

The Utilization of the Strategic Intervention Materials (SIM) In Improving the Mathematical Competence of the Grade 2 Learners in the Public Elementary Schools of Ragay Districts I And II, Division of Camarines Sur

Ponciana M. Reyes¹

¹School of Graduate Studies, University of Northeastern Philippines

ARTICLE INFORMATION	ABSTRACT
Article history: Published: October 2026	The study examined Grade 2 teachers' reported utilization of strategic intervention materials (SIM) in Mathematics 2 and their appraisals of mathematical competence in public elementary schools of Ragay Districts I and II, Division of Camarines Sur, during school year 2025–2026. A descriptive–evaluative–inferential survey used a 100-item questionnaire, weighted means, ranks, and reported Wilcoxon–Mann–Whitney comparisons. The abstract and methods state 53 teachers in 33 schools; the detailed district subtotals in the respondent table are 29 and 27, totaling 56, and require reconciliation. Overall SIM utilization was 4.30 (Much Utilized), led by monitoring and feedback (4.34). Teachers rated the competence domains at an overall 4.38 (Much Improved), highest for number sense and operations (4.46). The reported rank tests identified differences for three of five utilization domains and all four perceived competence domains, but several z–probability pairings and rank sums are inconsistent. The findings describe teachers' ratings, not independently measured learner improvement. The study identifies monitoring, regular implementation, and mathematical communication as areas for evaluation; the participant roster and inferential calculations need verification before publication.
Keywords: Strategic Intervention Materials Grade 2 mathematics Teacher Perceptions Numeracy Ragay Wilcoxon–Mann–Whitney test	

1. Introduction

Strategic intervention materials are designed to provide focused activities for learners who need additional practice with a particular competency. In early mathematics, such materials can represent number relationships, guide worked examples, and allow a teacher to see where a learner's reasoning fails. Their instructional value depends on alignment with the lesson, the quality of feedback, and opportunities to transfer a skill beyond the activity. A high rating for use does not, by itself, demonstrate that mathematical achievement rose. The study investigated this issue in two public school districts in Camarines Sur using Grade 2 teachers' appraisals.

The study's first set of domains addresses how SIMs are used: frequency, instructional fit, engagement, monitoring and feedback, and independent or remedial support. These represent distinct processes. A teacher may prepare a relevant activity but use it irregularly, or may use it frequently without checking what the learner understood. The second set concerns number sense, algebraic and logical thinking, geometry and measurement, and communication and disposition. The distinction matters because the questionnaire asks teachers to judge improvement in these areas without presenting a baseline and posttest or independently scored learner work.

Research involving a quick response embedded SIM with struggling mathematics learners has examined achievement with learner-level evidence (Abuda et al., 2019). That study concerned older learners and a particular material, so it cannot supply an effect size for Grade 2 pupils in Ragay. It does show why implementation ratings and achievement evidence should be kept separate. Effective professional development also involves content focus, active learning, collaboration, and sustained support rather than a single material distribution (Darling-Hammond et al., 2017). These perspectives frame the interpretation of the present teachers' responses without adding observations that were not collected.

The study asked (1) to what extent SIMs were utilized in teaching Mathematics 2 across five domains, (2) whether district rank orders differed in those domains, (3) how teachers rated Grade 2 mathematical competence across four domains, (4) whether district rank orders differed in those appraisals, and (5) what policy recommendations followed. This manuscript retains the original title and presents the reported statistics as printed. It distinguishes perceived competence from measured learner change and records material inconsistencies separately for author review.

2. Methods

2.1 Design, participants, and setting

The study described its design as descriptive–evaluative–inferential. It purposively selected Ragay Districts I and II as the setting and stated that the identified Grade 2 teachers were included by total enumeration. The methods and abstract report 53 respondents across 33 public elementary schools in school year 2025–2026. The respondent table lists 17 schools in District I and 16 in District II, but its school-level entries add to 29 and 27 teachers, respectively. These subtotals total 56, whereas the printed grand total elsewhere is 53. The table's 52% and 48% district shares align more closely with 29 and 27 out of 56. The number of teachers contributing to the analysis thus remains unresolved.

