

FROM BLOCKS TO NUMBERS: EFFECTS OF A MANIPULATIVE-BASED INTERVENTION ON THE EARLY NUMERACY OF LOW-PERFORMING GRADE TWO LEARNERS IN A PHILIPPINE PUBLIC SCHOOL

Authors Details

Norben Pabondia Agas^{1*}

Sully Elementary School, Loudon County Public Schools, Virginia, USA.

Corresponding Author: *Norben Pabondia Agas*

ABSTRACT: Background: Weak foundational numeracy in the early grades remains a pressing concern in the Philippines, and teachers need low-cost, classroom-feasible strategies for learners who fall behind. Physical manipulatives are widely recommended, yet evidence from Philippine public-school classrooms is thin and largely unpublished.

Purpose: This classroom action research examined whether a manipulative-based intervention, delivered through a concrete–representational–abstract (CRA) sequence, was associated with improved numeracy performance among low-performing Grade 2 learners.

Method: A single-group pretest–posttest design was used. Twelve Grade 2 learners with the lowest scores on a beginning-of-year numeracy diagnostic in one public elementary school in northern Luzon were purposively selected. Over eight weeks, lessons on place value, addition with and without regrouping, properties of addition, ordinal numbers, and basic shapes and slides were taught using base-ten blocks, number lines, and pattern and attribute blocks. A 15-item, expert-validated, researcher-made test was administered before and after the intervention. Data were analyzed with descriptive statistics, a paired-samples *t* test, a Wilcoxon signed-rank test, standardized effect sizes, and normalized gains.

Results: Mean scores rose from 3.17 (*SD* = 1.27; 21.1% correct) to 12.33 (*SD* = 1.37; 82.2% correct), a mean gain of 9.17 points (95% CI [8.23, 10.10]); *t* (11) = 21.65, *p* < .001,

$dz = 6.25$. All 12 learners moved from the low to the high performance band, and the Wilcoxon test confirmed the result ($p < .001$). The average normalized gain was .78.

Conclusion: Learners improved substantially while taught with manipulatives. Because the study lacked a comparison group and used the same items at pretest and posttest, the magnitude of the effect is likely overstated and causal claims remain tentative. The findings nonetheless support the structured, CRA-based use of manipulatives as a practical intervention for early-grade learners at risk of mathematics difficulties and justify a controlled replication.

Keywords: *manipulatives; early numeracy; concrete–representational–abstract; low-performing learners; place value; action research; Philippines*

1. Introduction

Numeracy—the capacity to understand and use numbers, quantities, and basic operations with confidence—is one of the most consequential outcomes of early schooling. Knowledge of number symbols, magnitudes, place value, and simple arithmetic built in the first grades forms the base on which later mathematics rests, and research on arithmetic fluency shows that children who fail to develop efficient and flexible strategies with whole numbers early on tend to struggle with more advanced content later (McNeil et al., 2025). Studies bridging numerical cognition and classroom practice likewise underline that early number instruction is most effective when teachers explicitly connect quantities, symbols, and spatial representations of number (Hawes et al., 2021).

In the Philippines, the need to strengthen these foundations is acute. In the Programme for International Student Assessment (PISA) 2022, Filipino 15-year-olds were among the lowest-performing participants in mathematics, with most students scoring below the baseline proficiency level (OECD, 2023). Secondary analyses of earlier PISA data show that low mathematics performance is concentrated among public-school learners and is associated with a cluster of student and home factors that begin to operate well before adolescence (Bernardo et al., 2022). School closures during the COVID-19 pandemic compounded these gaps; a systematic review of 42 studies found substantial learning deficits that were largest in lower- and middle-income countries and among socially disadvantaged children (Betthäuser et al., 2023). In response, the Department of Education

introduced the MATATAG curriculum, which reorients the early grades toward mastery of foundational literacy and numeracy (Department of Education [DepEd], 2024). Classroom-level interventions that help learners who are already falling behind are therefore a practical priority for teachers and school leaders.

