Presentation #2

Title: The Mechanisms behind the Results: Moderators of *Building Blocks* Curricular Effects

Author(s):
Dale Farran, PhD, Vanderbilt University (PRESENTER) (dale.farran@vanderbilt.edu)
Mark Lipsey, PhD, Vanderbilt University (mark.lipsey@vanderbilt.edu)
Douglas H. Clements, PhD, University at Buffalo, The State University of New York (clements@buffalo.edu)
Julie Sarama, PhD, University at Buffalo, The State University of New York (jsarama@buffalo.edu)
Kerry Hofer, PhD, Vanderbilt University (kerry.g.hofer@vanderbilt.edu)
Carol Bilbrey, PhD, Vanderbilt University (carol.bilbrey@vanderbilt.edu)
Elizabeth Vorhaus, Vanderbilt University (elizabeth.vorhause@vanderbilt.edu)

Background / Context:

In intervention research, it is critical to determine not just if an intervention is effective, but *for whom* it is effective and *under what circumstances* those effects occur. Moderators can be the key to answering those questions. A moderator is a variable that affects either the direction or the strength of the relationship between the predictor (curriculum condition, in this case) and the dependent variable (here, child outcomes) (Baron & Kenny, 1986). Identifying those variables that help specify the conditions under which interventions are most effective is central to social science research (Cohen et al. 2003). Moderators of curricular effects may be particularly important to scale-up studies. There may be no more challenging educational and theoretical issue than scaling up educational programs across a large number of diverse populations and contexts in the early childhood system in the U.S. The paucity of high-quality instruments, or the use of any measure of the fidelity of implementation, is one of the most important deficits in the field of scaling up educational innovations (Borman et al., 2003). This paper examines possible moderators, using fidelity as the primary variable of interest, of the effects of an early mathematics curriculum used in a scale-up study across three states.

Purpose / Objective / Research Question / Focus of Study:

The original project scaled up the implementation of a prekindergarten mathematics intervention that had been demonstrated in several randomized trials of increasing scope to increase foundational mathematics skills (Clements & Sarama, 2008). The key question for the scale-up project was whether the curriculum could be effective when provided under circumstances of routine practice on realistic scale—critical if it is to have any potential to help preschools across country improve math instruction. Many early childhood programs developed in universities or specialized research centers have proved initially effective but, when scaled up to be used by a “second-generation” (Farran, 1990, pg. 508), the effects have been diluted or proved non-existent [e.g., the Infant Health and Development Program (Brooks-Gunn, et al., 1994), Even Start (St. Pierre & Swartz, 1996), Head Start (U. S. GAO, 1997), and the Comprehensive Child Development Program (St. Pierre, Layzer, Goodson, & Bernstein, 1997)].

In contrast to that reported dilution, we found significant effects across outcomes in the children who participated in the original scale-up project. However, throughout the course of the study, we noticed that teachers implemented the curriculum to varying degrees, some displaying much greater fidelity of implementation than others. Additionally, we noticed that some children seemed to respond better to the intervention than their peers. This paper addresses those differences through a moderator approach, examining those variables like fidelity and child demographics that might explain some of the relationship between curriculum intervention and students' academic outcomes.

Setting:

This scale-up intervention took place in preschool classrooms in three urban school districts: the Buffalo Public School system in Buffalo, NY, the Boston Public School system in Boston, MA, and a combination of the Metropolitan Nashville Public School system and the Metropolitan Action Council Head Start system in Nashville, TN. A total of 62 sites (26 in Buffalo, 16 in Boston, and 20 in Nashville including 16 public schools and 4 Head Start centers) were randomly assigned to one of two conditions. This process yielded 103 classrooms that participated in the new math curriculum training and 60 classrooms that conducted business as usual. The original study sample included over 2000 children who had at least partial pretest information collected on them, whether by direct assessment, teacher ratings, or observations.

Population / Participants / Subjects:

Participants in this study included primarily at-risk preschoolers between the ages of 4 and 6 from low-income households. The analysis sample, defined as those students who had at least partial pretest information, was comprised of 2076 students. The sample was roughly half male and half female, and predominately African American. Attrition rates were low throughout the study.

