RESEARCH ARTICLE

JENER Journal of Empirical and Non-Empirical Research
Volume 2, Issue 10, Page 418-425 | Article No. 46
DOI: 10.4898/jenerv2i10.a46
ISSN: 7895-9216
Journal homepage: https://jenerjournal.com
Journal email: editor@jenerjournal.com

JENER Journal
JENER Journal of Empirical and Non-Empirical Research

Project ARAL in Context: Implementation Fidelity, Learner Engagement, and Reading Recovery in Aglipay, Quirino, Philippines

Kimberly P. Pandatuan[1] & Kimberly T. Tarubal[2]

[1,2]Northeastern College

ARTICLE INFORMATION

Article history:
Published: October 2026

Keywords:
Academic Recovery and Accessible Learning
ARAL Program
Learning Recovery
Reading Intervention
Implementation Fidelity
Tutoring
Aglipay
Quirino

ABSTRACT

The study evaluated the implementation of the Academic Recovery and Accessible Learning (ARAL) Program in Aglipay, Quirino, with emphasis on ARAL-Reading during School Year 2025-2026. A quantitative implementation-evaluation design combined two analytical components: a survey of 108 school heads, ARAL focal persons, and tutors, and 320 matched learner records with baseline and post-assessment reading scores. The implementer survey measured learner identification and assessment, tutor readiness, session fidelity, resource adequacy, monitoring and feedback, family and community support, and perceived implementation effectiveness. Internal consistency was examined through Cronbach's alpha. Means and standard deviations described implementation conditions; Pearson correlation and multiple linear regression assessed relationships between implementation dimensions and perceived effectiveness. Learner progress was analyzed through a paired-samples t test with a within-person effect size, while post-assessment reading performance was examined through linear regression controlling for baseline score, attendance rate, grade level, and sex, with standard errors clustered by school. Implementation dimensions were rated from high to very high, although family and community support and resource adequacy were comparatively lower than tutor readiness and session fidelity. Overall implementation quality was positively related to perceived effectiveness ($r = .635$, $p < .001$). Tutor readiness, session fidelity, and monitoring and feedback were significant independent predictors of perceived effectiveness. Learner reading scores increased from $M = 49.88$ ($SD = 8.89$) at baseline to $M = 55.80$ ($SD = 11.96$) at post-assessment, a mean increase of 5.92 points, 95% CI [5.04, 6.81], $t(319) = 13.16$, $p < .001$, $d_z = .74$. After adjustment for baseline score and learner characteristics, attendance remained a positive predictor of post-assessment performance. The findings supported the importance of targeted learner identification, prepared tutors, regular delivery, progress monitoring, and sustained attendance in learning-recovery implementation. Because the study was observational and did not include a comparison group, the results described improvement and implementation associations rather than definitive causal effects.

1. Introduction

Learning recovery became a central education-policy concern after prolonged disruptions exposed and intensified gaps in foundational skills. In the Philippines, national and international assessments documented persistent challenges in reading, mathematics, and science. In PISA 2022, only 24% of Filipino 15-year-olds reached at least Level 2 in reading, while 16% reached at least Level 2 in mathematics and 23% reached at least Level 2 in science. These results were substantially below OECD averages and showed that a large share of learners continued to experience difficulty with the competencies required for independent learning.[11]

Foundational reading has particular importance because it supports learning across subject areas. The World Bank's learning-poverty framework identifies the inability to read and understand a simple text by around age 10 as a critical signal of weak foundational learning. When reading difficulties persist, learners face cumulative disadvantages because later lessons increasingly depend on comprehension, independent study, and the ability to acquire information from text.[12]

The Philippine government responded to learning-recovery needs through Republic Act No. 12028, the Academic Recovery and Accessible Learning (ARAL) Program Act. The law established a free national intervention for Kindergarten to Grade 10 learners who were returning to school after a furlough, performing below minimum proficiency levels in reading, mathematics, and science, or failing classroom assessments. It required learner assessment, tutorial support, trained tutors or learning facilitators, appropriate learning resources, progress monitoring, and coordination among education stakeholders.[1]

DepEd Order No. 18, s. 2025 operationalized the law by establishing the implementing framework for ARAL across basic education. The subsequent implementing rules and supplemental issuances clarified governance responsibilities, school readiness, learner determination, intervention planning, tutorial delivery, progress checks, tutor engagement, and monitoring. The ARAL School Readiness and Responsiveness Audit was introduced to determine whether schools possessed the organizational conditions required for delivery, while ARAL-Reading guidelines specified phased implementation and targeted support for learners whose reading performance fell below expected levels.[2–5]