Table 1. Respondent accounting as reported

District	Schools listed	Teacher subtotal	Printed share
Ragay District I	17	29	52%
Ragay District II	16	27	48%
Combined listed subtotals	33	56	100%
Narrative and reported grand total	33	53	

Note. The combined subtotal of 56 is arithmetic from the printed school-level rows; the narrative and reported grand total state 53. No analytic denominator is inferred from the conflicting entries.

District I has more listed teachers than District II, although the magnitude depends on which count is correct. The inconsistency is consequential for the sampling frame, descriptive weighting, and any respondent-level comparison. It does not change the five and four domain means transcribed from the results tables, but those means cannot be independently reproduced without a corrected roster and response dataset. The report provides no teacher-level demographics, school-level clustering information, or nonresponse record. Generalization to schools outside the selected districts therefore remains limited.

2.2 Instrument and procedure

The researcher-developed questionnaire comprised 100 items according to its specification: 55 items in Part I covering five domains of SIM use, and 45 in Part II covering four mathematical-competence domains. The Part I item counts were 9, 11, 12, 11, and 12; Part II counts were 13, 12, 10, and 10. Responses used five-point scales. Scores from 3.50 to 4.49 were classified as Much Utilized in Part I and Much Improved in Part II; 4.50 to 5.00 received the corresponding Very Much labels. The methods describe questionnaire review, revision, and validation. A reported reliability coefficient was 0.91, with a reported *t* value of 22.18, but the reliability section identifies KR-21 for a five-point scale and does not provide the data or scoring needed to reproduce the calculation. Questionnaires were distributed to the selected Grade 2 teachers and retrieved for tallying. The chapter describes stages of approval and a dry run but does not specify exact administration dates, pilot size, missing-response handling, or how individual learner performance records were linked to teacher ratings. The category called support for independent and remedial learning in the specification is labeled “Reading” in the findings summary and test table, despite the mathematics focus. This manuscript retains the findings-table label in numerical presentations and uses the broader learning concept when interpreting the intended construct. No pupil test scores, comparison group, or pre/post assessment were reported.

2.3 Statistical treatment

Frequencies, percentages, weighted means, and ranks summarized ratings. The stated Wilcoxon–Mann–Whitney *U* procedure compared rank orders between the districts at $\alpha = .05$; the result tables report *z* statistics and associated probabilities. Item counts, rather than the 29/27 or 53/56 teachers, are shown as the two “No. of Samples” entries in each domain (for example, 9 and 9 in a nine-item domain). The unit entered into the comparison and the handling of ties are not explained. The printed *z*, *p*, and decision labels are reproduced as reported; they are not presented as independently validated respondent-level tests. A difference in ranks would concern the tested item summaries, not a causal effect of district or SIM on pupils.

3. Results

3.1 Reported SIM utilization

Table 2. Summary of SIM utilization in teaching Mathematics 2

Domain	District I Wx	District II Wx	Overall Wx	Interpretation	Rank
Frequency and Consistency of Use	4.26	4.20	4.23	MU	5
Instructional Integration and Appropriateness	4.30	4.34	4.32	MU	2.5
Learner Engagement during SIM Activities	4.28	4.33	4.31	MU	4
Monitoring and Feedback	4.44	4.23	4.34	MU	1
Support for Independent and Remedial Reading	4.34	4.29	4.32	MU	2.5
Average	4.32	4.28	4.30	MU	

Note. Wx = weighted mean; MU = Much Utilized. Values, category labels, and ranks follow the study’s Table 8. “Reading” is its printed label, although the instrument specification names “Learning.”