Intervention / Program / Practice:

Building Blocks was based on a comprehensive Curriculum Research Framework (Clements, 2007) and its efficacy validated by two Cluster Randomized Trial (CRT) evaluations, yielding effect sizes ranging from .5 to over 2 (Clements & Sarama, 2007, 2008). TRIAD's *Professional Development* includes multiple forms of training (15 full days over two years, the first year a "gentle introduction" with no data collection) and support (coaching and mentoring). Each of these uses the software application, *Building Blocks Learning Trajectories* (*BBLT)*, which presents and connects all components of the innovation. *BBLT* provides scalable access to the learning trajectories via descriptions, videos, and commentaries. The two main aspects of each learning trajectory—the *developmental progressions* of children's thinking and connected *instruction*—are linked to each other. *Building Blocks* is a *supplemental* mathematics curriculum designed to develop preschool children's early mathematical knowledge through various individual and small- and large-group activities. The curriculum embeds mathematical learning in children's daily activities, ranging from designated math activities to circle and story time, with the goal of helping children relate their informal math knowledge to more formal mathematical concepts.

The *Building Blocks* curricular intervention in this scale-up study was implemented during the preschool year after teachers had a year of training and practice. Children from both treatment and control classrooms were followed through their first grade years. Also during the full implementation preschool year, multiple classroom observations were conducted. Observations focused on a myriad of activities in the classroom, including implementation fidelity, teacher strategies, the general classroom environment, the specific math environment, and the behaviors of individual children. This is one of the few longitudinal studies to have detailed descriptions of classroom experiences in the prekindergarten year and to have descriptions that are directly related to the quantity and quality of mathematics instruction the children received.

Research Design:

This scale-up study was a randomized field trial in which schools/centers were randomly assigned to experimental conditions.

Data Collection and Analysis:

Child outcomes were measured with the *Research-based Early Mathematics Assessment* (REMA), which uses an individual interview format, with explicit protocol, coding, and scoring procedures. It assesses children's thinking and learning along research-based developmental progressions within areas of mathematics considered significant for preschoolers, as determined by a consensus of participants in a national conference on early childhood mathematics (Clements & Conference Working Group, 2004), rather than mirroring objectives or activities from any curriculum or state. Topics in number include verbal counting, number recognition and subitizing, object counting and counting strategies, number comparison and sequencing, number composition and decomposition, and adding and subtracting; geometry topics include shape identification, shape composition and decomposition, congruence, construction of shapes, and turns; and finally there are items on measurement and patterning. Content validity was assessed via expert panel review; concurrent validity was established with a .86 correlation with another instrument (Klein, Starkey, & Wakeley, 2000). For this study, Rasch scores for the total instrument were computed on correctness scores and logits transformed to T-scores ($M = 50$, $SD = 10$) for ease of interpretation. These T-scores were used for all statistical analyses. The REMA was given to children at the beginning and end of preschool, the end of kindergarten, and the end of first grade. In Nashville, children's math and literacy achievement were also assessed using two math subtests and one literacy subtest from the Woodcock Johnson III. Each of these tests was given to children twice during the preschool year, once at the end of the kindergarten year, and again at the end of the first grade year. Demographic information such as age, gender, ethnicity, IEP status, and ELL status were also collected.

Teacher implementation fidelity was measured with two instruments, one completed in both treatment and control classrooms, and the other completed only in treatment classrooms. The *Classroom Observation of Early Mathematics-Environment and Teaching* (*COEMET)* was used by observers in all participating classrooms. The COEMET is an instrument that measures the quality of the mathematics environment and activities, not connected to any curriculum. Thus, it allows for treatment-control group contrasts, no matter what the source of the enacted curriculum. Observers completed the *COEMET* during three separate observations across the full implementation year for all teachers. The *Near Fidelity* measure was used by observers in the treatment classrooms only. This measure was focused on the general quality of the mathematics environment in the classroom, as well as detailed information concerning specific math activities. This instrument evaluated the degree to which teachers were teaching the specific intervention curriculum. This data from the COEMET and the *Near Fidelity* were collected from three four-hour observations in classrooms during the preschool implementation year: fall, winter, and spring.

