Universitas: Journal of Global Research and Innovation
Excellence. Quality. Reliability. Service | Contributing to the Body of Knowledge

The Reading Reset: Examining the EOSY–BOSY Reading Classification Gap in Project SARA

Sharon M. Jaudian
Head Teacher I, Purificacion Salcedo Gustilo Memorial Elementary School,
Department of Education | PhD in Educational Management
https://doi.org/10.5281/zenodo.23679538

ABSTRACT

This quantitative study examined reading-classification patterns in Project SARA to determine what the available assessment records reveal about the perceived end-of-school-year (EOSY) to beginning-of-school-year (BOSY) reading-classification gap. Five district workbooks were analyzed: Grade 3 pretest and posttest for School Year (SY) 2024–2025, Grade 4 pretest and posttest for SY 2025–2026, and Grade 5 pretest for SY 2026–2027. The workbook-defined below-target group comprised 537 of 1,215 Grade 3 pretest records (44.2%), 251 of 535 Grade 3 posttest beneficiaries (46.9%), 635 of 1,210 Grade 4 pretest records (52.5%), 444 of the 1,160 Grade 4 learners assessed at posttest (38.3%), and 281 of 1,199 Grade 5 pretest records (23.4%). Of the 1,210 Grade 4 target learners, 50 (4.1%) were absent during the posttest; among those assessed, 716 (61.7%) were classified as having met the target. The Grade 3 posttest covered beneficiaries rather than the entire target cohort, limiting direct comparison with its pretest results. Across 21 school-level records provisionally matched between the Grade 3 and Grade 5 assessment snapshots, the unweighted mean below-target proportion decreased from 43.6% to 25.9%. Eighteen school records showed lower proportions in the later snapshot, while three showed higher proportions. An exploratory Wilcoxon signed-rank test indicated a statistically significant difference between the paired school-level proportions ($W = 28$, $p = .0014$). However, because the workbooks contain aggregate rather than learner-linked data, they cannot establish individual learners' transitions from EOSY to the subsequent BOSY or demonstrate that Project SARA caused the observed differences. The study recommends documenting assessment comparability and developing a privacy-protecting learner-linked dataset to support longitudinal analysis.

Keywords: *Project SARA; reading remediation; assessment comparability; school-level analysis*

INTRODUCTION

Reading is essential to learning across the elementary curriculum. Learners who struggle with reading may also experience difficulties understanding instructions, acquiring new knowledge, and completing tasks in other subjects. Reading assessment, therefore, should not simply classify learners; it should guide instruction, identify learning needs, and monitor progress.

In the Department of Education, literacy assessment initiatives such as the Comprehensive Rapid Literacy Assessment (CRLA) and the Philippine Informal Reading Inventory (Phil-IRI), together with learning recovery efforts under the Academic Recovery and Accessible Learning (ARAL) framework, emphasize the importance of identifying learners who need additional support. At the district level, Project SARA (Strengthening and Advocating Reading for All) serves as a reading initiative in the Schools District of Sara, Iloilo.

A concern arising from the district's assessment records is the apparent difference between end-of-school-year (EOSY) and beginning-of-school-year (BOSY) reading classifications. Learners who appear to have progressed by the end of one school year may receive lower classifications at the beginning of the next. This raises an important question: does the difference indicate a genuine decline in reading proficiency, or might it reflect changes in assessment procedures, learner participation, classification criteria, or reporting?

These explanations require different responses. Learners who struggle to retain or independently apply reading skills may need continued practice, while differences in assessment tools or classification thresholds may produce results that are not directly comparable. Attendance, enrolment changes, school transfers, and data-processing errors may also affect reported outcomes. Thus, lower BOSY classifications should not automatically be interpreted as learning loss, just as improved aggregate results should not automatically be considered proof of program effectiveness.

This study examines five Project SARA workbooks covering Grade 3 in School Year (SY) 2024–2025, Grade 4 in SY 2025–2026, and Grade 5 in SY 2026–2027. The records show varying proportions of learners classified below target. Consequently, the available data can describe reported reading classifications and support exploratory school-level comparisons, but they cannot establish whether the same learners experienced a decline between EOSY and the following BOSY. This study therefore examines what the records can validly demonstrate and identifies improvements needed to strengthen assessment interpretation and reading-progress monitoring under Project SARA.