All five utilization domains were classified Much Utilized. Monitoring and feedback was highest at 4.34, and frequency and consistency lowest at 4.23. The 0.21 gap between District I and District II in monitoring and feedback is more visible than the small overall differences between most domains, but no distribution of individual responses is provided. A high monitoring rating may reflect teacher attention to checking work and giving corrective guidance; the survey does not show the form, timing, or accuracy of that feedback. The lower frequency rating suggests a practical question about how regularly SIMs are incorporated into lessons, rather than proof of insufficient provision.

Table 3. Reported district rank comparisons for SIM utilization

Domain	Items per district	<i>z</i>	Reported <i>p</i>	Reported decision
Frequency and Consistency of Use	9	–1.06	0.2891	Not significant
Instructional Integration and Appropriateness	11	–1.51	0.1310	Not significant
Learner Engagement during SIM Activities	12	–2.17	0.0300	Significant (.05)
Monitoring and Feedback	11	–3.94	0.0001	Significant (.001)
Independent and Remedial Reading	12	–2.05	0.0404	Significant (.05)

Note. Statistics and decisions reproduce the study's Table 9. "Items per district" is the table's "No. of Samples," not a verified respondent count. Test direction and ranked observations require author confirmation.

The reported comparisons crossed the .05 threshold for engagement, monitoring and feedback, and independent/remedial support, but not for frequency or instructional integration. This pattern says that the ranked indicators differed under the reported test; it does not establish which district had better SIM implementation or whether pupils benefited. Monitoring and feedback had the most extreme printed *z* and *p* values. Its district means in Table 2 place District I higher, yet the precise rank-test direction depends on the actual ranked inputs. Results for the other two significant domains also need reanalysis after the participant roster and test unit are clarified.

3.2 Teachers' appraisals of mathematical competence

Table 4. Summary of perceived mathematical-competence domains

Domain	District I W _x	District II W _x	Overall W _x	Interpretation	Rank
Number Sense and Operations	4.51	4.42	4.46	MI	1
Algebraic and Logical Thinking	4.33	4.40	4.36	MI	3
Geometry, Measurements, and Data	4.40	4.34	4.37	MI	2
Mathematical Communication and Disposition	4.29	4.40	4.34	MI	4
Average	4.38	4.39	4.38	MI	

Note. W_x = weighted mean; MI = Much Improved. Values and labels reproduce Table 14. The preliminary abstract prints "4.3)" for geometry, whereas the findings table gives 4.37.

Teachers rated all four domains Much Improved. Number sense and operations was highest at 4.46; mathematical communication and disposition was lowest at 4.34. The word "improved" is the questionnaire's response category, not a calculated change score. District I's number-sense mean of 4.51 crosses the Very Much Improved threshold, while District II's 4.42 is Much Improved, but the combined mean remains below 4.50. The lower communication rating raises a question about whether pupils are given opportunities to explain, represent, and justify solutions. A review of actual pupil work would be necessary to determine the scale of any gap.

Table 5. Reported district rank comparisons for perceived mathematical competence

Domain	Items per district	<i>z</i>	Reported <i>p</i>	Reported decision
Number Sense and Operations	13	-3.18	0.0007	Significant (.001)
Algebraic and Logical Thinking	12	-2.17	0.0150	Significant (.01)
Geometry, Measurement, and Data	10	-2.38	0.0005	Significant (.001)
Mathematical Communication and Disposition	10	-3.14	0.0008	Significant (.001)

Note. Statistics and decisions reproduce Table 15 exactly. Several *z*-*p* combinations and a rank-sum total do not reconcile with a conventional two-sided test; the decision labels are therefore provisional.

Every perceived-competence domain was marked significant in the reported comparison. These tests concern two district rankings of item-level appraisals; no learner achievement scores entered the table. The reported *p* for algebraic/logical thinking is 0.0150, which is below .05 but above .01 even though its printed significance label is ".01." The geometry row pairs *z* = -2.38 with *p* = 0.0005, an especially pronounced probability mismatch. The number-sense rank sums also do not total the expected ranks for two sets of thirteen items. These issues require the raw ranks and calculations before an inferential claim about district differences can be relied on.

