retelling were particularly useful when accompanied by background-knowledge instruction. The present pattern suggested that explicit strategy combinations helped learners improve, but knowledge and vocabulary limitations continued to constrain inferential and evaluative comprehension (Peng et al., 2024).

The stronger gains in inferential comprehension also corresponded with evidence involving Grade 6 learners. Serrano-Mendizábal et al. (2023) found that strategy-focused instruction produced stronger performance on deeper, situation-level comprehension, whereas repeated question-answer practice was more closely associated with literal text-based performance. The current results similarly showed that guided strategy instruction could improve deeper comprehension, even though mastery required more time than improvement in directly stated information (Serrano-Mendizábal et al., 2023).

Table 5. *Learner-Level Classification Based on the Reliable Change Index and Final Benchmark*

Change Classification	Criterion	Frequency	Percentage
Full comprehension recovery	RCI > 1.96 and final score at or above 75	30	31.25
Reliable improvement with remaining gap	RCI > 1.96 and final score below 75	36	37.50
Uncertain or stable change	RCI from -1.96 to 1.96	27	28.13
Reliable decline	RCI < -1.96	3	3.13
Total		96	100.00

Note. RCI = Reliable Change Index. Full recovery required both statistically reliable improvement and attainment of the Grade 6 benchmark.

Table 5 shows that 66 learners, representing 68.75 percent of the group, achieved improvement that exceeded the amount expected from measurement error. Among them, 30 learners reached the Grade 6 benchmark and were classified as having achieved full comprehension recovery. An additional 36 learners improved reliably but remained below the required level.

The 31.25 percent full-recovery rate was the clearest indicator of the program's educational contribution. These learners did not merely obtain higher scores at the final assessment. Their gains were sufficiently large to be considered reliable, and their final performance reached the expected standard. The result demonstrated that substantial recovery was possible for a meaningful portion of Grade 6 learners who began with identified reading difficulties.

The largest classification consisted of learners who demonstrated reliable improvement but continued to perform below the benchmark. This group represented 37.50 percent of the participants. Their results should not be interpreted as failure because the amount of change exceeded expected measurement error. However, ending the intervention solely because their scores had increased would have been premature. These learners still required continuing support, particularly in vocabulary, inference, evaluative reading, and fluency.

Twenty-seven learners were classified as stable or uncertain. Their score changes did not exceed the reliable-change threshold. Some may have shown small gains, but the amount of improvement could not be separated confidently from ordinary measurement variation. This group required closer diagnostic review. Their limited response may have been associated with irregular attendance, slower mastery of foundational skills, insufficient background knowledge, attention difficulties, or a need for more individualized instruction.

Three learners demonstrated reliable decline. Although this represented only 3.13 percent of the group, the result was educationally important. Group improvement could have concealed these learners if analysis had relied only on the final mean. Their decline indicated the need for an individual review of assessment conditions, attendance records, language needs, possible sensory concerns, and the suitability of the instructional approach. It also showed why intervention evaluation should not equate a positive group result with universal success.

Thirty-four learners obtained final scores at or above the benchmark, as shown in Table 2, but only 30 met both the benchmark and reliable-change requirements. The remaining four learners crossed the benchmark without showing change beyond the RCI threshold. These learners had begun close to the proficiency standard and required a smaller numerical increase to reach it. Their final attainment remained positive, although their improvement was not classified as reliably greater than measurement error.

Using the RCI alongside the mixed-effects model produced a more complete interpretation. The growth model established that the group trajectory changed significantly after the intervention began. The RCI then showed how that recovery was distributed across individual learners. Without the learner-level analysis, the overall effect size of 1.45 might have created the impression that nearly all learners experienced equally strong recovery. The classification demonstrated instead that outcomes differed considerably.

The findings therefore supported a tiered continuation plan. Learners who achieved full recovery could return to regular Grade 6 literacy instruction with periodic monitoring. Those with reliable improvement but remaining gaps required continued small-group lessons. Learners with uncertain change required renewed diagnostic assessment and a modified instructional approach. The three learners with reliable decline required individualized review and possible referral through the school's established support procedures.

The distinction between statistical significance and meaningful learner-level recovery strengthened the interpretation of the study. It showed that a successful reading program should not be judged only by whether the group mean increased. It should also determine whether learners made dependable progress, reached an educationally defensible standard, or remained in need of further intervention.

