Abstract Title Page

Title:

Achievement Effects of Four Early Elementary School Math Curricula: Findings from First Graders in 39 Schools

Authors:

Roberto Agodini, Ph.D., Mathematica Policy Research, Inc.
Barbara Harris, Mathematica Policy Research, Inc.

Abstract Body

Background/context:

Math skills are critical for success in the workplace, more so today than was the case years ago. Scientific jobs have always required a strong math foundation, and growth rates in science- and technology-related jobs are exceeding job growth in the general labor force (National Science Board 2008). In addition, service jobs and jobs that once relied on strength and endurance now also require math skills for workers to perform successfully. For example, yesterday's assembly-line workers had to be physically fit and skillful with their hands. Today's assembly-line workers must possess math skills to effectively operate computerized equipment that automates tasks performed manually in the past.

Federal legislation recognizes the importance of beginning to develop math skills at an early age, and under Title I of the No Child Left Behind Act (NCLB), public schools must make adequate yearly progress (AYP) in bringing their students to state-specific targets for proficiency in math and reading. The intent of this provision is to ensure that all students are proficient in math and reading by 2014. However, at present, many students show mastery of only rudimentary mathematics, and only a small proportion achieve at high levels (Lee et al. 2007). In the 2007 National Assessment of Educational Progress, only 39 percent of fourth graders were judged "proficient" in mathematics, and 18 percent scored below "basic." Given the NCLB goal of ensuring that all students are proficient in math and reading by 2014, many schools have a lot of work ahead of them before all their students achieve proficiency in math.

What is taught to students and how it is taught (that is, curriculum and its pedagogical approach) may be important factors in a school's ability to improve student math achievement. Therefore, one way to help schools meet their AYP goals is to help them choose a math curriculum with a pedagogical approach that works best with their students. At present, a small number of curricula dominate elementary math instruction—in particular, seven curricula make up 91 percent of the curricula used by K-2 educators (Education Market Research 2008). The curricula are based on different theories for developing math skills, and debate exists over which theory is best.

Many of the curricula used in elementary schools are categorized as "reform" (student-centered) or "traditional" (teacher-directed). The former were developed with support from the National Science Foundation (NSF), and the latter were commercially developed. Reform approaches are based on the premise that children will develop a strong understanding of mathematical principles if learning takes place in the context of solving real-life problems through student-centered activities. Traditional approaches are based on the premise that children will develop a strong understanding of mathematical principles by first being taught facts and procedures, and then applying that knowledge to real-life problems. The debate about which is better is sometimes so intense that it is referred to as the "math war" (Whitehurst 2003; Schoenfeld 2004; and Klein 2007).

Despite the debate over the varying approaches to math instruction, there is little research-based evidence to support one theory over another (National Mathematics Advisory Panel 2008;

What Works Clearinghouse 2006; National Research Council 2004). The ongoing controversy over the different instructional approaches and lack of evidence on their relative effectiveness prompted the U.S. Department of Education (ED) to sponsor an evaluation of math curricula.

Purpose/objective/research question/focus of study:

The purpose of this large-scale, multi-year, national study is to determine whether some early elementary school math curricula are more effective than others at improving student math achievement, thereby providing educators with information that may be useful for making AYP. The main questions being addressed by the study are:

- What are the relative effects of different early elementary school math curricula on student math achievement in disadvantaged schools?
- Under what conditions is each math curriculum most effective?
- What is the relationship between teacher knowledge of math content/pedagogy and the effectiveness of the curricula?
- Which math curricula result in a sustained impact on student achievement?

The study is addressing these questions by examining the relative effects of four diverse curricula, by recruiting schools without prior experience using the study's curricula,[*] having publishers provide schools and teachers with curriculum materials and training, and evaluating the effects of the curricula (including materials, training, content, and pedagogy) on student achievement.

Setting:

The first year of the study included four geographically dispersed districts that were located in three regions of the country (Northeast, Midwest, and West). The districts also fall in areas with different levels of urbanicity—two districts are in urban areas, one is in a suburban area, and the other is in a rural area.

Population/Participants/Subjects:

The study team recruited districts for the evaluation that (1) have Title I schools, (2) are geographically dispersed, and (3) contain at least four elementary schools interested in participating so all four curricula could be implemented in each district. In the 2006-07 school year, four districts and 39 schools participated in the study. Curriculum implementation occurred in the first grade; data were collected from 131 teachers and from 1,309 students—a random sample of about 10 students in each classroom.[†] Each of the four curricula was assigned about 10 schools with 33 classrooms and 325 students.

