

The Effectiveness of the Utilization of the Instructional Tools in Improving the Learners' Performance in Grade 1 In the Public Elementary Schools in the Division of Albay

Damian D. Colarina¹

¹School of Graduate Studies, University of Northeastern Philippines

ARTICLE INFORMATION	ABSTRACT
Article history: Published: October 2026 Keywords: Instructional Tools Grade 1 Reading Teacher Perceptions Public Elementary Schools Kendall Concordance Albay	The study described Grade 1 teachers' utilization of instructional tools, perceived effectiveness for learners' performance, implementation challenges, and proposed solutions in public elementary schools in the Division of Albay during school year 2025–2026. A descriptive–evaluative–correlational design surveyed 160 teachers selected purposively across three geographic areas. A structured questionnaire supplied five-point ratings; weighted means, ranks, and Kendall's coefficient of concordance summarized and compared area-level rank orders. Utilization averaged 4.45 (Much Utilized); perceived effectiveness averaged 4.45 (Much Effective); and challenges averaged 4.36 (Much Challenging). All six utilization domains and all six challenge domains had nonsignificant rank concordance at .05. In perceived effectiveness, adequacy of instructional tools ($W = 0.73$, $\chi^2 = 24.00$, $df = 11$) and learners' engagement and reading improvement ($W = 0.66$, $\chi^2 = 17.82$, $df = 9$) exceeded the reported .05 critical values. The 23 proposed solutions showed significant concordance ($W = 0.40$, $\chi^2 = 79.00$, $df = 22$, $p < .001$). Teachers prioritized sufficient textbooks and supplementary reading resources (4.47) and daily time for tool use (4.46). These are perceptions and ranked appraisals, not measured learner gains or demonstrated effects of tools. Internal conflicts in the instrument's solution count and some significance statements require author review.

1. Introduction

The selection, preparation, and routine use of instructional tools are practical parts of early-grade teaching. Printed texts, teacher-made materials, audiovisual resources, and digital tools may support demonstration, practice, feedback, and participation when they fit the lesson and the learners. Their presence alone cannot show that learners have acquired a skill. The quality of the task, the teacher's use of the material, time available, and opportunities to respond all matter. Smaldino et al. (2019) frame media selection as part of an instructional design process, while Keller (2010) emphasizes attention and relevance in planning motivating learning experiences. These perspectives help explain why the study examined suitability, teacher proficiency, participation, and management rather than mere availability.

In Grade 1, the transition into formal literacy gives particular importance to accessible and suitable reading materials. A classroom may report frequent use of tools yet still face shortages of appropriately leveled or locally relevant texts, insufficient preparation time, or inconsistent support for teachers. A study of primary mathematics resources by Okoji and Olubayo (2022) likewise distinguishes availability from utilization, although its subject, setting, and measurement differ from this Albay inquiry. Such literature provides a context for examining practices and constraints; it cannot validate an inference that a high perception score corresponds to improved reading achievement in the present sample.

The phrase “effectiveness ... in improving learners' performance” is preserved in the study title. Operationally, the reported outcome consists of teachers' ratings of adequacy, relevance, competency, frequency, organization, and learners' engagement and reading improvement. No independently scored reading assessment, baseline, posttest, comparison group, or linked learner record is presented. The study therefore answers how teachers appraised instructional tools and their challenges, while a question about actual improvement remains unanswered by these data. This distinction is central to interpreting both high ratings and Kendall's agreement tests. The study asked four connected questions: How extensively did Grade 1 teachers use instructional tools across six domains? How effective did they perceive those tools to be for learner performance across six domains? What challenges did they encounter? Which proposed solutions did they judge necessary? For each set of item rankings, the analysis also examined whether the three areas showed significant agreement. The manuscript presents the area summaries and reported concordance results, then considers what they imply for instructional planning and a future direct evaluation of learning.