To examine the possible existence of moderator variables that altered the effectiveness of the curriculum in enhancing children's mathematics skills, a series of linear mixed models was conducted. Independent models predicted children's skills on each of the assessments at the end of prekindergarten, the end of kindergarten, and the end of first grade from the main effect of curriculum condition, the main effect of each moderator, and the interaction of each moderator with curriculum condition. In analyses of the *Near Fidelity* measure, condition was not included. Children were nested in their prekindergarten classrooms, schools, and sites. Due to the nested nature of the design, the effective sample size for analyses is decreased by a factor related to the Intraclass Correlation Coefficient (ICC), or the degree to which classroom and school units are non-independent. Because of this reduction in analytical sample size, a p-value of .10 was held as the significance marker rather than the more conservative .05.

Findings / Results:

Teachers in the *Building Blocks* classrooms had higher COEMET scores across all time points than did control teachers. On a 5-point scale with 5.0 being the highest implementation score, treatment teachers scored an average of 3.84 across time while control teachers scored an average of 3.18. Regarding the number of math activities, teachers in *Building Blocks* classrooms were observed with an average of 2.4 structured, planned math activities and 9.3 less formal, miniature math activities per day, while teachers in control classrooms were observed with an average of 1.5 structured math activities and 5.7 miniature math activities per day. Additionally, the quality with which math activities were taught and managed was higher for treatment teachers than control teachers (3.44 and 3.09, respectively, on a 1-5 scale). Though there was a range of implementation observed in the treatment classrooms, teachers on average implemented over half of the curriculum's suggestions in each instructional component (mean scores: General Curriculum: 12 out of 17; Hands-On Centers: 13 out of 22; Whole Group Activities: 26 out of 35; Small Group Activities: 54 out of 93; Computer Activities: 32 out of 49). While the differences in mathematics activities between treatment and control teachers were both significant and meaningful, the relatively low scores for the treatment teachers is evidence of the difficulty of changing the mathematics environment in early childhood classrooms.

Preliminary results regarding fidelity as a moderator of curriculum effects indicate that teachers with higher implementation scores had students who made greater gains on their mathematics measures (See Figures 1 and 2). Additional analyses are being conducted to examine the moderator effect of treatment-only fidelity and child demographics.

Conclusions:

Classrooms are complicated environments; educational interventions being scaled up rely on being able to change fundamental aspects of that environment, including teacher instructional behaviors and child learning behaviors. Data from this study demonstrated teacher adoption of many aspects of the curriculum, and the degree of adoption was related modestly to the amount children gained in math knowledge over the preschool year. While teachers enacted more math instructional activities, changing the quality of their instruction was much more difficult. Moreover, the overall instructional classroom environment in treatment and control classrooms remained very similar, as did children's learning behaviors. TRIAD is a multi-dimensional intervention that includes extensive professional development and in-classroom mentoring. The results from this scale up investigation demonstrate how difficult immediate and sustained change is to effect.

Appendix A. References

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73* (2), 125-230.

Brooks-Gunn, J., McCarton, C. M., Casey, P. H., McCormick, M. C., Bauer, C. R., Bernbaum, J. C., et al. (1994). Early intervention in low-birthweight premature infants: Results through age 5 years from the Infant Health and Development Program. *Journal of the American Medical Association, 272(*16).

Clements, D. H., & Conference Working Group. (2004). Part one: Major themes and recommendations. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 1–72). Mahwah, NJ: Lawrence Erlbaum Associates.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the *Building Blocks* project. *Journal for Research in Mathematics Education, 38,* 136-163.

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*, 443-494.

Clements, D. H. (2007). Curriculum research: Toward a framework for “research-based curricula.” *Journal for Research in Mathematics Education, 38*, 35-70.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Farran, D.C. (1990). Effects of intervention with disadvantaged and disabled children: A decade review. In S. Meisels & J. Shonkoff (Eds.), *Handbook of early intervention*. Cambridge, England: Cambridge University Press.

Government Accountability Office. (1997). *Head Start: Research Provides Little Information on Impact of Current Program* (HEHS-97-59). Washington, DC: Government Printing Office.