Grade 2 mathematics presents specific hurdles for such learners. Understanding that the digit 6 in 642 represents six hundreds, decomposing numbers into expanded form, and regrouping ten ones into one ten are conceptually demanding because the written symbols conceal the quantities they stand for. When instruction moves too quickly to symbolic procedures, some children memorize steps without understanding them, make persistent place-value errors, and begin to associate mathematics with failure. Mathematics anxiety, which shows a small-to-moderate negative relation with achievement across age groups (Barroso et al., 2021), can further depress the performance of learners who have repeatedly experienced difficulty.

1.1 Manipulatives and the concrete–representational–abstract sequence

Manipulatives are physical objects—such as counters, base-ten blocks, number lines, fraction tiles, and pattern blocks—that learners handle to model mathematical ideas. A scoping review of 102 intervention studies with pre-primary and primary-age children found that physical manipulatives are used most often in mathematics and frequently yield learning benefits, but it also noted that the evidence base comes almost entirely from high- and middle-income contexts, predominantly the United States (Byrne et al., 2023). For learners at risk of or identified with mathematics difficulties, a meta-analysis of 53 single-case studies reported a large omnibus effect of manipulative-based interventions ($\text{Tau-}U = 0.91$; Peltier et al., 2020).

Manipulatives are most effective when they are embedded in an explicit instructional sequence rather than used as free play. In the concrete–representational–abstract (CRA) sequence, learners first solve problems with objects, then with drawings or other visual representations of those objects, and finally with numerals and symbols alone. A recent meta-analysis confirmed that CRA instruction produces meaningful gains, especially for basic operations in the primary grades (Ebner et al., 2025). Integrated variants in which concrete, pictorial, and symbolic representations are presented together and then faded have improved number sense and addition performance among young learners receiving

supplementary support (Flores & Hinton, 2022; Hinton & Flores, 2022). Practical guidance for teachers emphasizes matching the manipulative to the targeted concept, knowing the learner, and planning how representations will be faded (Bouck et al., 2022).

The benefits of manipulatives are not automatic. Children who are guided to treat the objects as tools for doing mathematics, rather than as toys, show better learning, transfer, and conceptual understanding than those who are not (Donovan & Alibali, 2021). The theoretical premise of the present study therefore follows the CRA logic: concrete objects give learners a perceptual and physical model of quantity and structure (for example, a rod of ten unit cubes), drawings provide an intermediate representation, and teacher talk explicitly links each representation to its written symbol until learners can operate with symbols alone.

1.2 The present study

Although manipulatives are commonly recommended in Philippine teacher guides, published empirical studies from Philippine public elementary classrooms that describe the intervention in replicable detail, analyze repeated-measures data appropriately, and report effect sizes remain scarce. Much of the local evidence consists of unpublished graduate action research, which limits its usefulness for teachers and policymakers. This study responds to that gap by documenting a manipulative-based, CRA-sequenced intervention for low-performing Grade 2 learners in a public elementary school and by analyzing its outcomes with appropriate paired-data statistics and effect sizes.

The study addressed three research questions:

RQ1. What was the numeracy performance of low-performing Grade 2 learners before the manipulative-based intervention?

RQ2. What was their numeracy performance after the intervention?

RQ3. Was there a statistically significant difference between pretest and posttest scores?

The null hypothesis (H_0) stated that there is no significant difference between the learners' pretest and posttest numeracy scores ($\alpha = .05$, two-tailed).

2. Methods

2.1 Research design

The study was classroom action research conducted by the class teacher, using a single-group pretest–posttest (pre-experimental) design. One group of learners was assessed, received the intervention, and was assessed again with the same instrument. Because there was no randomization and no comparison group, the design can establish change over time but cannot, on its own, rule out alternative explanations such as maturation, regular classroom instruction, testing effects, or regression toward the mean (Miller et al., 2020). These threats are addressed in the Discussion.

2.2 Setting and participants

The study took place in a public elementary school in a rural municipality in northern Luzon, Philippines. To protect the privacy of the school, its learners, and its personnel, the school, district, and personnel are not named, and all learners are identified only by codes (L01–L12).