2. Methods

2.1 Design and setting

The study used a descriptive–evaluative–correlational approach during school year 2025–2026 in public elementary schools of the Division of Albay. Its descriptive component summarized teacher ratings; its evaluative component applied response-scale labels to judgments about instructional tools; its inferential component used Kendall's W to examine agreement among area-level rankings. Although the design is termed correlational, the presented inferential results are tests of rank concordance, not correlations between tool use and independently measured learner outcomes. No intervention or longitudinal comparison was reported.

2.2 Participants and sampling

The study reports purposive selection of 160 Grade 1 teachers in 17 public elementary schools. Schools were grouped into Areas A, B, and C, corresponding to the designated districts of Sto. Domingo, Bacacay East, Malilipot, Malinao, and Tiwi. Eligibility was described as an operating public elementary school offering Grade 1, with active Grade 1 enrollment and willingness to participate. Area A included six schools, Area B six, and Area C five in the detailed respondent table. The methods call the recruitment purposive-stratified and state 100% completion, but do not identify the invitation denominator or detailed nonresponse accounting.

Table 1. Respondent distribution by geographic area

Area	Schools in respondent table	Teachers	Reported percentage
A	6	56	35
B	6	55	34
C	5	49	31
Grand total	17	160	100

Note. Percentages and participant counts reproduce the respondent table; school counts are the enumerated rows in that table.

The three areas contributed similar numbers of teachers, with Area A accounting for 56, Area B 55, and Area C 49. Geographic grouping enabled an area-level comparison of item rankings, but purposive selection does not establish a probability sample of all Albay Grade 1 teachers. The table gives the composition of the analytic respondent pool rather than an estimate of how frequently tools were used in the wider division. The study does not give teachers' years of service, class size, school resource inventories, or learner characteristics, which restricts interpretation of differences across areas.

2.3 Questionnaire, validation, and procedure

A structured, researcher-developed questionnaire was reported as the primary data collection instrument. The table of specifications assigned 66 items to utilization, 66 to perceived effectiveness, 66 to challenges, and 22 to proposed solutions, for a stated total of 220. Six domains were specified within each of the first three parts. In the reported findings, utilization was organized as preparation; suitability with learners' needs; teachers' proficiency; supervision and management of materials; learners' participation and progress; and learners' engagement. The table of specifications uses different labels and positions for several utilization domains, so the final instrument and scoring key need to establish their precise mapping. The results present 23 numbered solution proposals, one more than the specified 22.

Each part used a five-point scale. Reported weighted means of 3.50–4.49 were labeled Much Utilized, Much Effective, Much Challenging, or Much Necessary as applicable; 4.50–5.00 received the corresponding Very Much labels. Lower bands were 2.50–3.49, 1.50–2.49, and 1.00–1.49 with part-specific descriptors. Content review by professionals in educational leadership and research and a pilot with respondents similar to the target group were described. Item wording was reportedly revised after feedback. The methods provide no number of validators or pilot participants and no reproducible reliability coefficient by domain. A section reporting "reliability ... 2.4563" and a computed value of 1.645 lacks a coherent identification of the coefficient and test statistic, so it cannot substantiate scale reliability without clarification.

The stated sequence included preparing and reviewing the questionnaire, a dry run, administration, retrieval, tabulation, and interpretation. The exact administration dates, survey mode, consent process, treatment of missing answers, and individual-level data structure were not provided. Optional interview and observation tools are mentioned prospectively, but no results from either are presented. Accordingly, only questionnaire-derived findings are analyzed here. Claims about classroom behavior or learning must be read as teacher reports rather than independent observations.

2.4 Analysis

The study used frequencies, percentages, weighted means, and rank orders, then Kendall's coefficient of concordance W for rankings across three areas. The associated chi-square statistic was compared with tabular critical values at $\alpha = .05$. The tables report item counts and degrees of freedom separately for each domain (12 items, $df = 11$; 11 items, $df = 10$; or 10 items, $df = 9$) and for solutions (23 items, $df = 22$). W describes agreement in the three areas' rank ordering of items; a nonsignificant result is not evidence that teachers within an area agree, and a significant result is not evidence of improvement in pupil performance. Exact p values were not generally reported; significance below is based on the printed statistics, critical values, and decision rows. No individual-level association or regression was reported.