Research Problem

The central research problem was to determine what the available Project SARA records establish about reported reading-classification patterns and the perceived EOSY–BOSY gap.

Specifically, the study sought to answer the following questions:

1. What numbers and proportions of learners were classified below target in the five available assessment workbooks?
2. How do below-target proportions differ across the provisionally matched schools between the Grade 3 assessment snapshot in SY 2024–2025 and the Grade 5 snapshot in SY 2026–2027?
3. What patterns emerge from the Grade 3 and Grade 4 pretest and posttest results, considering assessment coverage, absences, and denominators?
4. Can the available aggregate records establish individual learner regression from EOSY to the subsequent BOSY?
5. What data-quality, assessment-comparability, and monitoring improvements are needed to support valid longitudinal analysis?

Assumptions of the Study

The study was guided by four assumptions:

1. Record availability: The five supplied workbooks represent the available aggregate Project SARA records for the assessment periods examined.
2. Classification meaning: The reported reading categories reflect the rules used during each assessment period, but equivalence of instruments, languages, scoring procedures, and cutoffs across years is not assumed.
3. Assessment participation: Target enrolment and actual participation are distinct. The analysis accounts for the 50 Grade 4 posttest absentees.
4. Limits of aggregate analysis: Statistical differences can describe aggregate patterns but cannot establish individual learner transitions or causal program effects. Learner-linked EOSY and BOSY records would be needed to examine individual progress reliably.

Objectives of the Study

The study aimed to:

1. Describe the numbers and proportions of learners classified below target across the five assessment workbooks.
2. Compare school-level below-target proportions between the Grade 3 and Grade 5 assessment snapshots, subject to verification of school matching and denominators.

3. Examine Grade 3 and Grade 4 pretest and posttest patterns while accounting for beneficiary coverage and absences.
4. Determine whether aggregate records can establish individual learner regression between EOSY and the subsequent BOSY.
5. Identify data-quality and assessment-comparability limitations and recommend improvements in record validation, learner tracking, and longitudinal analysis to support evidence-based decisions under Project SARA.

Theoretical Framework

This study drew on two complementary perspectives: assessment for learning and progress monitoring, and implementation fidelity in educational interventions.

Assessment for learning emphasizes using assessment results to identify learning needs, guide instruction, and monitor progress. Applied to Project SARA, it suggests that a lower BOSY classification should prompt further assessment rather than an immediate conclusion that a learner has lost previously acquired skills. Differences in assessment tools, scoring criteria, and learner participation must first be considered.

Implementation fidelity refers to the extent to which an intervention is delivered as intended. Reading outcomes may be influenced by the frequency and duration of instruction, learner attendance, instructional focus, and quality of implementation. Recent reviews of literacy interventions and implementation fidelity support the importance of examining both outcomes and the conditions under which instruction is delivered (Al Otaiba et al., 2023; Hall et al., 2023; van Dijk et al., 2023).

Conceptual Framework

The study followed a four-stage analytical framework:

1. Assessment records: The five workbooks provided reading classifications, target enrolment, participation figures, and school-level totals.
2. Data validation and analysis: Frequencies, percentages, denominator checks, and exploratory school-level comparisons were used to examine reported patterns.
3. Evidence-based interpretation: Findings were assessed in light of data quality, assessment comparability, and the limitations of aggregate records.
4. Educational management and improvement: Results informed data-validation procedures, instructional follow-up, and the development of learner-linked longitudinal monitoring.

METHODOLOGY

Research Design

The study employed a retrospective quantitative, descriptive-comparative design using aggregate reading-assessment records. It combined descriptive analysis of district totals with an exploratory comparison of school-level proportions. The design was appropriate because the research examined existing administrative records rather than assigning an intervention or collecting new learner-level test scores.

The study was descriptive and diagnostic, not experimental or causal. Although the workbooks covered successive grade levels and school years, they did not contain stable learner identifiers linking EOSY results to the following BOSY. Accordingly, the study could not estimate individual learner transitions or establish that Project SARA caused the observed aggregate patterns.