For School Year 2025-2026, ARAL-Reading was prioritized as a major learning-recovery component. The guidelines for Key Stages 2 and 3 covered Grades 4 to 10 learners identified through Phil-IRI results, classroom assessments, or individualized reading assessment evidence. The program expected schools to provide appropriate instructional materials, build the capacity of tutors and school leaders, involve parents and communities, conduct close monitoring, and report learner progress. The guidelines also emphasized instruction matched to current reading ability rather than relying only on grade placement.[5]

The national framework was reinforced by regional and local preparation. DepEd Region II conducted upskilling for ARAL tutors and school heads on the ARAL-Reading Program in November 2025, with objectives that included clearer supervision roles, standardized tools, progress monitoring, and familiarity with updated learning resources. In Quirino Province, the Schools Division Office also introduced Project LAYAG (Learning Assistance for Youth Advancement and Growth) as a local initiative supporting the implementation of the ARAL Program. The Philippine Information Agency reported the launch of Project LAYAG during the province's 2025 anniversary activities.[8,9]

Aglipay was a relevant setting for examining how the policy translated into school-level practice. DepEd's current National Inventory Dashboard identifies multiple public elementary and secondary schools in Aglipay under the Schools Division of Quirino, including schools in both Aglipay East and Aglipay West. This distributed school context required implementation not only through formal policy compliance but also through tutor preparation, scheduling, attendance management, resource availability, learner progress monitoring, and family cooperation.[10]

The evidence base for tutoring supported the policy logic of ARAL. A systematic review and meta-analysis of experimental tutoring studies found a pooled positive effect on student learning and reported stronger effects when tutoring was frequent, embedded in school time, and delivered by teachers or trained paraprofessionals. Other syntheses similarly found that tutoring, feedback, and progress monitoring were among the more effective academic intervention components for learners at risk of low achievement.[13,14]

Randomized evaluations of intensive tutoring also demonstrated that adolescents who had fallen behind could make meaningful gains when instruction was individualized and delivered regularly. Guryan and colleagues reported positive mathematics impacts from high-dosage tutoring among ninth- and tenth-grade students, while evidence syntheses on struggling readers concluded that targeted tutoring and changes in instructional practice could improve reading outcomes when interventions were sufficiently structured and sustained.[15,22]

However, the existence of an evidence-based intervention did not guarantee that it would be implemented with sufficient quality. Implementation research has repeatedly distinguished the design of an intervention from the conditions under which it is delivered. Durlak and DuPre's review showed that variation in implementation was associated with variation in program outcomes, while Carroll and colleagues conceptualized implementation fidelity as adherence to intended content, frequency, duration, coverage, and quality of delivery.[17,18]

In education specifically, Century, Rudnick, and Freeman argued that fidelity measurement was necessary for determining whether outcomes could reasonably be interpreted in relation to an intervention. Proctor and colleagues likewise distinguished implementation outcomes from service and client outcomes, underscoring the importance of assessing whether a program was acceptable, feasible, adopted, and delivered as intended rather than judging implementation solely through final outcomes.[19,20]

These considerations were particularly relevant to ARAL. The program depended on a chain of implementation decisions: identifying eligible learners accurately, assigning and preparing tutors, protecting time for intervention sessions, providing suitable materials, monitoring progress, engaging families, and sustaining attendance. Weakness in any link could reduce the opportunity for learners to benefit even when the formal policy was well designed. Conversely, strong implementation did not by itself establish causality for learning gains because learner maturation, classroom instruction, attendance, prior achievement, and other factors could also influence performance.

The study therefore examined ARAL as both an implementation process and a learner-recovery intervention. It assessed the perceptions of personnel directly involved in implementation and separately analyzed matched learner assessment records. This dual approach reduced reliance on a single form of evidence and permitted a more disciplined interpretation of the relationship between implementation quality, participation, and observed reading improvement.

1.1 Research Questions

- What was the level of ARAL implementation in terms of learner identification and assessment, tutor readiness, session fidelity, resource adequacy, monitoring and feedback, and family and community support?
- What was the perceived level of overall implementation effectiveness?
- Was overall implementation quality significantly related to perceived implementation effectiveness?
- Which implementation dimensions significantly predicted perceived implementation effectiveness?
- Was there a significant difference between learners' baseline and post-assessment reading scores?
- Did ARAL attendance significantly predict post-assessment reading performance after controlling for baseline score, grade level, and sex?

The corresponding null hypotheses stated that implementation quality was not significantly related to perceived effectiveness, that the six implementation dimensions did not significantly predict perceived effectiveness, that baseline and post-assessment reading scores did not differ significantly, and that attendance did not significantly predict post-assessment reading performance after adjustment for the specified covariates.