3. Results

3.1 Utilization of instructional tools

Table 2. Summary of teachers' utilization ratings

Domain	Area A W_x	Area B W_x	Area C W_x	Overall W_x	Interpretation	Rank
Preparation of Instructional Tools	4.46	4.48	4.45	4.47	MU	1.5
Suitability with Learners' Needs	4.46	4.45	4.46	4.46	MU	3
Teachers' Proficiency	4.46	4.47	4.47	4.47	MU	1.5
Supervision and Management of Materials	4.44	4.44	4.42	4.43	MU	6
Learners' Participation and Progress	4.45	4.44	4.45	4.45	MU	4.5
Learners' Engagement	4.45	4.44	4.46	4.45	MU	4.5
Average	4.45	4.45	4.45	4.45	MU	

Note. W_x = weighted mean; MU = Much Utilized. Numbers, overall interpretations, and ranks follow the reported summary. The Area C average is printed as 4.45 “VMU” in that table despite the scale classifying 4.45 as MU; this manuscript uses the stated scale and the table’s overall MU label.

All six overall means fell in the Much Utilized range, with preparation and teachers’ proficiency tied at 4.47, and management of materials lowest at 4.43. The narrow spread of 0.04 suggests similarly favorable appraisals at the domain level. Yet it cannot resolve how often individual tools were used, whether they were available across all schools, or whether they matched particular reading tasks. The relatively lower management rating identifies an area for checking storage, maintenance, allocation, and classroom access. Smaldino et al. (2019) treat the integration of media as an instructional choice; a high self-report rating alone does not demonstrate that selection and use were aligned with learner needs.

Table 3. Kendall concordance for area rankings of utilization items

Domain	Items	W	χ^2	df	.05 critical χ^2	Decision
Preparation of Instructional Tools	12	0.54	17.85	11	19.68	NS
Suitability with Learners’ Needs	11	0.45	13.50	10	18.31	NS
Teachers’ Proficiency	10	0.53	14.31	9	16.92	NS
Supervision and Management of Materials	12	0.51	16.83	11	19.68	NS
Learners’ Participation and Progress	11	0.23	6.90	10	18.31	NS
Learners’ Engagement	10	0.54	14.58	9	16.92	NS

Note. W , χ^2 , df, critical values, and NS labels reproduce the reported Table 10. NS = not significant at .05; the study labels the null decision “Accepted.”

None of the six computed χ^2 values exceeded its .05 critical value. Thus the study did not detect statistically significant concordance among the three areas’ item rankings in any utilization domain. Learners’ participation and progress had the lowest reported W (0.23), whereas preparation and engagement each had $W = 0.54$. The coefficients are not correlations of utilization with reading gains. Different local priorities, tied or compressed ratings, and the small number of ranking groups could all affect the test; the available summaries do not isolate a cause. Domain means can be similar across areas even when finer item rankings lack statistically supported agreement.

3.2 Perceived effectiveness for learner performance

Table 4. Summary of teachers’ perceived effectiveness ratings

Domain	Area A W_x	Area B W_x	Area C W_x	Overall W_x	Interpretation	Rank
Adequacy of instructional tools	4.43	4.45	4.43	4.43	ME	6
Relevance of Materials to Learners’ Needs	4.44	4.44	4.43	4.44	ME	4.5
Teachers’ Competency in Using Tools	4.44	4.45	4.46	4.45	ME	2.5
Frequency and consistency of use	4.43	4.47	4.46	4.44	ME	4.5
Organization and maintenance of tools	4.45	4.47	4.46	4.46	ME	1
Learners’ engagement and reading improvement	4.45	4.42	4.46	4.45	ME	2.5
Average	4.44	4.45	4.45	4.45	ME	

Note. W_x = weighted mean; ME = Much Effective. Values, labels, and ranks reproduce the reported summary, including the printed 4.44 average for frequency and consistency of use.

