Abstract Title Page

Title: Preliminary Findings from a Multi-Year Scale-Up Effectiveness Trial of Everyday Mathematics

Authors and Affiliations: Michael Vaden-Kiernan (SEDL), Geoffrey Borman (University of Wisconsin-Madison), Sarah Caverly (SEDL), Nance Bell (SEDL), Veronica Ruiz de Castilla (SEDL), and Kate Sullivan (SEDL).

Abstract Body

Background / Context:

This Goal 4 study addresses the effectiveness of *Everyday Mathematics*, a widely used core mathematics curriculum that reflects over two decades of National Science Foundation (NSF)-sponsored research and development studies (Klein, 2007; National Research Council, 2004) and aligns well with recommended policies and practices by the National Council of Teachers of Mathematics (NCTM) Curriculum Focal Points (2006) and National Mathematics Advisory Panel (NMAP) (2008). This and other similar curricula are increasingly needed to strengthen student math skills and ensure that all children are at or above grade level in math proficiency. Statistics present concerns to reaching this goal - only 42% of fourth grade students and 35% of eighth grade students are at or above proficient-level on National Assessment of Educational Progress (NAEP) math scores (National Center for Educational Statistics, 2013). Recent initiatives emphasize the role of early math instruction and curricula, recognizing that students need to develop a well-defined set of skills in critical math content areas (e.g., whole numbers, fractions, elements of geometry, and measurement) in early grades in order to prevent difficulties in later grades. The NMAP (2008) recommended that all students receive effective preparation from an early age to ensure their later success in algebra and emphasized the need for early math programs that mitigate and prevent difficulties. To date, however, *Everyday Mathematics* is only one of a few elementary school math curriculum reviewed by the What Works Clearinghouse (WWC) (2007) that has demonstrated "potentially positive effects," with the evidence for effectiveness on math achievement rated as medium to large[1].

The *Everyday Mathematics* (EM) curriculum, developed by the University of Chicago School Mathematics Project (UCSMP) during the mid-1980s, and published by Wright Group/McGraw-Hill, relies on research-based practices, which according to the National Research Council (2004), are supported by more researchers and empirical studies than any other elementary mathematics curriculum. EM is a PK-6 curriculum emphasizing six content strands (numbers and numeration; operations and computation; data and chance; measurement and reference frames; geometry; and patterns, functions, and algebra) with learning targets or curriculum focal points identified for each of the six strands at each grade level. Developed largely from NSF-sponsored studies, the curriculum shows promise in preventing math difficulties in early grades. It emphasizes a constructivist philosophy, with a strong emphasis on real-life problem solving, manipulatives, concept development, and targeted use of technology and parent participation. Besides being the highest rated elementary school math curriculum in terms of effectiveness by the WWC (2007), it has also been widely adopted and used across the country. EM is used by nearly 4 million students in more than 11,000 schools in more than 3,000 districts in all 50 states (SRA/McGraw-Hill, 2009). Results from quasi-experimental evaluations have revealed statistically significant math achievement advantages of +0.16 relative to other math curricula (Carroll, 1998; Carroll & Isaacs, 2003; Riordan & Noyce, 2001; SRA/McGraw-Hill, 2003;

[1] "Potentially positive effects" is evidence of a positive effect in a domain with no overriding contrary evidence. Programs rated as having a "medium to large" amount of evidence require at least two studies that meet the WWC evidence screen with 2 schools and a total sample size of at least 350 students or 14 classrooms across the studies. The screened studies reviewed for EM included a total of approximately 12,600 students in grades 3-5 from a range of socioeconomic backgrounds and attending schools in urban, suburban, and rural communities in multiple states.

Waite, 2000; WWC, 2009; Woodward & Baxter, 1997). In addition, a cluster randomized controlled trial (RCT) currently being conducted by the publisher is documenting potential impacts of EM on math achievement in grades K-5 students in three schools in Washington D.C. The study is assessing classroom-level impacts of treatment assignment on the TerraNova math achievement test scores (SRA/McGraw-Hill, 2009b). Despite the curriculum's widespread use and promising research findings, EM has not been evaluated rigorously on a large scale as part of an objective, third-party evaluation.