2. Conceptual and Analytical Framework

The implementation component of the study drew from fidelity and implementation-outcome frameworks. Fidelity was treated as the extent to which the intervention was delivered in a manner consistent with policy expectations, while contextual supports were

treated as conditions that enabled or constrained implementation. The study operationalized six dimensions: learner identification and assessment, tutor readiness, session fidelity, resource adequacy, monitoring and feedback, and family and community support.[18–20] Learner identification and assessment referred to the systematic use of assessment evidence to identify participants and match instruction to reading needs. Tutor readiness referred to training, clarity of roles, confidence with session guides, and access to instructional support. Session fidelity referred to regularity, protected time, adherence to prescribed intervention activities, and continuity of tutoring. Resource adequacy covered learning materials, space, scheduling support, and basic logistical requirements. Monitoring and feedback referred to progress checks, documentation, data use, supervision, and adjustment of instruction. Family and community support referred to communication with parents or guardians, support for attendance, and reinforcement of learning outside tutorial sessions.[2,4,5]

The learner-outcome component was informed by the tutoring literature, which emphasizes targeting, frequency, trained personnel, aligned materials, data use, and sustained participation. Research has shown that tutoring effects vary substantially according to implementation features, and that attendance or exposure is one of the practical mechanisms through which an intervention can influence learning opportunities.[13,16]

3. Methods

3.1 Research Design

The study employed a quantitative implementation-evaluation design with two complementary components. The first component was a cross-sectional survey of personnel involved in ARAL implementation. The second component was a one-group matched-record analysis of learner reading performance from baseline to post-assessment. The survey component was explanatory because it tested relationships among implementation dimensions and perceived implementation effectiveness. The learner component was longitudinal at the record level because the same learners had baseline and post-assessment scores. The design supported analysis of implementation patterns and observed change but did not support a definitive causal estimate because no untreated comparison group was included.

3.2 Study Setting and Participants

The study was situated in selected public schools in Aglipay, Quirino, representing both elementary and secondary settings involved in ARAL-Reading implementation. The implementer dataset contained 108 valid responses from personnel with direct responsibilities in program delivery or oversight. The learner dataset contained 320 matched records from Grades 4 to 10 with complete baseline and post-assessment reading scores and documented session attendance.

Table 1. Profile of ARAL Implementer Respondents (n = 108)

Characteristic	Category	n	%
Primary ARAL role	Tutor/teacher	78	72.2
	School head	16	14.8
	ARAL focal person/coordinator	14	13.0
District cluster	Aglipay East	53	49.1
	Aglipay West	55	50.9
Years in teaching/education service	1–5 years	25	23.1
	6–10 years	31	28.7
	11–15 years	24	22.2
	More than 15 years	28	25.9

The implementer sample emphasized personnel who had direct exposure to learner identification, scheduling, tutoring, progress monitoring, and school-level supervision. The distribution across the two district clusters was nearly balanced, reducing the likelihood that one side of the municipality dominated the implementation ratings. The learner records covered Key Stages 2 and 3, where the ARAL-Reading guidelines required targeted intervention for learners who were below expected reading proficiency.

Table 2. Profile of Learner Records (n = 320)

Characteristic	Category	n	%
Sex	Female	162	50.6
	Male	158	49.4
Grade level	Grade 4	52	16.3
	Grade 5	51	15.9
	Grade 6	49	15.3
	Grade 7	49	15.3
	Grade 8	45	14.1
	Grade 9	41	12.8
	Grade 10	33	10.3
ARAL attendance	80% or higher	190	59.4
	Below 80%	130	40.6

3.3 Research Instruments and Measures

The implementation questionnaire used a five-point response scale ranging from 1 (strongly disagree) to 5 (strongly agree). Each implementation dimension contained four items, while perceived implementation effectiveness contained five items. Items were aligned with the requirements of the ARAL policy framework and with implementation-fidelity literature. The questionnaire did not reproduce proprietary instruments. Instead, statements were written specifically for the school context and focused on observable implementation conditions, such as whether learner selection was supported by assessment evidence, whether tutors understood the

session guides, whether scheduled sessions were delivered regularly, whether progress data were reviewed, and whether parents were informed about attendance and learner needs.

Dimension scores were computed as item means. Higher scores represented stronger implementation conditions. For descriptive interpretation, mean scores from 4.21 to 5.00 were described as very high, 3.41 to 4.20 as high, 2.61 to 3.40 as moderate, 1.81 to 2.60 as low, and 1.00 to 1.80 as very low. These bands were used only as descriptive labels and did not represent official DepEd performance standards.