Research Locale or Study Setting

The study was conducted in a schools district in the Province of Iloilo, Philippines, within the context of Project SARA (Strengthening and AdVocating Reading for All). The records covered Grade 3 in School Year (SY) 2024–2025, Grade 4 in SY 2025–2026, and Grade 5 in SY 2026–2027. The district workbooks listed 21 school records for the exploratory school-level comparison. The setting was relevant because the records document reading classifications across multiple assessment periods and can inform district-level monitoring and instructional planning.

Data Sources

The data sources constituted the complete set of available records for the defined scope, rather than a sample drawn from individual learners. The workbook-reported target totals were 1,215 for Grade 3 in SY 2024–2025, 1,210 for Grade 4 in SY 2025–2026, and 1,199 for Grade 5 in SY 2026–2027. The Grade 3 posttest included 535 beneficiaries. For the Grade 4 posttest, 50 of the 1,210 target learners were absent, leaving 1,160 assessed learners.

Workbook	Assessment period	Reported coverage	Methodological note
1. Grade 3 pretest	SY 2024–2025	1,215 target records	Full target cohort in workbook
2. Grade 3 posttest	SY 2024–2025	535 beneficiaries	Not directly comparable with full-cohort pretest
3. Grade 4 pretest	SY 2025–2026	1,210 target records	Pretest snapshot
4. Grade 4 posttest	SY 2025–2026	1,160 assessed; 50 absent	Use assessed denominator for test-taker proportions

5. Grade 5 pretest	SY 2026–2027	1,199 target records	Later-grade snapshot; not linked learner panel

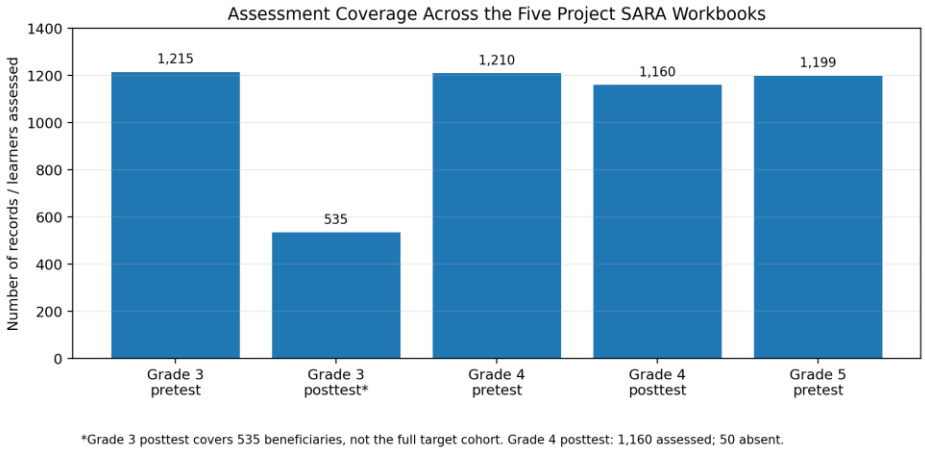


Figure 1. Assessment coverage represented in the five Project SARA workbooks.

Research Instruments or Materials

The primary materials were the five district Excel workbooks containing school-level enrolment or target counts and reading classifications. The workbooks included categories such as Pre-Primer, Primer, grade-level reading classifications, and reported passed/at-target or Failed/Non-Reader groupings.

Data Gathering Procedure

1. Inventory of Assessment Records. The five Excel workbooks covering Grade 3 in SY 2024–2025, Grade 4 in SY 2025–2026, and Grade 5 in SY 2026–2027 were examined. Relevant worksheets, school-level entries, and district summary rows were identified.
2. Data Extraction and Organization. Target enrolment, assessment participation, reading-category counts, school names, and district totals were extracted from each workbook. The extracted information was organized into a consolidated dataset for comparison and analysis.
3. Verification of Counts and Percentages. The percentages were recalculated using the appropriate numerators and denominators. This step helped distinguish percentages representing the full target enrolment from those representing only learners who participated in an assessment or benefited from the reading intervention.
4. Validation of Assessment Participation. The Grade 4 posttest records and 50 of the 1,210 absent target learners, leaving 1,160 learners who took the assessment were confirmed. The researcher accounted for this distinction when calculating and interpreting the posttest results.

