Title:

Impact and Implementation Analyses of the ELM Kindergarten Mathematics Intervention

Authors and Affiliations:
Christian Doabler, PhD (University of Oregon)
Scott K. Baker, PhD (University of Oregon)
Keith Smolkowski, PhD (Oregon Research Institute)
Hank Fien, PhD (University of Oregon)
Ben Clarke, PhD (University of Oregon)
Mari Strand Cary, PhD (University of Oregon)
David Chard, PhD (Southern Methodist University)

Abstract Body

Background:

For many years, educators have been concerned about the low level of mathematics performance of U.S. students in relation to national standards and in international comparisons (National Research Council, 2001). Difficulties in mathematics achievement are particularly severe for students from low income and minority backgrounds (Aud et al., 2011). Signs of these problems appear early. Significant differences between subpopulations of students can be reliably measured at school entry on measures related to counting principles and number knowledge to more complex understandings of quantities, operations, and problem solving (Griffin, Case, & Siegler, 1994; Jordan et al., 2006). Unless these differences are addressed as early as kindergarten, they are likely to persist and become more difficult to remediate over time (Jordan, Kaplan, & Hanich 2002). One potential approach to improving kindergarten and later math achievement is the delivery of effective instructional programs to all students as they enter school to bridge the emerging achievement gap and promote early mathematics learning.

The purpose of this 4-year efficacy trial, funded by IES under the Mathematics and Science Education topic (Baker, Chard, Clarke, Smolkowski, & Fien, 2008), is to study the efficacy of a (Tier 1) core kindergarten math curriculum, *Early Learning in Mathematics* (ELM), when implemented under rigorous experimental conditions. In the first year of this study (2008-09), kindergarten classrooms were randomly assigned, blocking on schools, to treatment and control conditions in two districts in Oregon.

There are two primary purposes of this presentation. The first is to briefly present impact findings from the first year of the 4-year study. The second focus of the presentation will be to discuss procedural fidelity and variations in the quality of implementation of the intervention. The focus of our discussion will be on the nature of the association between classroom observation data targeting instructional interactions between teachers and students, and student mathematics outcomes.

Focus of Study

Our overall aim is to test the efficacy of the ELM curriculum, and systematically investigate variables that mediate and moderate its impact.

Impact research question. What is the immediate impact of the ELM curriculum on the mathematics achievement of students in general education kindergarten classrooms?

Implementation research question. Specifically, we will examine whether (a) teachers in treatment classrooms provide more and better quality models or demonstrations of the learning outcomes they want their students to attain. We will also examine whether (b) treatment teachers provide more and better quality math feedback to their students. Regarding observed student behaviors in the context of teacher-student interactions during mathematics instruction, we will examine whether the frequency and quality of the opportunities students receive to engage in observable math practice (e.g., student math verbalizations), provided as part of mathematics instruction, is associated with student math achievement outcomes.

Setting

The first year of this 4-year study was conducted in Oregon during the 2008-2009 school year. All kindergarten classrooms in three school districts (urban and suburban) were recruited to participate. A total of 66 classrooms were randomly assigned to treatment and comparison conditions. A total of 64 classrooms were included in data analysis. Classrooms in the treatment condition implemented the ELM curriculum for the entire school year, whereas classrooms in the comparison condition implemented standard district mathematics instruction. The amount of time mathematics instruction was provided each day was the same in treatment and comparison classrooms.

Participants

The Oregon study sample included two participant groups: kindergarten classroom teachers and students in the participating kindergarten classrooms. The teacher sample includes 65 different teachers (i.e., one classroom included two teachers who shared one job). Most teachers were females (97%). The two male teachers involved in the study taught in treatment classrooms. Participating teachers had an average of 10.52 years of teaching experience, and a mean of 6.35 years teaching at the kindergarten level. Across conditions, 56% of the teachers held a graduate degree, and most of the teachers were identified as Caucasian (88%).

The student sample included 1,124 students. Across conditions, 56% of students were eligible for free or reduced lunch, 38% were English learners, and 8% were receiving special education services. Fifty percent of students were White, 36% were Hispanic, 5% were Asian American / Pacific Islander, 2% were African American, and 7% belonged to the Other category.

Intervention

The treatment condition consisted of the ELM curriculum. In these classrooms, the ELM program functions as the core (or Tier 1) mathematics program. The comparison condition consisted of standard district mathematics instruction.

ELM is a full-year kindergarten mathematics curriculum designed for use in whole classroom settings. ELM includes 120 core daily lessons, approximately 45 minutes in duration. Lessons incorporate 4-5 activities across four content strands: (a) whole number and operations; (b) measurement; (c) geometry; and (d) precise mathematics vocabulary. Mathematics content is explicitly introduced in each lesson, and systematically reviewed and extended across lessons.

Practicing ELM teachers implement all aspects of the program, using resources typically available for classroom instruction purposes. Teachers receive three days of professional development training related to program implementation: 1 day at the beginning of the intervention and 2 days during the school year once implementation has begun.