For learner outcomes, the study used matched baseline and post-assessment reading scores expressed on a 0–100 criterion-referenced scale. The analyses treated the scores as continuous measures of reading performance. Attendance was expressed as the percentage of scheduled ARAL sessions attended. The post-assessment analysis controlled for baseline performance because learners entered the intervention at different starting levels.

3.4 Data Collection and Data Management

Implementation questionnaires were consolidated after the main ARAL-Reading implementation period. Learner records were matched using de-identified study codes so that baseline scores, post-assessment scores, grade level, sex, attendance, and school cluster could be analyzed without using learner names in the analytical file. Records lacking either baseline or post-assessment scores were excluded from the matched analysis. Survey records with incomplete scale responses were also excluded before scale computation.

Data screening included range checks, duplicate checks, examination of missing values, and review of score distributions. Continuous variables were inspected for implausible values and extreme outliers. Regression analyses included diagnostic checks for linearity, residual distribution, heteroscedasticity, and multicollinearity. Because learner observations were grouped within schools, the learner-level regression used school-clustered standard errors to reduce underestimation of uncertainty arising from within-school dependence.

3.5 Statistical Treatment

Cronbach's alpha assessed the internal consistency of the seven survey scales. Frequencies and percentages summarized participant characteristics. Means and standard deviations described implementation dimensions and perceived effectiveness. Pearson product–moment correlation tested the bivariate association between the overall implementation composite and perceived implementation effectiveness. Multiple linear regression estimated the unique association of the six implementation dimensions with perceived effectiveness.

For learner outcomes, a paired-samples t test compared baseline and post-assessment reading scores. The mean difference was reported with a 95% confidence interval, and Cohen's d_z was computed by dividing the mean within-person change by the standard deviation of the change scores. Effect-size interpretation was cautious because conventional labels can overstate or understate practical importance in education, where context, outcome metric, cost, and scalability matter.[21]

A multiple linear regression then estimated post-assessment reading score as a function of baseline score, attendance percentage, grade level, and sex. Standard errors were clustered by school. The attendance coefficient represented the expected difference in post-assessment score associated with a one-percentage-point increase in attendance after adjustment for the other variables. Statistical significance was evaluated at alpha = .05, and exact p values were reported where practical.

4. Results

4.1 Reliability of the Survey Measures

Table 3. Internal Consistency of Implementation Measures

Scale	Items	Cronbach's α	Interpretation
Learner identification and assessment	4	.921	Excellent
Tutor readiness	4	.928	Excellent
Session fidelity	4	.920	Excellent
Resource adequacy	4	.900	Good
Monitoring and feedback	4	.935	Excellent
Family and community support	4	.917	Excellent
Perceived implementation effectiveness	5	.889	Good

All scales demonstrated strong internal consistency. Cronbach's alpha ranged from .889 to .935. The coefficients supported computation of mean scale scores because items within each dimension showed substantial covariance while still representing distinct aspects of implementation. Reliability alone did not establish validity, but the results indicated that the items within each scale functioned consistently in the analytical sample.

4.2 Level of ARAL Implementation

Table 4. Descriptive Statistics for ARAL Implementation Dimensions

Dimension	Mean	SD	Interpretation	Rank
Tutor readiness	4.00	.29	High	1
Session fidelity	3.94	.30	High	2
Monitoring and feedback	3.94	.29	High	3
Learner identification and assessment	3.93	.31	High	4
Resource adequacy	3.81	.25	High	5
Family and community support	3.74	.29	High	6
Overall implementation composite	3.89	.22	High	—

The overall implementation composite was high (M = 3.89, SD = .22). Tutor readiness received the highest mean (M = 4.00, SD = .29), indicating that respondents generally viewed tutors as adequately oriented to their roles and prepared to conduct ARAL

sessions. Session fidelity and monitoring and feedback followed closely, suggesting that scheduled intervention activities and progress-review processes were generally perceived as functioning consistently.

Learner identification and assessment also received a high rating (M = 3.93, SD = .31). This pattern was aligned with the policy requirement that ARAL participation be based on evidence of learning need rather than broad, undifferentiated enrollment. Accurate identification was important because the intervention was intended to provide targeted support based on learner performance and reading level.[1,5]

Resource adequacy (M = 3.81, SD = .25) and family and community support (M = 3.74, SD = .29) were the two lowest-rated dimensions, although both remained in the high range. The relative difference suggested that schools perceived implementation as stronger in the instructional and supervisory processes under their direct control than in external support conditions. This distinction was important because tutoring research has emphasized that effective programs require not only trained tutors but also reliable scheduling, materials, attendance systems, and support for sustained participation.[13,16]

4.3 Perceived Implementation Effectiveness

Table 5. Perceived Implementation Effectiveness

Measure	Mean	SD	Interpretation
Overall perceived implementation effectiveness	3.89	.18	High

Respondents rated overall implementation effectiveness as high (M = 3.89, SD = .18). The relatively small standard deviation indicated that responses were concentrated near the favorable end of the scale. The measure reflected implementer perceptions of whether the program was reaching identified learners, supporting regular intervention, producing observable learner progress, and functioning as a feasible school-level recovery mechanism.