Participants were 12 Grade 2 learners from one intact class. They were selected through purposive sampling (Campbell et al., 2020) because they obtained the lowest scores in the class on the school’s beginning-of-year numeracy diagnostic assessment. In this paper they are described as *low-performing learners*; none had a formal diagnosis of a learning disability. The term replaces the label “slow learners,” which is imprecise and potentially stigmatizing.

2.3 Intervention

The intervention ran for eight weeks during the school year 2024-2025.. It targeted first-quarter Grade 2 learning competencies. Table 1 maps each competency to the manipulatives used.

Table 1 Target Competencies and Manipulatives Used in the Intervention

Competency area	Manipulatives	Test items
Place value, value of a digit, and expanded form of three-digit numbers	Base-ten blocks (units, rods, flats); place-value mats	4, 5, 8, 12
Addition without and with regrouping	Base-ten blocks; number lines	3, 9, 10, 11, 13, 15

Competency area	Manipulatives	Test items
Properties of addition (order and grouping)	Base-ten blocks; counters	6, 14
Ordinal numbers	Picture/object sequences	7
Shapes and slides (basic transformations)	Pattern blocks; attribute blocks; pattern tiles	1, 2

Note. Item numbers refer to the pretest form. The posttest contained the same items in a different order.

Each lesson followed a CRA routine. In the concrete phase, the teacher modeled the concept with objects (for example, trading ten unit cubes for one rod when the ones column exceeded nine) while thinking aloud, and learners then manipulated the objects in pairs. In the representational phase, learners recorded the same actions as drawings of blocks and as entries on a place-value chart. In the abstract phase, learners solved the corresponding numeral-only exercises, and the teacher repeatedly asked them to explain how the written step matched what they had done with the blocks. Throughout, the manipulatives were introduced explicitly as tools for doing mathematics (Donovan & Alibali, 2021), and concrete support was faded as learners showed accuracy.

2.4 Instrument

Numeracy performance was measured with a 15-item multiple-choice test developed by the teacher-researcher from the Grade 2 curriculum guide for the first quarter. A table of specifications distributed the items across the targeted competencies and across cognitive levels of the revised Bloom's taxonomy. Each correct answer earned one point (possible range 0–15). The same 15 items, presented in a different order and with the options rearranged, served as the posttest.

Content validity was established by three evaluators—a school head, a master teacher, and the school numeracy coordinator—who rated the items for objectivity, specificity, and attainability on a five-point scale. The mean rating was 4.56, interpreted as very high. The test was then tried out with Grade 2 learners who were not part of the study to confirm that the items and language were developmentally appropriate, and items were revised on the basis of the try-out and evaluator comments.

2.5 Procedure

After written permission was obtained from the school head, parents were informed of the study and gave consent, and learners gave verbal assent. The pretest was administered in the week before the intervention, and the posttest in the week after it ended. Both were administered in the classroom by the teacher-researcher, with items read aloud to control for differences in reading ability.

2.6 Data analysis

Scores were summarized with means, standard deviations, medians, and ranges. For interpretation, scores were classified into three performance bands: low (0–5), average (6–10), and high (11–15). Because each learner was measured twice, pretest and posttest scores were compared with a paired-samples *t* test (two-tailed, $\alpha = .05$). The normality of the difference scores was checked with the Shapiro–Wilk test. Given the small sample, an exact Wilcoxon signed-rank test was run as a distribution-free robustness check. Effect size was reported as Cohen's *d*_z (mean difference divided by the standard deviation of the differences) and as Hedges' *g*_{av} (mean difference divided by the average of the pretest and posttest standard deviations, with small-sample correction), together with the 95% confidence interval of the mean difference. Individual normalized gains were computed as $(\text{posttest} - \text{pretest}) / (15 - \text{pretest})$. Analyses were run in Python (SciPy).

2.7 Ethical considerations

Participation was voluntary, and learners who declined would have received the same lessons as part of regular instruction. No names, photographs, or other identifying information about the learners or the school are reported. Data were stored by the researcher and used only for this study.