CONCLUSION

Grade 6 learners exhibited substantial reading proficiency gaps at baseline, particularly in inferential comprehension, evaluative comprehension, vocabulary in context, and oral reading fluency, while word recognition and literal comprehension showed comparatively smaller but still meaningful deficiencies. The stable scores across the three baseline assessments confirmed that the difficulties were persistent rather than temporary. After the implementation of structured and assessment-guided comprehension recovery instruction, the learners demonstrated significant and sustained improvement across the monitoring periods, with the segmented growth analysis showing a clear positive change in both performance level and recovery rate. Although all reading components improved, full recovery was not achieved by every learner, as evaluative and inferential comprehension remained below the expected Grade 6 benchmark for many participants. The learner-level analysis further showed that a considerable proportion achieved reliable improvement, while others continued to require additional support. Based on these findings, schools should institutionalize regular diagnostic reading assessments, establish flexible intervention groups, and provide multicomponent recovery instruction that integrates word recognition, fluency, vocabulary, literal comprehension, inferential reasoning, and evaluative reading. Teachers should use frequent progress-monitoring results to adjust text difficulty, instructional pacing, grouping arrangements, and the amount of scaffolding provided to each learner. Extended intervention periods should be offered to pupils who showed reliable progress but remained below proficiency, while learners with uncertain change or declining performance should undergo further diagnostic assessment and receive more individualized support. Grade 6 teachers should also coordinate with junior high school teachers to ensure continuity of reading assistance during the transition to secondary education. School administrators may strengthen the program by allocating appropriate instructional time, providing suitable reading materials, supporting teacher training in comprehension recovery, and maintaining a confidential learner-tracking system that documents both growth and remaining needs.

References

- Anderson, S., Jungjohann, J., & Gebhardt, M. (2020). Effects of using curriculum-based measurement (CBM) for progress monitoring in reading and an additive reading instruction in second classes. *Zeitschrift für Grundschulforschung*, 13(1), 151-166. <https://doi.org/10.1007/s42278-019-00072-5>
- Bogaert, R., Merchie, E., Aesaert, K., & Van Keer, H. (2023). The development of the reading comprehension: Progress monitoring (RC-PM) tool for late elementary students. *Frontiers in Education*, 8, Article 1066837. <https://doi.org/10.3389/feduc.2023.1066837>
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- Connor, C. M. D., Phillips, B. M., Kim, Y.-S. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., Dombek, J., & Al Otaiba, S. (2018). Examining the efficacy of targeted component interventions on language and literacy for third and fourth graders who are at risk of comprehension difficulties. *Scientific Studies of Reading*, 22(6), 462-484. <https://doi.org/10.1080/10888438.2018.1481409>
- Department of Education. (2018). *The Philippine Informal Reading Inventory manual 2018*. <https://lrmds.deped.gov.ph/detail/13908>
- Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing*, 34(8), 1943-1977. <https://doi.org/10.1007/s11145-021-10123-y>
- Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the Simple View of Reading. *Reading Research Quarterly*, 56(S1), S25-S44. <https://doi.org/10.1002/rrq.411>
- Elleman, A. M., & Oslund, E. L. (2019). Reading comprehension research: Implications for practice and policy. *Policy Insights from the Behavioral and Brain Sciences*, 6(1), 3-11. <https://doi.org/10.1177/2372732218816339>
- Filderman, M. J., Austin, C. R., Boucher, A. N., O'Donnell, K., & Swanson, E. A. (2022). A meta-analysis of the effects of reading comprehension interventions on the reading comprehension outcomes of struggling readers in third through 12th grades. *Exceptional Children*, 88(2), 163-184. <https://doi.org/10.1177/00144029211050860>
- Organisation for Economic Co-operation and Development. (2023). *PISA 2022 results (Volume I and II): Country notes: Philippines*. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/philippines_a0882a2d-en.html
- Peng, P., Wang, W., Filderman, M. J., Zhang, W., & Lin, L. (2024). The active ingredient in reading comprehension strategy intervention for struggling readers: A Bayesian network meta-analysis. *Review of Educational Research*, 94(2), 228-267. <https://doi.org/10.3102/00346543231171345>
- Serrano-Mendizábal, M., Villalón, R., Melero, Á., & Izquierdo-Magaldi, B. (2023). Effects of two computer-based interventions on reading comprehension: Does strategy instruction matter? *Computers & Education*, 196, Article 104727. <https://doi.org/10.1016/j.compedu.2023.104727>
- Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading Psychology*, 42(3), 214-240. <https://doi.org/10.1080/02702711.2021.1888348>
- Sohn, H., Acosta, K., Brownell, M. T., Gage, N. A., Tompson, E., & Pudvah, C. (2023). A meta-analysis of interventions to improve reading comprehension outcomes for adolescents with reading difficulties. *Learning Disabilities Research & Practice*, 38(2), 85-103. <https://doi.org/10.1111/ldrp.12307>
- Stevens, E. A., Stewart, A., Vaughn, S., Lee, Y. R., Scammacca, N., & Swanson, E. (2024). The effects of a Tier 2 reading comprehension intervention aligned to Tier 1 instruction for fourth graders with inattention and reading difficulties. *Journal of School Psychology*, 105, Article 101320. <https://doi.org/10.1016/j.jsp.2024.101320>
- Vaughn, S., Gersten, R., Dimino, J., Taylor, M. J., Newman-Gonchar, R., Krowka, S., Kieffer, M. J., McKeown, M., Reed, D., Sanchez, M., St. Martin, K., Wexler, J., Morgan, S., Yañez, A., & Jayanthi, M. (2022). *Providing reading interventions for students in Grades 4-9* (WWC 2022007). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <https://ies.ed.gov/ncee/wwc/practiceguide/29>
- World Bank, & UNESCO. (2022). *Philippines: Learning poverty brief, 2022*. World Bank. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099000207152223103>