The perception score was not treated as a substitute for learner achievement data. Implementation research distinguishes process outcomes from client outcomes, and favorable views of a program can coexist with variation in actual learning gains. For this reason, the study analyzed learner assessment change separately rather than inferring learning recovery directly from implementer ratings.[20]

4.4 Relationship Between Overall Implementation and Perceived Effectiveness

Table 6. Correlation Between Overall Implementation Quality and Perceived Effectiveness

Variables	r	95% CI	p	Interpretation
Overall implementation × perceived effectiveness	.635	[.51, .73]	< .001	Strong positive association

Overall implementation quality was positively and significantly related to perceived implementation effectiveness, $r = .635$, $p < .001$. The direction indicated that respondents who reported stronger learner identification, tutor preparation, session delivery, resource support, monitoring, and family engagement also tended to view ARAL implementation as more effective.

The magnitude of the correlation indicated substantial shared variation but not equivalence between the two constructs. Implementation quality and perceived effectiveness were related, yet they remained conceptually distinct. This was consistent with implementation theory, which treats fidelity and enabling conditions as proximal determinants of whether an intervention is delivered successfully, while recognizing that program outcomes are shaped by additional learner and organizational factors.[17–20]

4.5 Predictors of Perceived Implementation Effectiveness

Table 7. Multiple Regression Predicting Perceived Implementation Effectiveness

Predictor	Standardized β	t	p	VIF
Learner identification and assessment	.137	1.56	.122	1.43
Tutor readiness	.281	3.15	.002	1.47
Session fidelity	.227	2.75	.007	1.27
Resource adequacy	-.001	-.02	.988	1.46
Monitoring and feedback	.265	2.92	.004	1.53
Family and community support	.007	.09	.930	1.34
Model statistics	$R^2 = .456$; Adj. $R^2 = .424$	$F(6,101) = 14.10$	< .001	—

The six implementation dimensions jointly explained 45.6% of the variance in perceived implementation effectiveness, $F(6, 101) = 14.10$, $p < .001$. Tutor readiness was the strongest unique predictor ($\beta = .281$, $p = .002$), followed by monitoring and feedback ($\beta = .265$, $p = .004$) and session fidelity ($\beta = .227$, $p = .007$). These three dimensions remained significant after the shared variance among all six predictors was controlled.

The result suggested that respondents differentiated between conditions that were generally present and those most directly connected with perceived effectiveness. Resource adequacy and family support were positively rated descriptively but did not show unique effects after tutor readiness, session fidelity, monitoring, and other dimensions entered the regression simultaneously. Their nonsignificance did not indicate that they were unnecessary; rather, their unique statistical contribution was small in the presence of correlated implementation conditions.

Variance inflation factors ranged from 1.27 to 1.53, well below levels commonly associated with serious multicollinearity. The regression therefore did not appear to be distorted by excessive overlap among predictors. The pattern was consistent with broader implementation research emphasizing training and technical support, adherence to intended delivery, and systematic monitoring as central implementation mechanisms.[17–19]

4.6 Change in Learner Reading Performance

Table 8. Baseline and Post-Assessment Reading Scores (n = 320)

Assessment	Mean	SD	Mean change	95% CI of change	t(319)	p	d_z
Baseline	49.88	8.89					
Post-assessment	55.80	11.96	5.92	[5.04, 6.81]	13.16	< .001	.74

Learner reading performance increased from a baseline mean of 49.88 (SD = 8.89) to a post-assessment mean of 55.80 (SD = 11.96). The mean within-person increase was 5.92 points, and the 95% confidence interval ranged from 5.04 to 6.81 points. The paired-samples test showed that the difference was statistically significant, $t(319) = 13.16$, $p < .001$.

The standardized within-person effect was $d_z = .74$. This indicated that the mean increase was sizable relative to the variability of individual change scores. The effect should not be interpreted as a causal treatment effect because the study had no untreated comparison group. Learners could improve for reasons that included ordinary classroom instruction, repeated exposure to reading tasks, maturation, or concurrent school interventions. The result nevertheless documented that the matched learner group had higher reading scores at the end of the observation period than at baseline.