Purpose / Objective / Research Question / Focus of Study:

Given the importance of early mathematics instruction and curricula for preventing mathematics difficulties in later grades, it is necessary to identify effective mathematics curricula and instruction to ensure that children become proficient in early mathematics content and procedures. *Everyday Mathematics*, was reviewed by the What Works Clearinghouse and is reported to have "potentially positive effects" on students' mathematics achievement. However, most of the studies that have evaluated EM have used quasi-experimental designs or are small-scale randomized control trials. This study reports the preliminary year one findings for Kindergarten and 3rd grade cohorts of the first scale-up evaluation of this widely used curriculum. The results of this study will contribute to understanding whether EM is effective in promoting mathematic proficiency in the elementary grades when implemented "at scale" with typical "real world" levels of support. The study was designed to address the following research questions:

- **Overall Impacts**. Does school-level assignment to the *Everyday Mathematics* curriculum intervention produce stronger effects on math achievement than assignment to the "business-as-usual" control condition?
- **Impacts by Subgroups**. Is there significant variation in the outcomes of *Everyday Mathematics* or do the effects reliably replicate across student subgroups, the sampled classrooms/teachers, schools, and districts?
- **Fidelity of Implementation**. To what extent was the intervention delivered as the curriculum developers indicated it should be implemented? Was there significant variation in implementation fidelity of *Everyday Mathematics* among the classrooms/teachers, schools, and districts? In what ways were *Everyday Mathematics* students' experiences similar or different to those of students in the control condition?
- **Proximal Outcomes as Moderators of Impacts**. Is there a significant relationship between proximal student and teacher outcomes, such as fidelity of implementation or student motivation/engagement, and student math achievement outcomes and does this relationship vary by classrooms/teachers, schools, and districts?

Setting:

The study was conducted in a sample of 48 elementary schools (kindergarten through 5th grade) in 7 districts across the country.

Population / Participants / Subjects:

The study participants include approximately 4,500 elementary school students and 1,200 teachers per year.

Intervention / Program / Practice:

EM is a core mathematics curriculum for grades prekindergarten to six. The curriculum emphasizes six strands of mathematics knowledge with learning targets identified for each strand by grade level. It includes student materials, teacher manuals, assessment and practice guides, and home links to support parent involvement. The curriculum includes a 2- to 3-day summer workshop to train teachers to implement the curriculum, as well as follow-up support by EM consultants. The instructional format follows a consistent three-part lesson plan in all grades focusing on teaching the lesson, ongoing learning and practices for students, and differentiated instruction options. Teachers, as part of the intervention, can also use informal and formal assessments to monitor student progress and inform instruction.

Research Design:

The study design is a multi-site cluster-randomized trial in which 48 elementary schools from 7 districts across the country were randomized to training and delivery of the EM curriculum (treatment group) or to delivery of the standard mathematics instruction for the school (control group) blocking at the district level.

Data Collection and Analysis:

Data from teachers and students in two cohorts (grades K&3 and grades 1&4) were gathered over two school years. This paper is presenting findings from the first year. The pre- and post-test outcomes were assessed in the fall and spring using the *Group Mathematics Assessment and Diagnostic Evaluation* (GMADE) and Student Motivation Form (SMF). Fidelity of implementation was captured using classroom observations, interviews, and surveys with teachers and other key staff (e.g., curriculum trainers).

The main intent-to-treat impact analyses uses a three-level model with school-level fall pre-test scores on the GMADE as a covariate and spring post-test scores as the dependent variable, nested within schools, which in turn were nested within districts. Additional subgroup (moderator) analyses were used to investigate the effects of the EM program as a function of student baseline characteristics (e.g., age/grade, gender, baseline math proficiency, student engagement), teacher/classroom characteristics (e.g., class size, fidelity of implementation), and school characteristics (e.g., geographic region or locale). Student, teacher, and school characteristics will also be examined as potential mediators of the effects.

Findings / Results:

Table 1 provides school level characteristics of students in the study schools at baseline. The analytical sample is comprised of 4,520 students in 48 study schools in grades K and 3 with valid scores on the Spring GMADE assessment. Results indicate that at baseline, there were no significant differences between the demographic characteristics of schools in the treatment and control conditions. In terms of math achievement at the beginning of the year, there was no statistically significant difference between the mean pre-treatment score of EM schools (86.80, s.d.=6.72) and that of control schools (87.46, s.d.=6.26).

Insert Table 1 here

Preliminary results from the overall ITT impact analyses are presented in Table 2. The intraclass correlation (ICC) of student math achievement was .081 for schools and .087 for districts. The three level multilevel model (student, school, and district) includes grand-centered, school-mean GMADE pretest scores and an indicator for treatment condition. Both predictor variables of pretest and treatment condition are included in level 2 of the model.

Level 1:
$$Y_{ijkl} = \beta_{0jk} + \varepsilon_{ijkl}$$
Level 2:
$$\beta_{0jk} = \gamma_{00j} + \gamma_{01j}(MeanPretest)_{jk} + \gamma_{02}(Treatment)_{jk} + \eta_{0jk}$$
Level 3:
$$\gamma_{00} = \pi_{000} + \zeta_{000k}$$
$$\gamma_{01} = \pi_{010} + \zeta_{010k}$$
$$\gamma_{02} = \pi_{020} + \zeta_{020k}$$

The multilevel model results indicate that the school mean pretest is predictive of the posttest achievement. A one standard deviation increase in school pretest score is associated with a .88 increase in outcome scores. The effect of assignment to the treatment condition is positive, but it is not statistically significant.