Organization and maintenance ranked first (4.46), while adequacy ranked sixth (4.43); all domains were Much Effective by the study’s teacher-rating scale. The table supports a conclusion that respondents viewed the tools favorably, not that the tools caused academic improvement. The learners’ engagement and reading improvement domain blends a classroom process with an outcome in one perceived rating. Keller (2010) connects attention and relevance to motivational design, but motivation and engagement are not interchangeable with assessed reading proficiency. The lower adequacy score, despite its favorable label, also merits a resource audit before interpreting the high organizational rating as sufficient provision.

Table 5. Kendall concordance for area rankings of perceived-effectiveness items

Domain	Items	W	χ^2	df	.05 critical χ^2	Reported decision
Adequacy of Instructional Tools	12	0.73	24.00	11	19.68	Significant (.025)
Relevance of Materials to Learners’ Needs	11	0.29	8.70	10	18.31	NS
Teachers’ Competency in Using Tools	10	0.18	4.36	9	16.92	NS
Frequency and Consistency of Use	12	0.45	14.85	11	19.68	NS
Organization and Maintenance of Tools	11	0.40	12.00	10	18.31	NS
Learners’ Engagement and Reading Improvement	10	0.66	17.82	9	16.92	Significant (.05)

Note. The significance column follows the decision and significance rows of the reported Table 18. Its “.025” and “.05” are tabular thresholds, not exact p values. All other comparisons are labeled NS.

Rank agreement across areas exceeded the reported .05 critical value for adequacy and for learners’ engagement and reading improvement. The other four domains did not. The strongest concordance, adequacy ($W = 0.73$), suggests that the areas placed the adequacy items in relatively similar rank order despite adequacy being the lowest mean domain. The engagement and reading improvement concordance ($W = 0.66$) points to a similar prioritization of its constituent items, not a verified increase in reading. Teacher competency had the smallest W (0.18), so its favorable overall mean masks little demonstrable agreement across area rankings. The findings cannot identify which tools raised achievement, because no achievement outcome entered these tests.

3.3 Challenges and proposed solutions

Table 6. Summary of challenges encountered by teachers

Domain	Area A Wx	Area B Wx	Area C Wx	Overall Wx	Interpretation	Rank
Material Access	4.36	4.37	4.36	4.36	MC	3.5
Tools Quality	4.36	4.37	4.36	4.36	MC	3.5
Teacher readiness	4.36	4.37	4.39	4.38	MC	1
Time/Resources Constraints	4.36	4.36	4.37	4.36	MC	3.5
Support and Monitoring	4.37	4.36	4.36	4.36	MC	3.5
Learners' Engagement	4.38	4.35	4.35	4.36	MC	6
Average	4.37	4.36	4.37	4.36	MC	

Note. Wx = weighted mean; MC = Much Challenging. Values, interpretations, and ranks follow the reported Table 25. Equal rounded means have the printed ranks, including the rank 6 for learner engagement.

Teacher readiness had the highest challenge rating (4.38), although all six means were clustered within 0.02 and remained Much Challenging. High reported utilization and high reported challenges can coexist: teachers may use available tools regularly while still needing preparation, better material quality, time, and technical assistance. The means do not measure a shortfall in training hours or inventories, and a 0.02 difference is too small to justify an exclusive focus on teacher readiness. A local response would examine the six constraints together and collect concrete counts of usable resources and opportunities for supported practice.

Table 7. Selected proposed solutions and reported priority ratings

Proposal as reported	Area A Wx	Area B Wx	Area C Wx	Overall Wx	Rank
Provide a sufficient stock of textbooks and supplementary reading resources for learners.	4.48	4.45	4.47	4.47	1
Dedicate 30 minutes each day to using instructional tools	4.48	4.43	4.48	4.46	2
Provide access to a wide range of reading materials in the local language.	4.42	4.46	4.40	4.43	4
Make sure teachers have access to necessary teaching materials, resources, and technology	4.40	4.47	4.41	4.43	4
Allow teachers adequate planning time to incorporate new instructional tools and approaches.	4.47	4.35	4.46	4.43	4
Provide regular professional development for teachers.	4.32	4.36	4.30	4.33	23

Note. These are selected rows from the 23-row solution table, retaining wording, means, and printed ranks. All selected overall ratings were classified Much Necessary (MN); tied rank 4 is printed for the three proposals at 4.43.