The direction of improvement was consistent with the evidence base supporting intensive and targeted tutoring. Nickow, Oreopoulos, and Quan found positive pooled effects of tutoring across experimental studies, while Dietrichson and colleagues reported favorable effects for tutoring and progress monitoring among disadvantaged learners. Reading-specific syntheses have likewise found benefits from structured tutoring and instructional-process interventions for struggling readers.[13,14,22]

4.7 Attendance and Post-Assessment Reading Performance

Table 9. Baseline-Adjusted Regression Predicting Post-Assessment Reading Score

Predictor	B	Cluster-robust SE	z	p	95% CI
Baseline reading score	.988	.057	17.30	< .001	[.876, 1.100]
ARAL attendance (%)	.075	.036	2.08	.038	[.004, .146]
Grade level	-.182	.254	-.72	.473	[-.680, .315]
Sex (female = 1)	-.704	.826	-.85	.394	[-2.323, .915]
Model statistics	$R^2 = .553$; Adj. $R^2 = .548$		Wald F = 93.05	< .001	School-clustered SEs

Baseline reading score was the dominant predictor of post-assessment performance. Holding attendance, grade, and sex constant, each additional baseline point was associated with approximately .99 post-assessment points. This result reflected the expected persistence of prior achievement: learners who entered ARAL with stronger reading performance generally remained stronger at post-assessment.

Attendance also showed a positive adjusted association with post-assessment performance ($B = .075$, $p = .038$). A ten-percentage-point increase in session attendance corresponded to an estimated .75-point increase in post-assessment score after baseline performance, grade, and sex were controlled. Although the coefficient was modest, it was educationally relevant because attendance represented actual exposure to intervention opportunities rather than a background characteristic.

Grade level and sex were not significant predictors after baseline performance and attendance were included. The use of school-clustered standard errors accounted for dependence among learners within the same school. The attendance result was consistent with the design principles of high-impact tutoring, which emphasize regular participation and sufficient dosage as key conditions for achieving learning benefits.[13,16]

5. Discussion

The study examined ARAL implementation in Aglipay through the combined lenses of implementation quality and learner reading recovery. Four findings were central. First, implementation conditions were rated highly overall, with tutor readiness, session fidelity, monitoring, and learner identification receiving the strongest ratings. Second, overall implementation quality was strongly related to perceived effectiveness. Third, tutor readiness, session fidelity, and monitoring independently predicted perceived effectiveness. Fourth, matched learner records showed higher post-assessment reading scores, and attendance remained positively associated with post-assessment performance after baseline adjustment.

The strong rating for tutor readiness was consistent with the policy architecture of ARAL. RA 12028 and subsequent DepEd issuances did not treat tutoring as informal volunteer assistance; they established specific roles for tutors, learning facilitators, school heads, and supervisors and required mechanisms for preparation and support. Region II's upskilling activities similarly emphasized supervision, session guides, progress monitoring, and updated materials. The regression results suggested that implementers viewed this preparation as central to whether the program functioned effectively in schools.[1,6,8]

Session fidelity also emerged as an important predictor. The result reinforced a basic principle of implementation science: an intervention cannot be evaluated meaningfully if it is delivered sporadically or substantially differently from its intended design. In the ARAL context, fidelity included protecting tutorial time, delivering sessions regularly, using appropriate materials, and maintaining continuity across the implementation period. This finding aligned with Century and colleagues' argument that fidelity data are necessary to interpret educational interventions and with Carroll and colleagues' emphasis on adherence, dose, and quality of delivery.[18,19]

Monitoring and feedback showed an independent association with perceived effectiveness and occupied a central position in the ARAL policy framework. The ARAL-Reading guidelines embedded progress checks and midline or post-assessment activities in the intervention sequence. Monitoring was not merely a compliance requirement; it enabled tutors and school leaders to determine whether learners were responding to instruction and whether intervention plans required adjustment. The tutoring literature similarly identifies data use and ongoing assessment as features of effective tutoring systems.[5,16]

The lower relative ratings for resource adequacy and family and community support identified potential implementation pressure points. Even when instructional processes were functioning well, schools could encounter constraints involving materials, suitable spaces, scheduling, competing teacher responsibilities, transportation, household obligations, or inconsistent parent communication.

These conditions can reduce attendance or interrupt continuity. Implementation reviews have repeatedly shown that organizational capacity and external context can affect whether programs are delivered with sufficient quality.[17]

The learner analysis provided a separate form of evidence. The average reading score increased by approximately six points, and the within-person standardized change was substantial. This finding was compatible with the broader tutoring evidence but should be interpreted with methodological restraint. Meta-analyses and randomized trials provide causal estimates because they compare intervention participants with appropriate counterfactual groups. The present design did not. Consequently, the observed improvement showed temporal change among ARAL participants, not the amount of change uniquely attributable to ARAL.[13–15]

This distinction mattered because education outcomes are influenced by several concurrent processes. Learners continued to receive regular classroom instruction while participating in ARAL. Teachers could have changed classroom practices, learners could have become more familiar with assessment formats, and ordinary developmental progress could have contributed to higher scores. The absence of a comparison group meant that these influences could not be separated statistically from program participation.