5. Review of Category Totals. The reported reading-category counts were reviewed for consistency with the corresponding enrolment or assessment totals. Any differences identified were documented for appropriate interpretation rather than adjusted without supporting evidence.
6. School-Level Record Matching. School names and reported denominators across the Grade 3 and Grade 5 workbooks were compared to prepare the exploratory school-level analysis.

Data Analysis

Descriptive statistics were used to summarize the available records. Frequencies and percentages described the number and proportion of learners in the workbook-defined below-target group. For the Grade 4 posttest, the proportion among learners assessed used 1,160 as the denominator; the 1,210 target enrolment was retained separately to report attendance coverage.

Pretest and posttest patterns were summarized using counts, percentages, and percentage-point differences. Comparisons were interpreted cautiously when denominators differed or when assessment comparability was undocumented. The Grade 3 posttest percentage was treated as beneficiary-specific and was not interpreted as a whole-cohort pre/post change.

An exploratory school-level comparison was conducted between the Grade 3 SY 2024–2025 and Grade 5 SY 2026–2027 below-target proportions for 21 provisionally matched school records. The reported analysis used the Wilcoxon signed-rank test, with a paired-samples t-test and confidence interval as sensitivity summaries. No learner-level transition matrix or McNemar test was conducted because the supplied workbooks did not contain paired records for the same learners. The analysis therefore emphasizes descriptive patterns and the boundaries of valid inference rather than attributing changes to Project SARA.

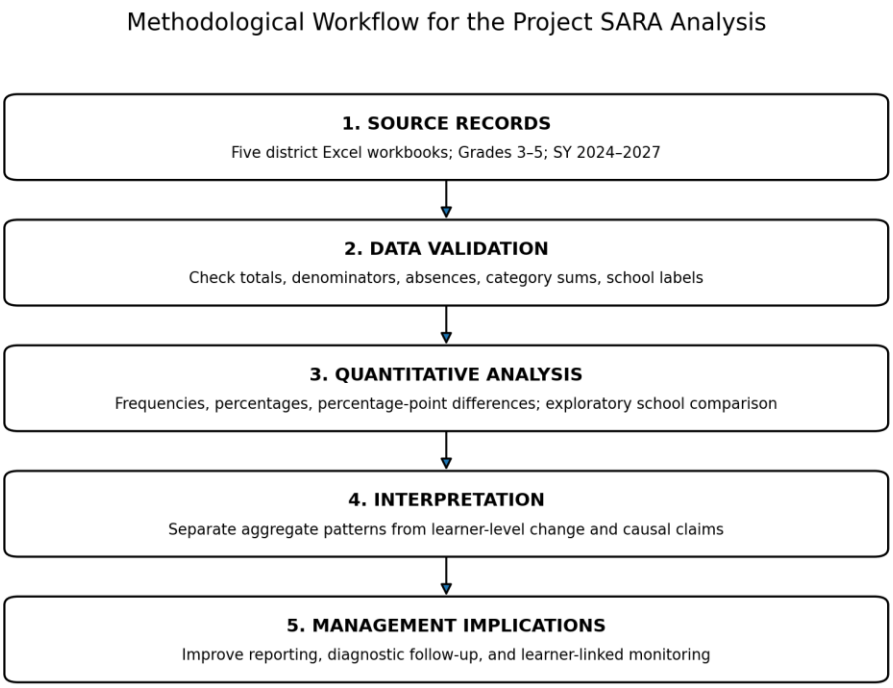


Figure 2. Data-processing and interpretation workflow used in the study.

Ethical Considerations

The study analyzed aggregate school-level records and did not report individual learner identities. Source files shall be stored securely, access shall be limited to authorized personnel, and the researcher's operational role in Project SARA shall be disclosed where required.

Any future learner-linked dataset shall use pseudonymous identifiers, restrict access, define data-retention procedures, and comply with applicable privacy and institutional requirements. Missing assessments, including absences, shall be documented explicitly rather than coded as reading failures.

RESULTS AND DISCUSSION

The findings are presented in the order of the study objectives. Percentages were recalculated from the reported counts and denominators in the five Project SARA workbooks. In this study, "below target" refers to the workbook-defined Failed/Non-Reader or struggling classification; it is not a standardized diagnosis. Because the workbooks contain aggregate counts rather than learner-linked records, the results describe distributions at specific assessment occasions and do not identify individual learner transitions.