Respondents placed textbooks and supplementary reading resources first (4.47), followed by a dedicated 30 minutes daily for tools (4.46). Three proposals at 4.43 concerned local-language reading material, broad resource access, and planning time. Even the lowest ranked proposal shown, regular professional development, remained Much Necessary at 4.33. The ordering offers a practical sequence for discussion, not evidence that the first option is more effective than training. Resource provision and protected instructional time require implementation and outcome measurement. It would be premature to mandate a uniform 30-minute period solely from preference ratings without examining timetable feasibility and learner-level effects.

Table 8. Kendall concordance for area rankings of challenge items and solutions

Category	Items	W	χ^2	df	.05 critical χ^2	Decision
Material Access	12	0.29	9.64	11	19.68	NS
Tools Quality	11	0.45	13.38	10	18.31	NS
Teacher Readiness	10	0.33	9.40	9	16.92	NS
Time/Resources Constraints	12	0.43	14.19	11	19.68	NS
Support and Monitoring	11	0.39	11.79	10	18.31	NS
Learners' Engagement	10	0.43	11.62	9	16.92	NS
Solutions Offered	23	0.40	79.00	22	33.92	Significant (.001)

Note. Statistics and decisions reproduce the reported Table 27. NS = not significant at .05. The solution statistic exceeds the printed .001 critical value of 48.27; the table's ".001" denotes its significance threshold, not an exact p value.

None of the six challenge domains had a χ^2 value exceeding its .05 critical value. The three areas therefore did not show statistically significant agreement in rankings within those domains, even though their domain means in Table 6 were nearly alike. The solution ranking was different: $\chi^2 = 79.00$ exceeded both the .05 and .001 critical values, so the reported analysis indicates concordance in prioritizing the 23 proposals. A W of 0.40 is partial agreement, not unanimity. Its statistical significance does not establish the feasibility, cost, or effectiveness of a proposed intervention; those require separate assessment.

3.4 Discussion

The first research question produced favorable reports of utilization across preparation, suitability, proficiency, management, participation, and engagement. The modest domain differences do not identify a single dominant practice; instead, they indicate a broadly positive response pattern. Self-report may reflect perceived expectations, common response tendencies, or genuinely frequent practice, and the study cannot separate these explanations. The lack of significant area-level item concordance in Table 3 also cautions against interpreting uniform domain means as uniform priorities. Schools could check the specific tools actually used, their condition, and their place in lesson plans before adopting a division-wide allocation rule.

The second question concerns "effectiveness," yet the questionnaire recorded teachers' appraisals rather than direct changes in learner performance. Table 4's high scores suggest perceived instructional value. Table 5 shows that adequacy and engagement/reading improvement items had significant rank concordance across the three areas, while teacher competency and

several other domains did not. Agreement about which adequacy items matter can guide a common audit, but it cannot estimate the learning effect of new materials. Francisco and Celon (2020) examined teachers' instructional practices alongside student records in a distinct Philippine setting and reported no significant combined effects in their analysis. Their design illustrates the difference between including performance data and assuming it from teacher perception; its findings cannot be imported as an Albay effect estimate.

The challenges in Table 6 complicate any simple interpretation of the positive effectiveness ratings. Teachers could view their current tools as useful while simultaneously encountering shortages, quality concerns, limited preparation time, and uneven support. Teacher readiness narrowly led the list, but the categories' rounded means were almost indistinguishable. Okoji and Olubayo (2022) similarly address resource availability and use as different constructs in a primary-school context. Subject, location, and sampling differences limit direct comparison, but the conceptual distinction supports collecting an inventory and usage log instead of assuming that a resource judged necessary is already accessible.