3.3 Selected item patterns

Table 6. Selected frequency and consistency items for SIM use

Item	District I W _x	District II W _x	Overall W _x	Interpretation
SIMs regularly incorporated daily or weekly to reinforce targeted competencies	4.31	4.26	4.28	MU
SIMs used across introduction, development, and mastery phases	4.24	4.26	4.25	MU
SIMs used in scheduled remediation for difficulty in specific competencies	4.34	4.22	4.28	MU
SIMs applied across a wide range of Mathematics topics	4.17	4.19	4.18	MU
SIMs used immediately when assessment indicates non-mastery	4.14	4.26	4.20	MU
SIMs systematically included in lesson plans	4.41	4.07	4.24	MU

Note. Selected items and figures follow the study's Table 3 and continuation. MU = Much Utilized. Item wording is shortened without changing its meaning.

The frequency domain's item-level pattern sharpens the interpretation of its 4.23 mean. Regular daily or weekly incorporation and scheduled remediation both reached 4.28, whereas use across a wide range of topics was 4.18 and immediate use following non-mastery was 4.20. Systematic inclusion in lesson plans differed between districts (4.41 versus 4.07), although the overall item was 4.24. The data suggest that scheduled or familiar SIM use may be more established than flexible, assessment-triggered use. This is an interpretation of ratings; the questionnaire did not track dates of sessions or match a learner's error to a particular SIM. A lesson-log review would test whether the reported pattern reflects actual practice.

Table 7. Selected mathematical communication and disposition items

Item	District I W _x	District II W _x	Overall W _x	Interpretation
Clearly explain mathematical thinking and solutions	4.24	4.37	4.31	MI
Use appropriate mathematical terms and vocabulary	4.28	4.33	4.30	MI
Express ideas using symbols, drawings, and written explanations	4.24	4.33	4.29	MI

Item	District I Wx	District II Wx	Overall Wx	Interpretation
Participate in discussions and group activities	4.38	4.41	4.39	MI
Justify and defend answers using logical reasoning	4.31	4.44	4.38	MI

Note. Selected items reproduce the means and labels in the study's Table 13 and continuation. MI = Much Improved, the teachers' rating category rather than an observed change score.

The lowest of these communication items was expression through symbols, drawings, and writing (4.29), while participation reached 4.39 and justification reached 4.38. Participation is easier to see during an activity than the quality of mathematical explanation. A child may contribute to a group without articulating why a solution works. Conversely, a strong written representation may not be captured by a participation rating. These distinctions suggest a need to score actual verbal and written responses with a clear rubric. The reported means cannot show whether SIMs improved either form of communication over time. They also do not distinguish a mathematics vocabulary need from a general language demand in a Grade 2 classroom.

Taken together, Tables 6 and 7 suggest a practical pairing for later evaluation: define a least-mastered competency, document why a SIM was selected, use it at an identifiable point after assessment, and collect a short explanation from the pupil. Such a record would make the connection between assessment-triggered use and mathematical communication explicit. It would also distinguish a teacher's intention to use SIMs from actual exposure and a learner's completed task from demonstrated reasoning. The present survey did not contain these linked observations; the sequence is a proposed method for testing the study's interpretation.

4. Discussion

The first research question yields a favorable profile of SIM utilization, but the lowest domain is regular use. A material may be highly valued in a targeted remediation session while being hard to sustain across the curriculum. The high monitoring and feedback domain suggests that teachers associate SIM with identifying errors and responding to learner work. In a useful instructional cycle, a specific misconception would trigger a suitable SIM task, the teacher would observe the response, and feedback would inform a subsequent task. The questionnaire gives no direct record of that cycle, so its presence remains a hypothesis for classroom-based study rather than an observed mechanism.