3. Results

3.1 Numeracy performance before the intervention (RQ1)

Before the intervention, the learners answered on average about one in five items correctly ($M = 3.17$, $SD = 1.27$, $Mdn = 3.00$, range 1–5). As Table 2 shows, all 12 learners fell in the low performance band. This was expected because the learners were selected on the basis of low diagnostic scores, and it confirms that the group entered the intervention with limited mastery of the targeted Grade 2 competencies.

Table 2 Descriptive Statistics and Distribution of Performance Levels at Pretest and Posttest (N = 12)

Statistic / performance level	Pretest	Posttest
<i>M</i> (<i>SD</i>)	3.17 (1.27)	12.33 (1.37)
Mean percentage correct	21.1%	82.2%
<i>Mdn</i>	3.00	12.00
Range	1–5	11–15
High (11–15), <i>f</i> (%)	0 (0.0)	12 (100.0)
Average (6–10), <i>f</i> (%)	0 (0.0)	0 (0.0)
Low (0–5), <i>f</i> (%)	12 (100.0)	0 (0.0)

Note. Maximum possible score = 15.

3.2 Numeracy performance after the intervention (RQ2)

After eight weeks, the mean score rose to 12.33 ($SD = 1.37$, $Mdn = 12.00$, range 11–15), or about 82% of items answered correctly. All 12 learners reached the high performance band, and one learner (L11) obtained a perfect score. Table 3 presents the individual results. Every learner improved, with raw gains ranging from 7 to 12 points and normalized gains from .67 to 1.00 ($M = .78$, $SD = .11$), meaning that, on average, learners closed about three-quarters of the distance between their pretest score and the maximum.

Table 3 Individual Pretest and Posttest Scores, Gains, and Normalized Gains

Learner	Pretest	Posttest	Gain	Normalized gain
L01	2	11	9	.69
L02	4	12	8	.73
L03	5	12	7	.70
L04	3	11	8	.67
L05	5	13	8	.80
L06	4	14	10	.91
L07	3	11	8	.67
L08	2	11	9	.69
L09	1	12	11	.79
L10	4	14	10	.91
L11	3	15	12	1.00
L12	2	12	10	.77
<i>M</i> (<i>SD</i>)	3.17 (1.27)	12.33 (1.37)	9.17 (1.47)	.78 (.11)

Note. Normalized gain = (posttest – pretest) / (15 – pretest).

3.3 Difference between pretest and posttest scores (RQ3)

The difference scores did not depart from normality (Shapiro–Wilk $W = 0.94$, $p = .448$), so the paired-samples t test was appropriate. As Table 4 shows, posttest scores were significantly higher than pretest scores, $t(11) = 21.65$, $p < .001$, with a mean gain of 9.17 points, 95% CI [8.23, 10.10]. The exact Wilcoxon signed-rank test led to the same conclusion ($W = 0$, $p < .001$; all 12 differences positive; rank-biserial $r = 1.00$). The null hypothesis was therefore rejected. The standardized effect was very large ($dz = 6.25$; $gav = 6.70$).

Table 4 Paired Comparison of Pretest and Posttest Scores (N = 12)

Test	Statistic	<i>df</i>	<i>p</i>	Effect size
Paired-samples <i>t</i> test	$t = 21.65$	11	$< .001$	$dz = 6.25$; $gav = 6.70$
Wilcoxon signed-rank (exact)	$W = 0$	—	$< .001$	rank-biserial $r = 1.00$
Mean difference [95% CI]	9.17 [8.23, 10.10]			

Note. The original thesis report applied an independent-samples (Welch) t test to these data. Because the same learners were tested twice, the paired analysis reported here is the appropriate test; the conclusion is unchanged.

4. Discussion

This study examined whether a manipulative-based, CRA-sequenced intervention was associated with better numeracy among 12 low-performing Grade 2 learners. Learners who began the intervention answering roughly one item in five correctly were answering more than four in five correctly eight weeks later; every learner improved, and every learner moved from the low to the high performance band.