A central feature of the program, established through the design of the curriculum and the professional development that teachers receive to implement the program as intended, is high rates of teacher-student interactions during mathematics instruction. Instructional interactions through explicit instruction and student practice opportunities have been shown to be important for students with and without math difficulties (Gersten et al., 2009). Teachers are expected to model and demonstrate what they want students to learn, and provide specific and frequent feedback to students as they engage in learning activities. Teachers are also expected to provide students with many opportunities to practice learning key mathematics concepts and content. These opportunities can occur in multiple formats but particularly important are practice opportunities that involve verbal interactions between teachers and students. Although many of these modeling, feedback, and practice opportunities for students will occur as a standard part of instruction in comparison classrooms, it is our explicit hypothesis that the quantity and quality of the observable instructional interactions between teachers and students, related to these dimensions, will be more pronounced in the ELM classrooms than comparison classrooms.

Research Design

This study uses a *group-randomized controlled trial* (Murray, 1998, 2001) with students nested within classrooms and classrooms nested within condition. Math achievement data were collected from individual students, and random assignment and instructional delivery took place at the classroom level. Classrooms within schools were matched on full / half day schedules and randomly assigned to treatment (ELM) and control conditions. Thus, in all sites and in both conditions, the same overall amount of mathematics instruction occurred in both treatment and comparison classrooms. The internal validity of our design controls for variables, such as maturation, exposure to mathematics instruction, and clustering, that may limit our ability to infer a causal connection between curriculum exposure and student learning outcomes.

Data Collection and Analysis

Trained staff members collected all student and classroom observation data. Data collection met acceptable reliability criteria for all student and classroom measures included in the analysis.

Student outcomes. Student mathematics performance measures included the Test of Early Mathematics Ability (TEMA) and Early Numeracy – Curriculum Based Measurement (Oral Counting, Number Identification, Missing Number, and Quantity Discrimination). These measures demonstrate strong psychometric properties.

We analyzed intervention effects on each of the primary outcomes with a *nested time by condition analysis* (Murray, 1998) to test differences between conditions on change in outcomes from the beginning of kindergarten (T_1) to the end of kindergarten (T_2). This analysis approach included all data—whether or not a student's scores were present at both time points—to estimate differences between assessment times and between conditions. The nested time by condition analysis accounts autocorrelation among assessments within individual students and the intraclass correlation associated with multiple students nested within same schools. As a test of net differences, it also provides a more straightforward interpretation of the results. We tested moderators within this framework. To test for student risk as a moderator of impact, we added pretest scores and other covariates and their interactions with condition (Jaccard & Turrisi, 2003).

Fidelity of implementation. Fidelity of implementation was measured using the *Classroom Observation Student-Teacher Interactions – Mathematics* (*COSTI-M;* Doabler, Fien, & Smolkowski, 2010), a quantitative measure of classroom instruction. The *COSTI-M* systematically measures the instructional interactions that occur between teachers and students during kindergarten mathematics instruction. Observation data were collected three times during the year by observers using the *COSTI-M*. Adapted from the Classroom Observation Student-Teacher Interactions instrument (Smolkowksi & Gunn, 2011), the *COSTI-M* is grounded in the converging knowledge base of effective mathematics instruction (Gersten et al., 2009; NMAP, 2008). It includes sections for documenting information about the context of the observation and the instructional interactions hypothesized to influence mathematics achievement. The context section documents (a) instructional start and finish times for the math activity, (b) type of content targeted in the activity, and (c) type of instructional format (small-group or whole-class). The mathematical content areas include (a) number and operations, (b) geometry, and (c) measurement.

The instructional interaction section focuses on six behaviors: (a) teacher models, (b) group responses, (c) individual responses, (d) covert responses, (e) student mistakes, and (f) teacher-provided academic feedback. Observers code behavior occurrences in a continual, serial fashion. Consequently, this allows for consideration of the total number of observed behaviors and the rate in which they occurred.

Observers conducted 191 direct observations of classroom instruction across treatment and control conditions using the *COSTI-M*. Direct observations took place across the fall, winter, and spring of the 2008-2009 school year. Observers were required to meet a .80 reliability cutoff before being cleared to conduct observations on their own. Approximately six weeks separated each observation round. Each round planned for one observation per classroom for a total of 65 observations per round. Of the 191 observations, approximately 24% (n = 46) consisted of paired observations or reliability checks. A paired observation entailed two observers independently measuring kindergarten math instruction using the *COSTI-M*. Intraclass correlation coefficients representing inter-rater reliability range from .67 to .95.

We used hierarchical linear models to test the relationship between instructional interactions and student math outcomes. We entered rates of (a) teacher models / demonstrations, (b) student practice opportunities to provide group level responses, and (c) student practice opportunities to provide individual level responses as mediators to determine if they decreased the condition effect, which, if it occurred, would provide support for their role in mediating outcomes.