Results for Objective 1: Describe the Reported Below-Target Counts and Proportions Across the Five Assessment Snapshots

Table 1 summarizes the district-level counts and proportions. Among the full-target pretest snapshots, the largest reported below-target proportion was in the Grade 4 pretest for SY 2025–2026 (635/1,210; 52.5%). The Grade 3 pretest recorded 537 of 1,215 learners below target (44.2%), while the Grade 5 pretest recorded 281 of 1,199 (23.4%). The Grade 3 posttest included 535 beneficiaries, of whom 251 (46.9%) were below target. At the Grade 4 posttest, 444 of the 1,160 learners assessed were below target (38.3%); 50 of the 1,210 target learners were absent.

Assessment snapshot	Below target	Denominator	Below-target proportion
Grade 3 pretest, SY 2024–2025	537	1,215 target learners	44.2%
Grade 3 posttest, SY 2024–2025	251	535 beneficiaries	46.9%
Grade 4 pretest, SY 2025–2026	635	1,210 target learners	52.5%
Grade 4 posttest, SY 2025–2026	444	1,160 assessed learners	38.3%
Grade 5 pretest, SY 2026–2027	281	1,199 target learners	23.4%

Table 1. District-level below-target counts and proportions by assessment snapshot.

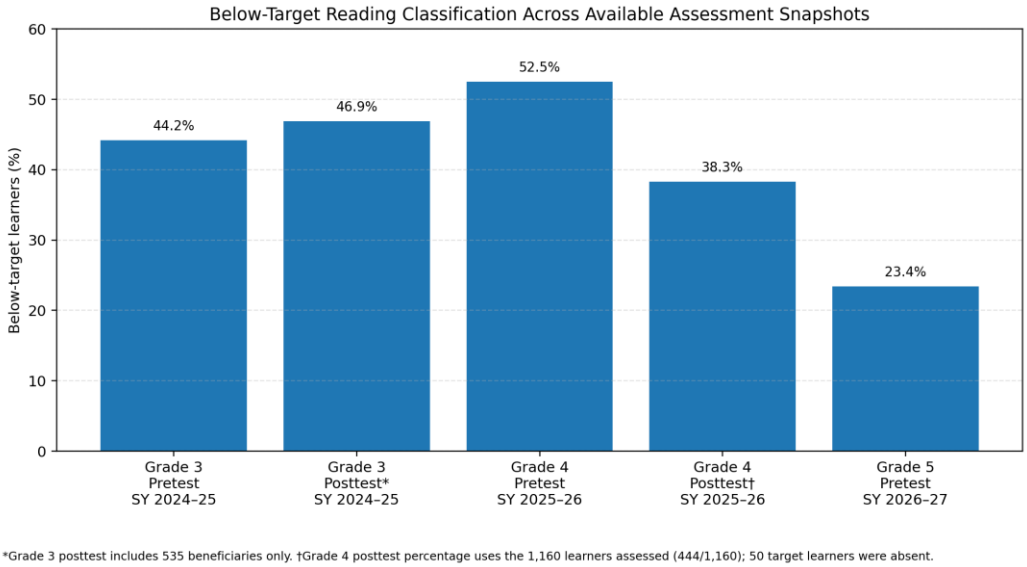


Figure 3. Below-target proportions across the five available assessment snapshots.

Discussion of Objective 1. The snapshots showed that the proportion classified below target varied across grades and assessment occasions. The Grade 5 pretest proportion was lower than the Grade 3 and Grade 4 pretest proportions, but this pattern alone does not demonstrate that the same learners improved over time. Grade level, cohort composition, assessment timing, classification cutoffs, school participation, and assessment comparability may differ. The Grade 3 posttest also represents beneficiaries rather than the full Grade 3 target group, so its 46.9% proportion should not be interpreted as a whole-cohort pre/post outcome. The defensible conclusion is that the workbooks document different distributions, not a verified longitudinal trajectory.

Results for Objective 2: Compare School-Level Below-Target Proportions Between the Grade 3 and Grade 5 Snapshots

In the exploratory school-level comparison, 21 school records were provisionally paired across the Grade 3 SY 2024–2025 and Grade 5 SY 2026–2027 snapshots. The unweighted mean below-target proportion was 43.6% in the Grade 3 snapshot and 25.9% in the Grade 5 snapshot, a mean decrease of 17.7 percentage points. Eighteen school records had a lower proportion in the Grade 5 snapshot, while three had a higher proportion. An exploratory Wilcoxon signed-rank test yielded $W = 28$, $p = .0014$.