[*] New users of the curricula were recruited so that no one curriculum has an advantage over the other.

[†] Given the number of schools and classrooms included in the study, the statistical power benefits of pre- and post-testing more than 10 students per classroom are minimal, though the costs are significant because the study used an individually administered assessment, as described below.

When compared to the average U.S. district, those that agreed to participate have a higher fraction of school-wide Title I eligible schools, students eligible for free/reduced-price meals, and minority students. A similar pattern exists when comparing U.S. elementary schools with those that agreed to participate. Across the four curriculum groups, students and teachers had similar characteristics at baseline.

Intervention/Program/Practice:

Four curricula were selected for inclusion in the study using a competitive process in which developers and publishers of early elementary school math curricula were invited to submit a proposal to include their curricula in the evaluation. A panel of experts in math and math instruction reviewed the proposals and recommended to ED curricula suitable for the study. The following four curricula were selected for the evaluation to represent many of the diverse approaches used to teach elementary school math:

- ***Investigations in Number, Data, and Space,*** published by Pearson Scott Foresman, uses a student-centered approach encouraging metacognitive reasoning and drawing on constructivist learning theory. The lessons focus on understanding, rather than on "correct answers," and build on students' knowledge and understanding. Students first investigate, then discuss and reason about problems and strategies.

- ***Math Expressions***, published by Houghton Mifflin Company, blends student-centered and teacher-directed approaches to mathematics. Students question and discuss mathematics, but are explicitly taught effective procedures. The curriculum emphasizes using multiple objects, drawings, and language to represent concepts, and learning through the use of real-world situations.

- ***Saxon Math***, published by Harcourt Achieve, is a scripted curriculum that blends teacher-directed instruction of new material with daily distributed practice of previously learned concepts and procedures. Teachers introduce concepts and efficient strategies for solving problems. Students hear the correct answers and are explicitly taught procedures and strategies.

- ***Scott Foresman-Addison Wesley Mathematics,*** published by Pearson Scott Foresman, is a basal curriculum that combines teacher-directed instruction with a variety of differentiated materials and instructional strategies. The curriculum is based on a consistent daily lesson structure which includes direct instruction, hands-on exploration, the use of questioning, and practice of new skills.

The curricula were implemented in study schools during the 2006-07 school year. Publishers provided curriculum materials and training to teachers in summer 2006; additional follow-up training was provided by the publishers to study teachers during the school year.

Research Design:

The study is based on a school-level experimental (random assignment) design. In particular, the participating elementary schools in each district were randomly assigned to the four curricula. For example, in a district that has eight participating elementary schools, the study

randomly selected two schools to implement curriculum A, two schools to implement curriculum B, and so on. Within each participating school, all first-grade teachers participated in the study.

Because students must take math, the study does not include a control group with students who do not receive math instruction. The study team also chose not to include a 'business-as-usual' control group that uses the curriculum previously used by a school, because it would be difficult to interpret results based on such a control group that would contain numerous curricula. Before joining the study, the participating districts used a variety of curricula, and in some of those districts, individual schools could choose their math curriculum. Rather than include a control group with a variety of curricula, the study chose to compare the effects of four curricula that (as described above) represent many of the diverse approaches to teaching mathematics.

The relative effects of the curricula were estimated using hierarchical linear modeling (HLM) techniques which compare the average math achievement of students in schools assigned to the various curricula. For example, the relative effect of curriculum A versus curriculum B is estimated as the difference in average achievement between students in the schools assigned to curriculum A and students in the schools assigned to curriculum B.[‡]

Statistical tests were used to assess the significance of all results. When comparing results for pairs of curricula, the Tukey-Kramer method was used to adjust the statistical tests for the six unique pair-wise comparisons that can be made with four curricula (Tukey 1952, 1953; Kramer 1956). HLM techniques were used to conduct the statistical tests. Only results that are statistically significant at the 5 percent level of confidence will be discussed.

Data Collection and Analysis:

The study team collected all data necessary to evaluate the four curricula. To measure the effects of the curricula, the study team tested students at the beginning and end of the school year using the math assessment developed for the Early Childhood Longitudinal Study. The assessment met the study's requirements regarding validity, reliability, math content areas covered, individual administration, and accuracy in capturing achievement of students from a wide range of backgrounds and ability levels (Rock and Pollack 2002). The assessment also is adaptive, which limits the amount of time students are away from their classrooms and supports an accurate assessment of achievement gains over the study's grade range.