Even with that limitation, attendance provided an informative implementation indicator. The positive association between attendance and post-assessment performance remained after baseline reading score was controlled. This finding did not prove that each additional session caused a fixed amount of learning, but it supported the practical importance of ensuring that identified learners actually received the intervention. A program can have strong materials and trained tutors yet produce limited benefits if learners attend irregularly. High-impact tutoring frameworks consistently emphasize frequency and sustained participation for this reason.[13,16]

The results also supported a systems interpretation of ARAL implementation. The program's effectiveness depended on more than tutor skill. Learner identification determined who received intervention; tutor readiness shaped instructional capacity; session fidelity determined exposure; monitoring supported adjustment; resources affected feasibility; and family support influenced participation. These dimensions interacted rather than operating as isolated components. A school that performed strongly on one dimension could still encounter implementation difficulty if another component was weak.

For Aglipay, the findings indicated that implementation-strengthening efforts should prioritize the dimensions that were both statistically important and operationally actionable. Continued tutor development, protected ARAL schedules, routine review of learner progress, and attendance follow-up represented immediate management levers. Resource planning and parent engagement required parallel attention because they could affect continuity even when they did not emerge as unique regression predictors in the implementer model.

6. Implications for Policy and Practice

First, tutor development should extend beyond one-time orientation. School heads and instructional supervisors should use coaching, demonstration sessions, peer observation, and learning-action-cell activities to strengthen diagnostic reading instruction and the use of ARAL materials. Periodic support is particularly important when tutors handle learners with varied reading levels and when instructional adjustments are required based on progress data.

Second, session fidelity should be treated as a measurable school-management indicator. Schools should document planned versus completed sessions, reasons for cancellations, make-up arrangements, and learner attendance. These records would make implementation review more concrete and would permit district supervisors to distinguish between instructional challenges and simple loss of intervention time.

Third, progress monitoring should be linked directly to instructional decisions. Assessment results are most useful when tutors can interpret them and modify lesson focus, grouping, pacing, or support. Schools should avoid treating assessment as an end-of-program reporting requirement only. The ARAL framework explicitly places assessment before, during, and after intervention, creating an opportunity for responsive instruction rather than static remediation.[5]

Fourth, attendance should be monitored as a core participation outcome. Learners with repeated absences could be flagged early for follow-up with parents or guardians. Attendance analysis should consider transportation, health, household responsibilities, timetable conflicts, and learner motivation rather than treating nonattendance solely as a compliance problem. Where feasible, schools could adjust schedules or provide catch-up sessions to protect intervention exposure.

Fifth, resource adequacy should be reviewed at the school and district levels using the readiness indicators established by DepEd. Materials, spaces, tutor assignments, schedules, and basic logistical supports should be evaluated before and during implementation. Resource constraints that recur across schools may require division-level procurement, material reproduction, or technical assistance rather than school-by-school solutions.[4]

Finally, future evaluation should link implementation records with verified learner outcomes in a design capable of stronger causal inference. Where ethically and operationally feasible, researchers could use phased implementation, matched comparison groups, regression discontinuity around eligibility thresholds, or interrupted time-series designs. Such approaches would provide more defensible estimates of program impact while preserving the program's objective of serving learners with documented need.

7. Conclusion

The study showed how the ARAL Program could be evaluated as both an implementation process and a learning-recovery intervention. In Aglipay, implementation conditions were generally rated highly, and the strongest operational features involved tutor readiness, regular session delivery, monitoring, and learner identification. Overall implementation quality was strongly associated with perceived effectiveness, while tutor readiness, session fidelity, and monitoring remained significant predictors in the multivariable analysis.

Matched learner records also showed higher reading performance at post-assessment than at baseline. Attendance retained a positive association with post-assessment performance after baseline reading score, grade, and sex were controlled. Together, the findings emphasized that learning recovery depended on whether intervention opportunities were actually delivered and received: appropriate learners had to be identified, tutors had to be prepared, sessions had to occur regularly, progress had to be monitored, and learners had to attend.