The unweighted school mean gives each school equal influence regardless of enrolment size. It therefore answers a different question from the district-wide proportion, which weights learners through the overall count and denominator. The comparison is also across

Grade 3 and Grade 5 snapshots, not a matched learner panel. The observed school-level difference cannot be translated into an estimate of individual reading gains or a causal effect of Project SARA.

Discussion of Objective 2. The direction of the school-level pattern is encouraging as a monitoring signal: most provisionally matched schools had lower below-target proportions in the later snapshot. However, the result requires cautious interpretation because school names and denominators vary across files, the pairing has not been validated against an official school crosswalk, and the equivalence of assessment instruments and cutoffs is undocumented. A low p-value does not resolve these design limitations.

Results for Objective 3: Interpret the Grade 3 Beneficiary Posttest and Grade 4 Pretest–Posttest Totals

The Grade 3 posttest workbook recorded 251 below-target and 284 passed/at-target beneficiaries, totaling 535. The corresponding pretest workbook covered 1,215 target learners. Because the posttest denominator is limited to beneficiaries and the workbooks do not provide learner-level matching or explain selection into the beneficiary group, the two Grade 3 percentages do not represent a directly comparable full-cohort pre/post change.

For Grade 4, the pretest reported 635 below-target learners (52.5% of 1,210) and 575 passed/at-target learners. The posttest reported 444 below target and 716 passed/at target, totaling 1,160 assessed learners. Fifty of the 1,210 target learners were absent. Among those assessed, 38.3% were below target (444/1,160) and 61.7% were passed/at target (716/1,160). When the below-target count is divided by the full target enrolment, the proportion is 36.7% (444/1,210).

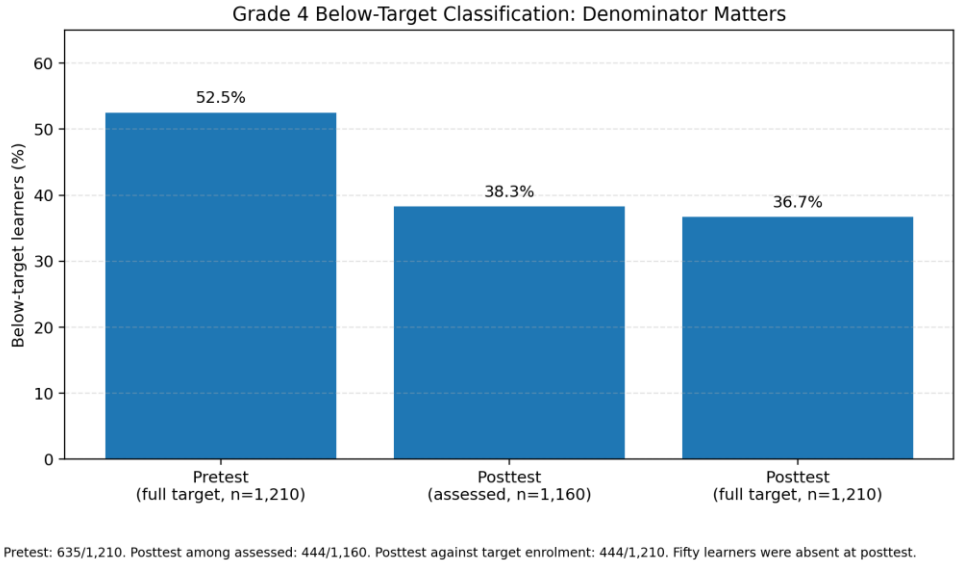


Figure 4. Grade 4 below-target proportions using the assessed and target-enrolment denominators.

Discussion of Objective 3. The Grade 4 posttest showed a lower below-target proportion than the pretest in the district summary, but aggregate totals cannot reveal whether individual learners moved upward, remained stable, or moved downward. The difference could reflect learner change, assessment conditions, participation, or other factors; the supplied workbooks do not isolate these explanations.