The direction of these results is consistent with the international evidence. Manipulative-based interventions have produced large effects for learners at risk of mathematics difficulties (Peltier et al., 2020), CRA instruction has shown meaningful gains in basic operations in the primary grades (Ebner et al., 2025), and integrated CRA approaches have improved young learners' number sense and addition (Flores & Hinton, 2022; Hinton & Flores, 2022). The competencies on which this group made progress—place value, expanded form, and regrouping—are precisely those in which a physical model makes the hidden structure of the base-ten system visible: a learner who has repeatedly exchanged ten unit cubes for one rod has a concrete referent for “carrying one ten,” which the symbolic

procedure alone does not provide. The explicit linking of objects, drawings, and numerals in each lesson, together with the framing of the materials as mathematical tools, reflects conditions that research associates with effective manipulative use (Bouck et al., 2022; Donovan & Alibali, 2021). The findings also add a classroom example from a Philippine public school to a literature that, as Byrne et al. (2023) observed, draws almost entirely on higher-income settings.

The magnitude of the change, however, must be interpreted with considerable caution. An effect of $dz = 6.25$ is far larger than those typically found for educational interventions; by commonly used benchmarks, effects of about 0.20 standard deviations on standardized achievement measures are already considered large (Kraft, 2020). Several features of the design likely inflated the observed effect. First, the outcome was a short, researcher-made test closely aligned with the lessons taught, and proximal measures of this kind consistently yield larger effects than broad standardized tests (Kraft, 2020). Second, the posttest reused the pretest items, so familiarity with the questions may have contributed to the gain. Third, learners were selected because they scored lowest, so part of the improvement may reflect regression toward the mean. Fourth, eight weeks of regular Grade 2 instruction and ordinary maturation occurred alongside the intervention, and without a comparison group their contribution cannot be separated from that of the manipulatives (Miller et al., 2020). Finally, the teacher who delivered the intervention also administered and scored the tests. Standardized effect sizes are further inflated when, as here, the spread of scores within a small, homogeneous group is narrow. The results are best read as strong evidence that these learners improved while taught with manipulatives, and as encouraging but not conclusive evidence that the manipulatives caused the improvement.

4.1 Implications for practice

For teachers, the study illustrates a feasible routine that uses inexpensive, often locally available materials: model the concept with objects, have learners record the same action pictorially, and then move to numerals while repeatedly asking learners to connect each written step to what they did with the objects. For school heads, the results support provisioning classrooms with basic sets of base-ten blocks and number lines and building time for supplementary numeracy sessions into the school learning recovery plan. Within

the foundational-numeracy emphasis of the MATATAG curriculum (DepEd, 2024), structured manipulative use offers one practical option for learners who are not yet ready for purely symbolic work, although its effects should be monitored with measures that go beyond the lessons taught.

4.2 Limitations and directions for future research

The study has limitations beyond those already noted. The sample was small and drawn from one class in one school, which limits generalizability. The instrument was brief, its reliability was not established with a formal coefficient, and it included a few geometry items (shapes and slides) alongside number items, so the total score is a broad indicator of first-quarter mathematics rather than a pure numeracy measure. No delayed posttest was given, so it is unknown whether the gains were retained, and implementation fidelity was not systematically recorded.

Future studies should (a) include a comparison group receiving regular instruction or an alternative intervention, ideally with random assignment of classes; (b) use parallel test forms or a standardized early-numeracy assessment, with reported reliability; (c) add a delayed posttest to examine retention and transfer to symbol-only tasks; (d) document dosage and fidelity; (e) compare concrete and virtual manipulatives, which may be more scalable; and (f) collect qualitative data on learners' strategies and attitudes, including mathematics anxiety, to explain how the manipulatives support learning.