Insert Table 2 here

This preliminary analysis indicates that the EM program does not have a statistically significant impact on students' math achievement compared to the business as usual curriculum in place in control schools after the first year of implementation. The results however do indicate that variation in the treatment effect by school was significant (i.e., greater than two standard deviations) and district variation was greater than one standard deviation. This possible effect will be explored further in additional analyses focusing on subgroup variation and differences in fidelity of implementation.

Insert Figure 1 here

Conclusions:

Early preliminary results indicate no statistically significant impacts of EM after one year when implemented at scale relative to business as usual math programs in a sample of 48 schools Year one findings will be further explored by assessing fidelity of implementation and subgroup impacts as well as potential mediation pathways in the full paper.

Appendices

Appendix A. References

Carroll, W.M. (1998). Geometric knowledge of middle school students in a reform-based mathematics curriculum. *School Science and Mathematics*, *98*(4), 188-197.

Carroll, W.M., & Isaacs, A. (2003). Achievement of students using the University of Chicago School Mathematics Project's Everyday Mathematics. In S.L. Senk & D.R. Thompson (Eds.), *Standards-based school mathematics curriculum: Where are they? What do students learn?* (pp. 79-108). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Klein, D. (2007). A quarter century of U.S. "math wars" and political partisanship. *Journal of the British Society for the History of Mathematics, 22*(1) 22-33.

National Center for Educational Statistics. (2013). *A First Look: 2013 Mathematics and Reading (National Assessment of Educational Progress at Grades 4 and 8).* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through Grade 8 mathematics: A quest for coherence*. Reston, VA: National Council of Teachers of Mathematics.

National Mathematics Advisory Panel. (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. U.S. Department of Education: Washington, DC.

National Research Council. (2004). *On Evaluating Curricular Effectiveness: Judging the Quality of K-12 mathematics Evaluations*. Committee for a Review of the Evaluation Data on the Effectiveness of NSF-Supported and Commercially Generated Mathematics Curriculum Materials. Mathematical Sciences Education Board, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Riordan, J.E., & Noyce, P.E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, *32*(4), 368-398.

SRA/McGraw-Hill (2003). Everyday Mathematics student achievement studies: Volume 4. Chicago: Author.

SRA/McGraw-Hill (2009). Kate Powell, Marketing Manager, Everyday Mathematics, Wright Group McGraw-Hill, personal communication, May 21, 2009.

SRA/McGraw-Hill (2009b). District of Columbia Public Schools Everyday Mathematics, Third Edition Evaluation Research. Unpublished manuscript.

Waite, R.D. (2000). A study of the effects of Everyday Mathematics on student achievement of third-, fourth-, and fifth-grade students in a large north Texas urban school district. *Dissertation Abstracts International*, *61*(10), 3933A. (UMI No. 9992659)

What Works Clearinghouse (2007, July). *What Works Clearinghouse topic report: Elementary School Math.* Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from: http://ies.ed.gov/ncee/wwc/reports/elementary_math/topic/

What Works Clearinghouse. Revised (2009, March). *What Works Clearinghouse intervention report: Everyday Mathematics.* Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from: http://ies.ed.gov/ncee/wwc/reports/elementary_math/eday_math/

Woodward, J., & Baxter, J. (1997). The effects of an innovative approach to mathematics on academically low-achieving students in inclusive settings. *Exceptional Children*, *63*(3), 373-388.

Appendix B. Tables and Figures

Table 1. School Characteristics for Study Schools by Treatment Assignment at Baseline

		All Schools	Everyday Math Treatment Schools	Business as Usual Schools	p-value
Title 1 Eligible (%)[1]		83%	83%	84%	1.000
Schoolwide Title I Eligible (%)[1]		83%	83%	83%	1.000
Students Eligible for Free or Reduced Price Meals (%)[1]		73%	75%	71%	0.633
Student Enrollment (average)[2]	Grade K	69	68	70	0.773
	Grade 3	68	65	71	0.349
Student Gender (%)[2]	Male	51%	52%	50%	0.278
	Female	48%	47%	48%	0.599
Student Race/Ethnicity (%)[2]	White	40%	40%	40%	0.997
	Non-Hispanic black	51%	52%	50%	0.853
	Hispanic	6%	5%	6%	0.822
	Other	4%	3%	5%	0.333
ELL (%)[2]		6%	5%	8%	0.517
Special Educaton (%)[2]		7%	7%	8%	0.461
Schools		**48**	**24**	**24**	
Students enrolled at Baseline[2]		6,566	3,196	3,370	