Results for Objective 4: Determine Whether the Aggregate Files Establish Individual EOSY-to-BOSY Regression

The files did not establish whether the same learners received a higher end-of-school-year classification and then a lower beginning-of-school-year classification. The workbooks provided school-level category totals without stable learner identifiers, individual scores, or linked EOSY and BOSY records.

Discussion of Objective 4. An aggregate increase in the below-target percentage does not prove that specific learners lost skills; likewise, a lower aggregate percentage does not prove that every learner improved. Changes in group composition, attendance, transfers, assessment language, passages, scoring rules, cutoffs, and timing may influence the distribution.

Results for Objective 5: Identify Data-Quality and Assessment-Comparability Limitations and Recommend Improvements for Longitudinal Monitoring

The review identified four immediate needs: (a) a verified school-name crosswalk; (b) explicit documentation of assessment versions, language, scoring procedures, and category cutoffs; (c) clear separation of target enrolment, assessed learners, absentees, and intervention beneficiaries; and (d) learner-linked records that permit valid matching across EOSY and BOSY.

Discussion of Objective 5. A privacy-protecting longitudinal dataset would allow the district to move from comparing snapshots to examining actual learner transitions. Each record should use a pseudonymous identifier and include grade, school, assessment date, instrument/version, language, raw or subskill scores where available, classification, attendance, transfer status, and intervention exposure. With validated repeated records, the district could construct transition matrices and estimate downward, stable, and upward classification changes.

General Discussion

Taken together, the five workbooks revealed substantial variation in the share of learners classified below target across the available Project SARA assessment snapshots. The Grade 4 district results showed a lower below-target share at posttest than at pretest, and the Grade 5 pretest recorded a lower below-target proportion than the earlier Grade 3 and Grade 4 pretests. Most provisionally paired school records also showed a lower below-target proportion in the Grade 5 snapshot. These patterns can inform instructional planning and identify where additional validation or support may be warranted.

The evidence did not, however, establish the specific mechanism behind the observed differences. The Grade 4 posttest excluded 50 absent learners from its assessed total; school-level matching was provisional; and assessment equivalence across years and grades had not been documented. Most importantly, aggregate counts cannot reveal individual EOSY-to-BOSY transitions. The findings should not be described as proof of reading loss, proof of individual growth, or proof that Project SARA caused the observed changes.

Research on structured literacy interventions supports systematic, targeted instruction while emphasizing that outcomes depend on the intervention, the skills measured, and the quality of implementation (Al Otaiba et al., 2023; Hall et al., 2023). Research on implementation fidelity also highlights the importance of documenting dosage, adherence,

and instructional quality (Van Dijk et al., 2023). These studies provide a rationale for strengthening monitoring, but they do not establish the effectiveness of Project SARA.

REFERENCES

Al Otaiba, S., McMaster, K., Wanzek, J., & Zaru, M. W. (2023). What we know and need to know about literacy interventions for elementary students with reading difficulties and disabilities, including dyslexia. Reading Research Quarterly, 58(2), 313–332. https://doi.org/10.1002/rrq.458

Department of Education. (2024, October 18). President Marcos signs ARAL Law to accelerate learning recovery. https://www.deped.gov.ph/2024/10/18/president-marcos-signs-aral-law-to-accelerate-learning-recovery/

Department of Education. (2025, March 4). Administration of the Comprehensive Literacy Assessment (CRLA) for Key Stage 1 learners and the Philippine Informal Reading Inventory (Phil-IRI) for Key Stage 2 and 3 learners for the end-of-the-school-year 2024–2025 [Joint memorandum]. https://ozamiz.deped.gov.ph/wp-content/uploads/2025/03/DM-120-S-2025.pdf

Hall, C., Dahl-Leonard, K., Cho, E., Solari, E. J., Capin, P., Conner, C. L., Henry, A. R., Cook, L., Hayes, L., Vargas, I., Richmond, C. L., & Kehoe, K. F. (2023). Forty years of reading intervention research for elementary students with or at risk for dyslexia: A systematic review and meta-analysis. Reading Research Quarterly, 58(2), 285–312. https://doi.org/10.1002/rrq.477

Van Dijk, W., Lane, H. B., & Gage, N. A. (2023). How do intervention studies measure the relation between implementation fidelity and students' reading outcomes? A systematic review. The Elementary School Journal, 124(1), 54. https://doi.org/10.1086/725672