To help interpret measured effects, the study team conducted several complementary data collection efforts:

- ***Assessment of Teacher Knowledge of Math Content and Pedagogy.*** Teachers' math content and pedagogical knowledge was assessed at the initial teacher training sessions before the curricula were introduced, using an assessment developed by researchers at the University of Michigan (Hill 2004). The study team included the test scores in the analysis to examine the relationship between teachers' knowledge and the relative effects of the curricula.

[‡] The HLM includes school, classroom, and student characteristics (such as baseline math achievement) to increase the precision of the results.

- ***Classroom Observations.*** Each classroom in the study was observed once during the school year by the study team. Observers collected information on the frequency with which teachers used various instructional practices during math instruction. This information will be used to assess the fidelity of implementation and to compare instructional practices across the four curricula.

- ***Teacher Surveys.*** Two surveys were administered to teachers. The first, administered in the fall, focused on teacher background information, classroom characteristics, curriculum training provided by the publishers up to that point, and math instruction approaches used before the study. These data were used as control variables in the HLM and to describe the participating teachers. The second survey, administered in the spring, focused on follow-up training provided by the publishers, use of the assigned curriculum, and math instructional practice during the school year. Data from that survey were used to assess the fidelity of implementation, as reported by teachers.

Findings/Results:

The results that will be presented during the session are based on the first cohort of 39 schools that began study participation during the 2006-07 school year, in which about 10 schools with 33 classrooms and 325 students were randomly assigned to each of the four curricula.[§] The results include the relative effects of the curricula on average student achievement, the relative effects of the curricula for a variety of subgroups, and an assessment of teacher adherence to their assigned curriculum.

Conclusions:

As math skills have grown in importance in the workplace, so has the debate over how best to improve students' math skills. This study is an important step toward resolving that debate. The study's results present rigorous evidence of the relative effectiveness of four commercially available math curricula that are widely used in the nation's classrooms. The study's first report (upon which the presentation will be based) will provide evidence on the relative effectiveness of the curricula in first grade, which may help educators make more informed decisions about the math curriculum (and pedagogical approach) that will work best with their students. Future study reports will provide information based on second- and third-grade classrooms.

[§] The effect size that can be detected with the first cohort is as small as 0.22, where the effect size equals the difference in average student math scores of any two curriculum groups, divided by the pooled standard deviation of the score for the two curricula being compared.

Appendix A. References

Education Market Research. (2008). *Mathematics Market, K-12, 2008.* Rockaway Park, NY: EMR.

Hill, Heather. (2004). *Learning mathematics for teaching & study of instructional improvement:* Content *knowledge for teaching mathematics measures, introduction to CKTM forms.* Ann Arbor, MI: University of Michigan.

Klein, David. (2007). A quarter century of U.S. 'math wars' and political partisanship. *Journal of the British Society for the History of Mathematics, 22,* 22-33.

Kramer C.Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrica, 12,* 307-310.

Lee, J., W. Grigg, & G. Dion. (2007). *The nation's report card: mathematics 2007.* Publication NCES 2007-494. Washington, D.C: U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.

National Mathematics Advisory Panel. (2008). *Foundations for success: the final report of the national mathematics advisory panel.* Washington, DC: U.S. Department of Education.

National Research Council. (2004). *On evaluating curricular effectiveness: judging the quality of K-12 mathematics evaluations.* Washington, DC: National Academies Press.

National Science Board. (2008). *Science and engineering indicators 2008 (Vol. 1 & Vol. 2).* Arlington, VA: National Science Foundation.

Rock, Donald A., & Judith M. Pollack. (2002). *Early childhood longitudinal study—kindergarten class of 1998-99: Psychometric report for kindergarten though first grade.* Publication NCES 2002-05. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Schoenfeld, Alan H. (2004). The math wars. *Educational Policy, 18,* 253–286.

Tukey, J.W. (1952). *The problem of multiple comparisons: allowances for various types of error rates.* Unpublished IMS address, Chicago, IL.

Tukey, J.W. (1953). *The problem of multiple comparisons.* Unpublished manuscript.

What Works Clearinghouse. (2007). *Elementary School Math.* Retrieved February 7, 2007, from www.whatworks.ed.gov/Topic.asp?tid=04&ReturnPage=default.asp.

Whitehurst, G. J. (2004). *Research on Mathematics Instruction.* Retrieved February 7, 2007, from http://ies.ed.gov/director/speeches2003/02_06/2003_02_06.asp.