Presentation #1

Title: An Examination of the Building Blocks Math Curriculum: Results of a Longitudinal Scale-Up Study

Authors:

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

Background / Context:

Studies show that the mathematics test-score gap is evident at every level of schooling and can be linked to students' earlier performance. For example, a mathematics performance gap was found in children as young as three years of age (Case & Griffin, 1990; Jordan, Huttenlocher, & Levine, 1992). This gap has effects into kindergarten and 1st grade (Denton & West, 2002; Entwisle & Alexander, 1993; West, Denton, & Reaney, 2001) that continue into middle school and high school (Berkner & Chavez, 1997; Brasswell, Lutkus, Grigg, Santapau, Tay-Lim, & Johnson, 2001). However, the gap when children are young is relatively smaller than the gap at older grades (Bodovski & Farkas, 2007; Kilpatrick, Swafford, & Findell, 2001). Thus, children who come to school with less knowledge about mathematics are most at risk of falling behind in elementary mathematics, which affects their overall math achievement across the K-12 years. Addressing the mathematics performance gap early on, before children start school, has therefore become a priority for preschool programs serving children from low-income backgrounds (Clements, 2004).

Although the successes of some research-based educational practices have been documented, equally recognized is the "deep, systemic incapacity of U.S. schools, and the practitioners who work in them, to develop, incorporate, and extend new ideas about teaching and learning in anything but a small fraction of schools and classrooms" (see also Berends, Kirby, Naftel, & McKelvey, 2001; Cuban, 2001; Elmore, 1996, p. 1; Tyack & Tobin, 1992). 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., avoiding the dilution and pollution that usually plagues such efforts to achieve broad success. We created a research-based model to meet this challenge in the area of mathematics, with the intent that the model generalize to other subject matter areas and other age groups. The field needs transferable, practical examples of scale up (McDonald, Keesler, Kauffman, & Schneider, 2006); empirical evidence of the effectiveness of these examples; and focused research on critical variables—all leading to refined, generalizable theories and models of scale up. This paper examines the longitudinal effects of a randomized field trial involving the scale-up of a preschool mathematics curriculum.

Purpose / Objective / Research Question / Focus of Study:

The specific goal of our implementation of the TRIAD (Technology-enhanced, Research-based, Instruction, Assessment, and professional Development) model was to increase math achievement in young children, especially those at risk, by means of a high-quality field-centered implementation of the Building Blocks math curriculum, with all aspects of the curriculum—mathematical content, pedagogy, teacher's guide, technology, and assessments—based on a common core of learning trajectories. The TRIAD intervention provides (a) these curriculum materials; (b) ongoing professional development, including scalable distance education, a web-based application with extensive support for teaching based on learning trajectories, and classroom-based coaching during the school year; and (c) supportive roles and materials for parents and administrators. In this project, we evaluated a large-scale implementation of the TRIAD intervention in distant geographical areas with diverse populations. The primary research question of interest is as follows:

Do children who are exposed to the Building Blocks mathematics curriculum in preschool perform better on measures of mathematics skills through the end of first grade than do children who were not exposed to that curriculum?

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 2,100 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, the I in the TRIAD acronym, 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). The Assessment component of TRIAD includes both formative assessment performed by the teachers training to use learning trajectories for this purpose, supplemented by the Building Blocks Software management system. 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 the 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.

Research Design:

This scale-up study was a cluster 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, as was the Renfrew Bus Story, a measure of children's narrative recall skills (although only administered in the preschool year). 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. In order to examine the effectiveness of the curriculum in enhancing children's mathematics skills, a series of linear mixed models were 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 curriculum condition, controlling for children's pretest skills and a host of demographic covariates. 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:

HLM analyses revealed that the two experimental groups differed significantly in math achievement in preschool. The Building Blocks group outperformed the control group on REMA measures with effect sizes ranging from .35 to .69. The Nashville site found effect sizes for Woodcock Johnson math subtests that ranged from .18 to .32. At no site were there significant effects on children's letter/word identification. However, significant effects were found for expressive language in some sites but not others. While some significant curricular effects were found at the end of kindergarten, most differences between experimental groups had dissipated by the end of first grade.

Conclusions:

Our 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, including different measures of mathematics achievement and one expressive language measure.

Research has suggested that early curricular effects may fade over time, resulting in very little, if any, discernable difference in elementary school between students who had been exposed to a given curriculum prior to formal schooling and students who were not exposed to such a program, as those without early curriculum exposure "catch up" to their peers (Barnett et al., 1995). In the Preschool Curriculum Evaluation Research (PCER) project, across all 14 curricula, kindergarten effects were nonexistent, prompting a decision not to collect any further longitudinal data. Similarly, with the scale-up project, we saw evidence of curricular effects across outcomes at the end of prekindergarten, but very few differences at the end of kindergarten, and virtually none at the end of first grade. Longitudinal research, including follow through interventions in these grades, is needed to determine if these early gains truly "fade," or if, as we posit, the problem is that primary grade curricula and teachers do not build upon them.

Appendix A. References

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5*(3), 25-50.

Berends, M., Kirby, S. N., Naftel, S., & McKelvey, C. (2001). *Implementation and performance in New American Schools: Three years into scale-up*. Santa Monica, CA: Rand Education.

Berkner, L., and Chavez, L. (1997) Access to Postsecondary Education for the 1992 High School Graduates (NCES 98-105). U.S. Department of Education. Washington DC: U.S. Government Printing Office.

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal, 108*(2), 115-130.

Brasswell, J. S., Lutkus, A. D., Grigg, W. S., Santapau, S. L., Tay-Lim, B., & Johnson, M. (2001). *The nation's report card: Mathematics 2000* (NCES 2001–517). Washington, DC: National Center for Education Statistics.

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).

Case, R., & Griffin, S. (1990). Child cognitive development: The role of central conceptual structures in the development of scientific and social thoughts. In C. A. Hauert (Ed.), *Advances in psychology-developmental psychology: Cognitive, perception-motor and neurological perspectives*. Amsterdam: North Holland.

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. (2004). Geometric and spatial thinking in early childhood education. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 267–297). Mahwah, NJ: Lawrence Erlbaum Associates.

Cuban, L. (2001). *Oversold and underused*. Cambridge, MA: Harvard University Press.

Denton, K., & West, J. (2002, March 20). *Children's reading and mathematics achievement in kindergarten and first grade* (NCES 2002125). Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002125

Elmore, R. F. (1996). Getting to scale with good educational practices. *Harvard Educational Review, 66*, 1-25.

Entwisle, D. R. & Alexander, K. L. (1993). Entry into school: The beginning school transition and educational stratification in the United States. *Annual Review of Sociology, 19*, 401-423.

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.

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28*, 644–653.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

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

McDonald, S.-K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, *35*(3), 15-24.

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.

Tyack, D., & Tobin, W. (1992). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal*, *31*, 453-479.

West, J., Denton, K., & Reaney, L. (2001). *The kindergarten year: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998–1999*. US Department of Education, National Center for Education Statistics (NCES 2001-023). Washington, DC: US Government Printing Office.