The second research question produced three reported district differences in item rankings. Because the tables list equal numbers of items as "samples" on both sides, the comparison appears to operate on sets of domain indicators rather than on all teachers' independent responses. The study does not document the tested unit, ties, or score construction. It would be inappropriate to attribute the contrasts to district management or teacher competence. Schools may nevertheless inspect the specific monitoring and remedial practices that differ locally, using common observation criteria and equivalent tools, rather than drawing allocation decisions from the provisional p values alone.

For the third question, number sense led and mathematical communication was lowest. This ordering is educationally plausible but remains a report of teacher judgment. Abuda et al. (2019) examined a specific mathematics SIM with learner outcomes; its very different grade level and intervention caution against treating any SIM as a single uniform exposure. A local follow-up could identify the particular numeracy skills targeted, analyze pupil solutions, and distinguish procedural accuracy from explanation and representation. Repeated assessments would allow the title's claim of improvement to be tested directly, preferably with a suitable comparison and documented SIM exposure.

The fourth question reports statistical differences in all four competence rankings, but the probability and rank-sum inconsistencies weaken them. A reader cannot resolve whether the p values were one-sided, two-sided, calculated from ties, or transcribed incorrectly. Importantly, the two districts' overall competence means are close, 4.38 and 4.39. Such means do not prove equivalence, but they also do not support a broad assertion that one district's pupils learned more. The statistical analysis must be rerun on clearly identified observations and accompanied by transparent sample sizes, rank sums, U values, tie handling, test tails, and exact probabilities. The practical policy question is what evidence would guide improvement. Teachers could catalogue which SIMs target least-mastered Grade 2 competencies, record when each is used, and collect short work samples before and after targeted lessons. Monitoring should include the feedback that pupils receive, not only completion of activities. Darling-Hammond et al. (2017) describe professional development that combines content, active learning, collaboration, and coaching; this supports a plan for teachers to review pupil work together and refine SIM design. The study did not evaluate such training, so it is a proposed response to the descriptive findings, not a measured result.

The study has additional limitations. The conflicting teacher totals prevent a reliable response rate and denominator. The five-point scales may produce favorable response patterns, and no independent observation or learner-level assessment triangulates the ratings. School effects and teacher experience were not modeled. The label "Reading" in two mathematics tables may be a copy or construct error. The reported KR-21 coefficient for a five-point questionnaire needs a suitable reliability analysis and item-level documentation. Finally, the two districts were purposively chosen. These limitations restrict generalization beyond the sampled schools and preclude a causal claim that SIM utilization improved mathematical competence.

An evaluable next phase could compare targeted competencies before and after a defined SIM sequence while documenting alternative instruction, learner attendance, baseline attainment, and teacher implementation. Pupil work would be scored with an explicit rubric for number sense, reasoning, geometry/data, and mathematical communication. The units of assignment and analysis would need to reflect pupils nested within classes and schools. This design would answer a different, stronger question than the present survey: whether, how much, and for whom a specified intervention contributes to learning. The current ratings remain useful for deciding what to investigate first.

5. Conclusion

The study reports SIM utilization as Much Utilized overall (4.30), with monitoring and feedback highest (4.34), and perceived mathematical competence as Much Improved (4.38), with number sense highest (4.46). Three utilization domains and four perceived-competence domains were labeled significant in district rank comparisons. Those inferential labels require verification

because the participant totals, some rank sums, and several z - p pairs conflict. The study provides a local picture of teacher perceptions and a rationale for closer examination of regular implementation and mathematical communication; it does not independently demonstrate learner improvement.

References

- [1] Abuda, B. F. Q., Balazo, G. F. O., Orque, J. C., Cabili, M. C. D., & Maestre, M. F. A. M. (2019). Struggling learners' mathematics achievement level using quick response embedded strategic intervention material. *International Journal in Information Technology in Governance, Education and Business*, 1(1), 39–45. <https://doi.org/10.32664/ijitgeb.v1i1.11>
- [2] Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. Learning Policy Institute. <https://learningpolicyinstitute.org/product/effective-teacher-professional-development-report>