5. Conclusion

Low-performing Grade 2 learners in a Philippine public elementary school made large and consistent gains in first-quarter numeracy competencies—especially place value and addition with regrouping—during an eight-week intervention in which manipulatives were used within an explicit concrete–representational–abstract sequence. Although the single-group design and the reuse of test items mean that the size of the effect is likely overstated and cannot be attributed to the manipulatives alone, the consistency of improvement across all learners, together with converging international evidence, supports the structured use of manipulatives as a practical, low-cost intervention for learners at risk of early mathematics difficulties. A controlled replication with parallel measures and a delayed follow-up is the necessary next step.

References

1. Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, 147(2), 134–168. <https://doi.org/10.1037/bul0000307>
2. Bernardo, A. B. I., Cordel, M. O., II, Lapinid, M. R. C., Teves, J. M. M., Yap, S. A., & Chua, U. C. (2022). Contrasting profiles of low-performing mathematics students in public and private schools in the Philippines: Insights from machine learning. *Journal of Intelligence*, 10(3), Article 61. <https://doi.org/10.3390/jintelligence10030061>
3. Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, 7(3), 375–385. <https://doi.org/10.1038/s41562-022-01506-4>
4. Bouck, E. C., Anderson, R. D., Long, H., & Sprick, J. (2022). Manipulative-based instructional sequences in mathematics for students with disabilities. *TEACHING Exceptional Children*, 54(3), 178–190. <https://doi.org/10.1177/0040059921994599>
5. Byrne, E. M., Jensen, H., Thomsen, B. S., & Ramchandani, P. G. (2023). Educational interventions involving physical manipulatives for improving children's learning and development: A scoping review. *Review of Education*, 11(2), Article e3400. <https://doi.org/10.1002/rev3.3400>
6. Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing*, 25(8), 652–661. <https://doi.org/10.1177/1744987120927206>
7. Department of Education. (2024). *Policy guidelines on the implementation of the MATATAG curriculum* (DepEd Order No. 10, s. 2024). <https://www.deped.gov.ph>
8. Donovan, A. M., & Alibali, M. W. (2021). Toys or math tools: Do children's views of manipulatives affect their learning? *Journal of Cognition and Development*, 22(2), 281–304. <https://doi.org/10.1080/15248372.2021.1890602>

9. Ebner, S., MacDonald, M. K., Grekov, P., & Aspiranti, K. B. (2025). A meta-analytic review of the concrete-representational-abstract math approach. *Learning Disabilities Research & Practice, 40*(1), 31–42. <https://doi.org/10.1177/09388982241292299>
10. Flores, M. M., & Hinton, V. M. (2022). The effects of a CRA-I intervention on students' number sense and understanding of addition. *Remedial and Special Education, 43*(3), 183–194. <https://doi.org/10.1177/07419325211038009>
11. Hawes, Z., Merkley, R., Stager, C. L., & Ansari, D. (2021). Integrating numerical cognition research and mathematics education to strengthen the teaching and learning of early number. *British Journal of Educational Psychology, 91*(4), 1073–1109. <https://doi.org/10.1111/bjep.12421>
12. Hinton, V., & Flores, M. (2022). Concrete-representational-abstract–integrated as a tier 2 instruction to teach addition. *Rural Special Education Quarterly, 41*(3), 169–178. <https://doi.org/10.1177/87568705221075756>
13. Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
14. McNeil, N. M., Jordan, N. C., Viegut, A. A., & Ansari, D. (2025). What the science of learning teaches us about arithmetic fluency. *Psychological Science in the Public Interest, 26*(1), 10–57. <https://doi.org/10.1177/15291006241287726>
15. Miller, C. J., Smith, S. N., & Pugatch, M. (2020). Experimental and quasi-experimental designs in implementation research. *Psychiatry Research, 283*, Article 112452. <https://doi.org/10.1016/j.psychres.2019.06.027>
16. OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
17. Peltier, C., Morin, K. L., Bouck, E. C., Lingo, M. E., Pulos, J. M., Scheffler, F. A., Suk, A., Mathews, L. A., Sinclair, T. E., & Deardorff, M. E. (2020). A meta-analysis of single-case research using mathematics manipulatives with students at risk or identified with a disability. *The Journal of Special Education, 54*(1), 3–15. <https://doi.org/10.1177/0022466919844516>