Tables 7 and 8 add a policy-oriented result: the areas agreed in ranking the offered responses more than they agreed in ranking the challenges. Sufficient stock and daily instructional time were highest. That pattern is consistent with teachers seeking tangible access and protected time, while local-language materials and planning time remained close priorities. A significant W for the 23 solution items does not mean all schools have the same shortage or that consensus proves a solution works. Budget, language, reading level, durability, and teacher preparation should guide a pilot. Schools could first document the supply of usable texts, which children can access them, and when tools are used, then evaluate any change in reading outcomes with an independent measure.

Keller's (2010) attention and relevance framework gives a reason to consider engaging, age-appropriate materials, but a motivational rationale cannot repair the absence of learner assessment. Smaldino et al. (2019) likewise connect media selection with instructional objectives and evaluation. The study's proposed 30-minute daily use could be implemented in varied ways: guided reading, manipulative-supported practice, or audiovisual instruction. These differ in content and likely in how their outcomes should be assessed. A future evaluation should define the specific tool, learning target, exposure, teacher preparation, and comparison condition before estimating an effect.

The inferential tests warrant a narrow interpretation. Kendall's W ranks questionnaire items across three geographic groups and tests whether their order is more concordant than expected under the study's null model. It neither estimates association between individual teachers' utilization and learner scores nor adjusts for clustering of teachers within schools. The original narrative occasionally describes the concordance as a "correlation," and its summary assigns significance to domains whose Table 18 statistics fall below their .05 critical values. This manuscript follows the tabulated decision rows, with the conflicts identified for author review. No causal attribution is made from these cross-sectional questionnaire summaries.

Several limitations follow from the measurement and reporting. The purposive sample and unspecified invitation denominator constrain representativeness. Teacher ratings on all parts of a single instrument may share response-style bias. The final questionnaire, validation records, pilot size, and coherent subscale reliability estimates are unavailable. The specification's 22 solution items conflict with 23 solution rows and the reported df of 22. Domain names and item placements differ between specification and results. The methods do not report independent observation or learner performance assessments, and exact administration dates are not supplied. These issues need resolution before a claim of journal-ready empirical reproducibility, though the descriptive record can be presented faithfully as reported.

The practical implication is to use the ratings as a needs-assessment starting point. A division or school could pair a materials inventory with teacher interviews about readiness and support, then select a feasible response from the ranked proposals. Predefined implementation measures might include the number of usable and appropriately leveled texts, lesson-level tool-use logs, planning time, and observed instructional tasks. Repeated learner reading assessments, with an appropriate comparison or careful longitudinal design, could then address the title's improvement claim. Until such evidence is collected, the study supports prioritization and further evaluation, not a conclusion that instructional tools improved performance.

4. Conclusion

The study reports high teacher ratings for utilization (4.45), perceived effectiveness (4.45), and challenges (4.36). Area-level item rankings showed no significant concordance for any utilization or challenge domain; significant concordance was reported for adequacy and learners' engagement and reading improvement within perceived effectiveness, and for the 23 proposed solutions. Teachers ranked textbooks and supplementary reading resources highest among the solutions. These results describe perceptions and priorities in the participating Albay schools. The study cannot determine whether instructional tools improved Grade 1 learners' performance, and the instrument and inferential-reporting conflicts require clarification before submission.

References

- [1] Francisco, C. D. C., & Celon, L. C. (2020). Teachers' instructional practices and its effects on students' academic performance. *International Journal of Scientific Research in Multidisciplinary Studies*, 6(7), 64–71. https://www.isroset.org/journal/IJSRMS/full_paper_view.php?paper_id=1998
- [2] Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. Springer. <https://doi.org/10.1007/978-1-4419-1250-3>
- [3] Okoji, O. O., & Olubayo, M. A. (2022). Assessment of availability and utilization of instructional resources for teaching primary school pupils' mathematics in Ilorin Metropolis, Nigeria. *African Journal of Educational Studies in Mathematics and Sciences*, 17(2), 1–12. <https://doi.org/10.4314/ajesms.v17i2.1>
- [4] Smaldino, S. E., Lowther, D. L., & Mims, C. (2019). *Instructional technology and media for learning* (12th ed.). Pearson.