Klein, A., Starkey, P., &Wakeley, A. (2000). *Child Math Assessment: Preschool Battery (CMA)*. Berkeley: University of California, Berkeley.

St. Pierre, R. G., & Swartz, J. P. (July 1996*). The Even Start Family Literacy Program: Early Implementation.* Cambridge, MA: Abt Associates Inc.

St.Pierre, R., Layzer, J., Goodson, B., & Bernstein, L. (June 1997). *National impact evaluation of the Comprehensive Child Development Program: Final report*. Cambridge, MA: Abt Associates Inc.

Appendix B. Tables and Figures

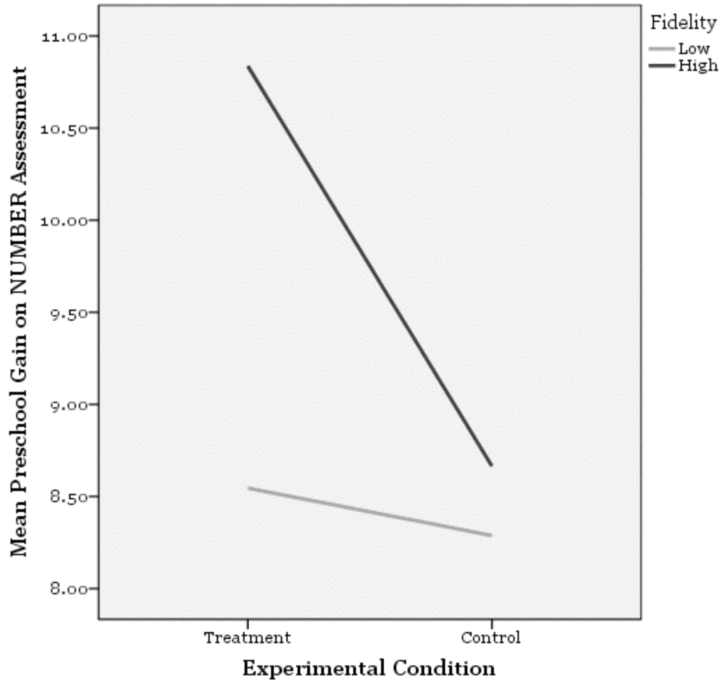

Figure 1. Mean Preschool NUMBER Student Gain by Curriculum/COEMET Grouping

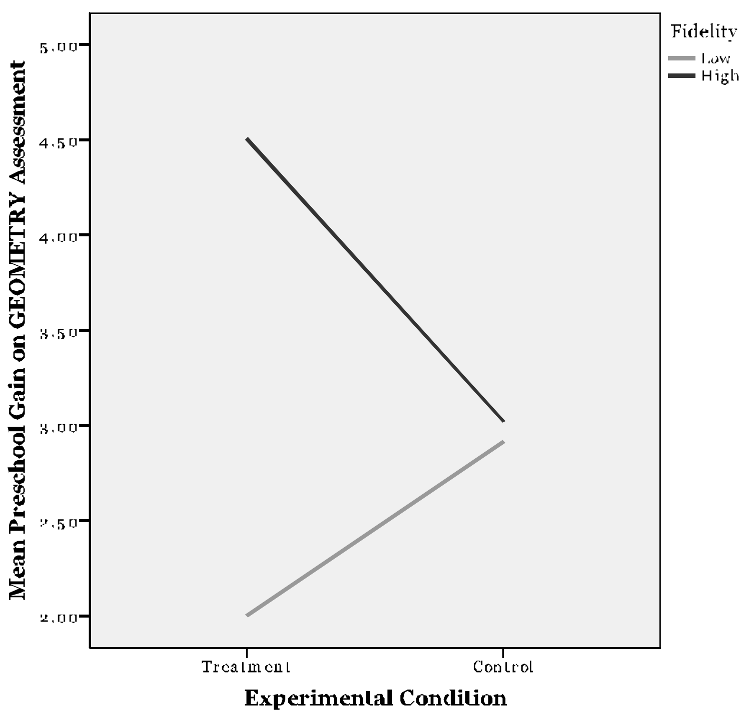

Figure 2. Mean Preschool GEOMETRY Student Gain by Curriculum/COEMET Grouping