References

[1] Republic of the Philippines. (2024). Republic Act No. 12028: Academic Recovery and Accessible Learning (ARAL) Program Act. LawPhil. https://lawphil.net/statutes/repacts/ra2024/ra_12028_2024.html

[2] Department of Education. (2025). DepEd Order No. 018, s. 2025: Implementing Guidelines of the Academic Recovery and Accessible Learning Program. https://www.deped.gov.ph/2025/06/13/june-13-2025-do-018-s-2025-implementing-guidelines-of-the-academic-recovery-and-accessible-learning-program/

[3] Department of Education. (2025). DepEd Memorandum No. 058, s. 2025: Dissemination of the Implementing Rules and Regulations of Republic Act No. 12028. https://www.deped.gov.ph/wp-content/uploads/DM_s2025_058r.pdf

[4] Department of Education. (2025). DepEd Memorandum No. 056, s. 2025: Supplemental Guidelines on the Implementation of the ARAL School Readiness and Responsiveness Audit. https://www.deped.gov.ph/wp-content/uploads/DM_s2025_056_r.pdf

[5] Department of Education. (2025). DepEd Memorandum No. 064, s. 2025: Implementing Guidelines for the Academic Recovery and Accessible Learning Program in Reading for Key Stages 1 to 3. https://www.deped.gov.ph/wp-content/uploads/DM_s2025_064r.pdf

[6] Department of Education. (2026). DepEd Memorandum No. 001, s. 2026: Guidelines on the Engagement of Academic Recovery and Accessible Learning Program Tutors and Support Aides. https://www.deped.gov.ph/wp-content/uploads/DM_s2026_001r.pdf

[7] Department of Education. (2026). DepEd Order No. 010, s. 2026: Guidelines for the Implementation of the 2026 Summer Remediation Programs. https://www.deped.gov.ph/wp-content/uploads/DO_s2026_010r.pdf

[8] Department of Education Region II. (2025). Regional Memorandum No. 509, s. 2025: Regional Upskilling of ARAL Tutors and School Heads on ARAL Reading Program. https://region2.deped.gov.ph/wp-content/uploads/2025/10/REGIONAL-MEMORANDUM-NO.-509-S.-2025-REGIONAL-UPSKILLING-OF-ARAL-TUTORS-AND-SCHOOL-HEADS-ON-ARAL-READING-PROGRAM.pdf

[9] Philippine Information Agency. (2025). Quirino at 54: Honoring heritage, celebrating first-class dreams. https://pia.gov.ph/news/luzon/cv/quirino-at-54-honoring-heritage-celebrating-first-class-dreams/

[10] Department of Education. (2026). National Inventory Dashboard: Quirino, Region II - Schools. Retrieved September 30, 2026, from https://www.nid.deped.gov.ph/public-dashboard/region/Region%20II/division/Quirino

[11] OECD. (2023). PISA 2022 Results (Volumes I and II) - Country Notes: Philippines. OECD Publishing. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/philippines_a0882a2d-en.html

[12] World Bank. (2022). Ending Learning Poverty. https://www.worldbank.org/en/brief/2022/10/12/ending-learning-poverty

[13] Nickow, A., Oreopoulos, P., & Quan, V. (2024). The promise of tutoring for PreK-12 learning: A systematic review and meta-analysis of the experimental evidence. American Educational Research Journal, 61(1), 74-107. https://doi.org/10.3102/00028312231208687

[14] Dietrichson, J., Bøg, M., Filges, T., & Jørgensen, A.-M. K. (2017). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. Review of Educational Research, 87(2), 243-282. https://doi.org/10.3102/0034654316687036

[15] Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M. V., Dodge, K., Farkas, G., Fryer, R. G., Mayer, S., Pollack, H., Steinberg, L., & Stoddard, G. (2023). Not too late: Improving academic outcomes among adolescents. American Economic Review, 113(3), 738-765. https://doi.org/10.1257/aer.20210434

[16] Robinson, C. D., Kraft, M. A., Loeb, S., & Schueler, B. E. (2024). Design principles for accelerating student learning with high-impact tutoring. EdResearch for Action. https://edresearchforaction.org/research-briefs/design-principles-for-accelerating-student-learning-with-high-impact-tutoring/

[17] Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. American Journal of Community Psychology, 41, 327-350. https://doi.org/10.1007/s10464-008-9165-0

[18] Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. Implementation Science, 2, 40. https://doi.org/10.1186/1748-5908-2-40

[19] Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. American Journal of Evaluation, 31(2), 199-218. https://doi.org/10.1177/1098214010366173

[20] Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Administration and Policy in Mental Health and Mental Health Services Research, 38, 65-76. https://doi.org/10.1007/s10488-010-0319-7

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

[22] Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. Educational Research Review, 6(1), 1-26. https://doi.org/10.1016/j.edurev.2010.07.002

[23] Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. Reading Research Quarterly, 43(3), 290-322. https://doi.org/10.1598/RRQ.43.3.4