Findings / Results

Student outcome findings. Analysis of the Oregon data indicates that overall, there was a significant treatment effect. Specifically, students in the treatment condition, ELM, outperformed their control classroom peers on two distal measures of math proficiency. ELM students significantly outperformed control students on both the Test of Early Mathematics Ability – 3rd Edition (TEMA; $t = 2.41$, $p = .02$) and Early Numeracy Curriculum Based Measurement (EN-CBM; $t = 1.99$ $p = .05$). Overall Hedges g effect sizes were .13 on the TEMA and .14 on CBM (Clarke et al., 2011). In condition by risk status analyses, at-risk students (defined as performing below the 40th percentile on the TEMA at pretest) demonstrated the greatest treatment benefit (Clarke et al., 2011). At-risk treatment students significantly outperformed at-risk control students on both the TEMA ($t = 3.29$, $p = .0017$, $g = .24$) and EN-CBM total score ($t = 2.54$, $p = .0138$, $g = .22$). Also, at-risk students in the treatment condition grew more than students who were not at risk for mathematics difficulties, effectively closing the mathematics achievement gap.

COSTI-M findings. After entering rates of (a) teacher models / demonstrations, (b) student practice opportunities to provide group level responses, and (c) student practice opportunities to provide individual level responses as mediators in our HLM model, the condition effect was still significant. Thus, there was no evidence to support this mediation hypothesis.

Next, we explored the interaction between rate of student practice (a combination of items on the COSTI-M measure) and condition. On the TEMA, we found that within ELM condition, high-practice classrooms outperformed low-practice classrooms ($p = .02$) but there was no difference between high-practice and low-practice classrooms in the comparison condition ($p = .76$). We also found that when collapsing across condition, ELM high-practice classrooms trended toward outperforming comparison high-practice classrooms ($p = .07$), but within low practice classrooms there was no difference between ELM and comparison classrooms ($p = .89$). This same general outcome pattern occurred when the outcome measure was the proximal Early Numeracy – Curriculum Based Measurement.

Conclusions

In Oregon, we found there was an overall significant impact on proximal and distal measures and impact was moderated by student risk status, as expected based on the development of the intervention (Clarke et al., 2011). That is, effect sizes were largest for students at risk for mathematics difficulties, and students at risk for mathematics difficulties in the treatment group made greater gains than students who were not at risk.

Preliminary analysis in Oregon has revealed that teacher models and student practice may not mediate impact, but student practice in particular may play a moderating role. There is evidence that the nature and function of practice may be different in ELM classrooms versus comparison classrooms.

There is convincing evidence that many students experience an early and lasting onset of math difficulties. Our findings indicate that getting students on track for success in early mathematics entails establishing a prevention-oriented framework of instruction and implementing carefully designed core (Tier 1) curricula. Independent of the curriculum used for teaching core mathematics instruction in kindergarten, the frequency and quality of observable instructional interactions between teachers and students during mathematics instruction also play an important role in student learning. We expect to further unpack the nature and function of these instructional interactions and determine their relevance for boosting kindergarten mathematics achievement as we delve deeper into our data.

Appendices

Appendix A. References

Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., & Tahan, K. (2011). The Condition of Education 2011. NCES 2011-033. *National Center for Education Statistics*, 410.

Baker, S. K., Chard, D., Clarke, B. S., Smolkowski, K., & Fien, H. Early Learning in Mathematics: Efficacy in Kindergarten Classrooms (ELM).(Oregon: U.S. Department of Education, Institute for Education Science, 2008-2012, $4,280,188).

Clarke, B., Smolkowski, K., Baker, S. K., Fien, H., Doabler, C., & Chard, D. (2011). The impact of a comprehensive tier 1 kindergarten curriculum on the achievement of students at-risk in mathematics. *The Elementary School Journal, 111*(4), 561-584.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Doabler, C., Fien, H., & Smolkowski, K. (2010). Measuring instructional interactions in kindergarten math. Poster presented at the Eighteenth Annual Pacific Coast Research Conference, San Diego, CA

Ginsburg, H., & Baroody, A. J. (2003). *Test of early mathematics ability* (3rd ed.). Austin, TX: Proed.

Griffin, S., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24-49). Cambridge, MA: MIT Press.

Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2 ed.). Thousand Oaks: Sage Publications.

Jordan, N., Kaplan, D., Olah, L., & Locuniak, M. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at-risk for mathematics difficulities. *Child Development, 77*(1), 153-177.

Jordan, N. C., Kaplan, D., & Hanich, L. B. (2002). Achievement growth in children with learning difficulties in mathematics: Findings of a two-year longitudinal study. *Journal of Educational Psychology, 94*(3), 586-597.

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

Murray, D. M. (2001). Statistical models appropriate for designs often used in group-randomized trials. *Statistics in Medicine, 20*(9-10), 1373-1385. doi: 10.1002/sim.675

National Research Council. (2001). Adding it up: Helping children learn mathematics. In J. Kilpatrick, J. Swafford & B. Findell (Eds.), (pp. xvii, 454). Washington, DC: Mathematics Learning Study Committee.

Smolkowski, K., & Gunn, B. (2011). Reliability and validity of the classroom observations of student–teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly*. doi: 10.1016.

Appendix B. Tables and Figures

Not included in page count.