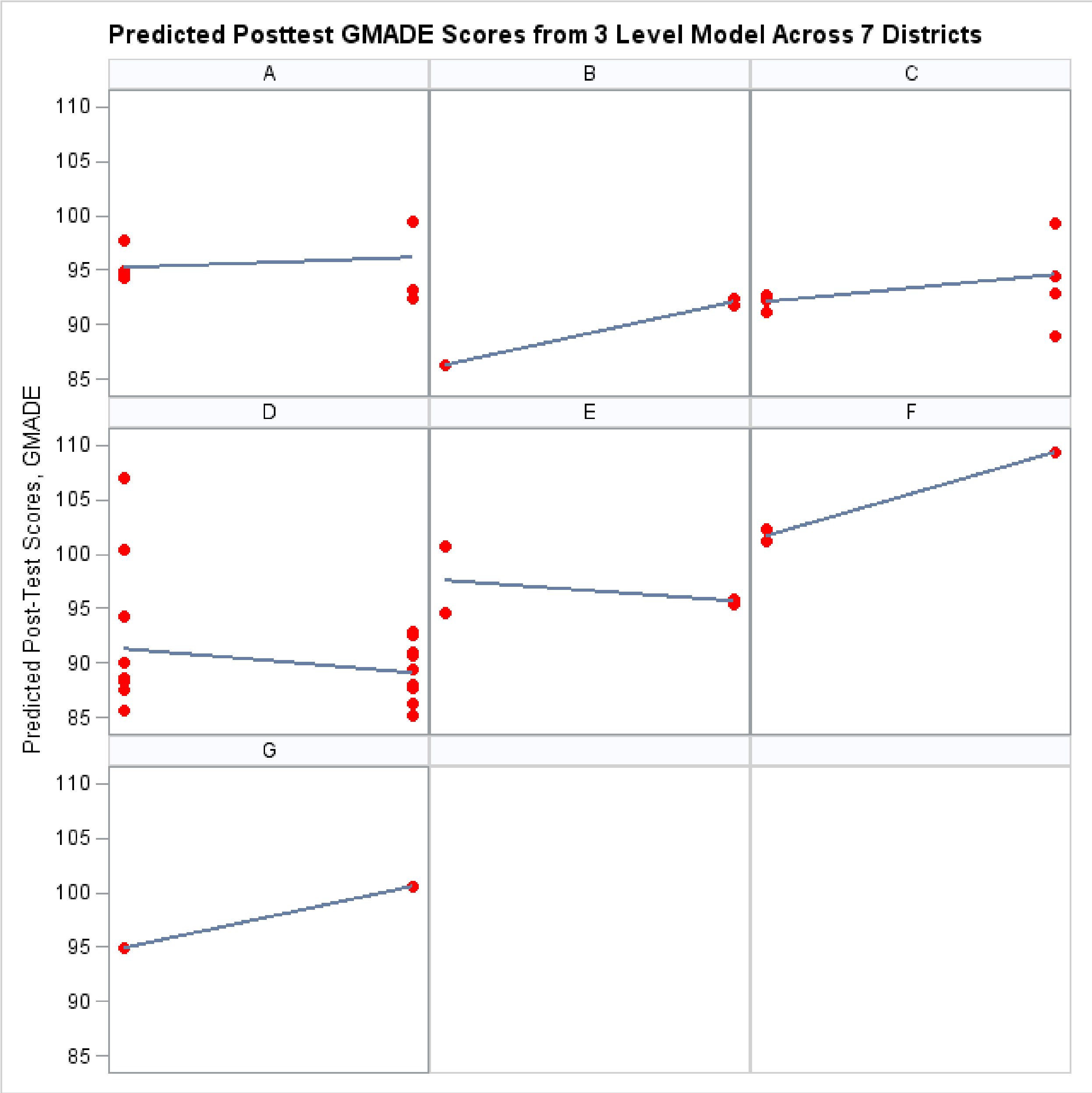
Note 1:Data source: CCD Public School Data 2011-2012 school year, National Center for Education Statistics.

Note 2:Data Source:student-level data collected from participating schools.

Table 2. Multilevel Model Estimates for Impact of Everyday Mathematics on Student Math Achievement

Fixed effects model	Coefficient	Standard Error	Degrees of freedom	t-value	p-value
Intercept	93.833	0.623	10.9	150.570	<.0001
GMADE Pretest (grand-mean, school mean score)	0.882	0.055	40.3	16.100	<.0001
Treatment	0.705	0.577	40.5	1.220	0.229

Random effects	Variance component	Standard Deviation
School Level	1.838	0.883
District Level	1.291	1.096



Predicted Posttest GMADE Scores from 3 Level Model Across 7 Districts
A
B
C
D
E
F
G
110
105
100
95
90
85
Predicted Post-Test Scores, GMADE