Exceptional Children

Vol. 80, No. 4, pp. 423–437.

DOI: 10.1177/0014402914527240

Effects of Blended Instructional Models on Math Performance

BRIAN A. BOTTGE
XIN MA
LINDA GASSAWAY
MICHAEL D. TOLAND
MARK BUTLER
University of Kentucky
SUN-JOO CHO
Peabody College of Vanderbilt University

ABSTRACT: *A pretest-posttest cluster-randomized trial involving 31 middle schools and 335 students with disabilities tested the effects of combining explicit and anchored instruction on fraction computation and problem solving. Results of standardized and researcher-developed tests showed that students who were taught with the blended units outscored students in Business As Usual classes. Students made the largest gains in computing with fractions and on problems related to ratios, proportions, and geometry. The findings suggest important implications for the way curriculum is designed for middle school students with disabilities who exhibit low performance in math.*

The National Assessment of Educational Progress (NAEP; National Center for Education Statistics, 2011) reported that almost two thirds (65%) of students with disabilities in eighth grade scored below the *Basic* level compared to 23% of students without disabilities. To achieve *Basic* on the NAEP, students must have an understanding of arithmetic operations with whole and rational numbers and be able to solve "real world" problems using charts, graphs, and technology tools such as calculators and computers.

Much of the students' low math performance can be traced to computing with fractions and problem solving (see Misquitta, 2011), two areas of difficulty that have been identified for many years on large-scale assessments and in research studies. For example, results of the 1996 NAEP (Silver & Kenney, 2000) indicated many eighth grade students (65%) could not order five fractions less than 1 from least to greatest and more than one third of students (35%) did not choose 1/3 as the fraction equivalent to 4/12. In problem solving, students have difficulty focusing on relevant parts of problems (Kauffman, 2001), organizing information (Geary, 2004), and monitoring their work (Gallico, Burns, & Grob, 1991; Montague, Bos, & Doucette, 1991).

These findings have put more pressure on teachers to emphasize in earlier grades complex math concepts such as fractions and algebra (e.g., Butler, Miller, Crehan, Babbitt, & Pierce, 2003; Scheuermann, Deshler, & Schumaker, 2009). The *Common Core State Standards for Mathematics* (CCSS-M, National Governors Association Center for Best Practices, Council of Chief State School Officers 2010) recommends that students develop understanding of fractions in Grade 3, equivalence in Grade 4, and fluency with adding and subtracting fractions by the end of Grade 5. Stressing the importance of fractions, the Brookings Institution (2004) and the National Mathematics Advisory Panel (NMAP, 2008) have suggested adding more fraction items on the NAEP test noting their underrepresentation.

In problem solving, the National Council of Teachers of Mathematics (NCTM; Briars, Asturias, Foster, & Gale, 2012) recommends that all students have opportunities to solve quality problems that are motivating and help students build confidence in their ability to figure things out on their own. Problem solving poses special challenges for low-performing students because problems are typically embedded in text that is difficult to decipher for students with comorbid difficulties in reading and math (Knopik, Alarcón, & DeFries, 1997; Parmar, Cawley, & Frazita, 1996).

Much of the students' low math performance can be traced to computing with fractions and problem solving, two areas of difficulty that have been identified for many years on large-scale assessments and in research studies

Over a decade ago we began testing the effects of enhanced anchored instruction (EAI), an instructional method we developed for improving the computation and problem solving skills of middle students with disabilities in math. EAI is based on the concept of anchored instruction (AI; Cognition and Technology Group at Vanderbilt, 1997) and follows closely the theoretical base in our model for teaching and learning math (Bottge, 2001). Realistic problems (i.e., anchors) are embedded in interesting contexts and presented in interactive, video-based formats. The anchors consist of an 8- to 15-min video in which adolescents are shown attempting to solve a challenging problem.

First versions of EAI included one video-based problem and one simple hands-on application. Based on a series of studies, we made improvements to EAI by adding more complex hands-on applications (enhancements) to help students "visualize" abstract concepts, a strategy that aligns closely with recommendations in the Institute of Education Sciences' *Practice Guide for Organizing Instruction to Improve Student Learning* (Pashler et al., 2007). The hands-on projects afforded multiple opportunities for students to improve their procedural skills and helped them gain a deeper understanding of the math concepts embedded in each problem. We also developed computer-based modules to boost students' performance with fractions. By representing math concepts in engaging interactive media and providing multiple practice opportunities, we hoped to reinforce the working memory that many low-achieving students lack (Swanson, Jerman, & Zheng, 2008).

Math educators now suggest that teachers work on developing students' problem-solving skills, procedural fluency (e.g., add and subtract fractions), and conceptual understanding of key math concepts (e.g., represent fractions on a number line, identify equivalent fractions) (e.g., Baroody, Feil, & Johnson, 2007) at the same time rather than teaching them in isolation. The NMAP (2008) stated that "a conceptual understanding of fractions and decimals and the operational procedures for using them are mutually reinforcing" and that "curriculum must simultaneously develop conceptual understanding, computational fluency, and problem solving skills" (p. xix). Results of recent studies (e.g., Bottge, Rueda, Grant, Stephens, & LaRoque, 2010; Cho, Bottge, Cohen, & Kim, 2011) showed the potential of teaching as NMAP suggests.

The purpose of this study was to test the effects of five EAI units on a large sample of teachers and students in a randomized experiment. Research with EAI has progressed through several stages, beginning from small-scale basic and applied studies to efficacy and effectiveness trials. Based on these early findings, we tested combinations of the EAI units in a variety of instructional settings. Using Sloane's (2008) continuum of research phases for developing mathematics education interventions, we consider the present experiment a large-scale implementation trial that tests "the effectiveness of an efficacious intervention under real-world conditions" (p. 628). Unlike previous studies that were smaller in scale and more controlled, this experiment tested the treatment potential of the complete package of EAI units taking into consideration the realities of day-to-day classroom challenges. Specifically, we designed the study to answer the following research questions:

1. What are the differential effects of EAI and business as usual (BAU) instruction on students' computation with fractions?
2. What are the differential effects of EAI and BAU instruction on students' problem solving performances?

RESEARCH DESIGN AND METHOD

Design

A pretest/posttest, cluster-randomized school-based trial tested the efficacy of the two instructional conditions (EAI vs. BAU) on students' ability to compute fractions and solve problems. Overall test scores and item-error response patterns were analyzed to assess changes in computation skills and conceptual understanding attributable to instruction. Random assignment was conducted at the school level and not at the class or student level for two main reasons. First, administrators and teachers at each school scheduled students to classes. Second, we wanted to avoid contamination within schools due to teachers sharing instructional strategies across treatment conditions (Raudenbush, 2008). This would likely happen because of the school-wide attention some of the applied EAI instructional modules (e.g., hovercraft races) often draw.

Participants and Setting

A total of 335 students with disabilities and their special education teachers in 31 middle schools participated in the study. Parents, students, and teachers signed consent forms approved by the university institutional review board for participating in the study. Schools were located in and around a large metropolitan area in the Southeast. Table 1 shows demographic information about the teachers and students. Originally, 32 schools

TABLE 1
Teacher and Student Characteristics

	Teachers				
	BAU (n = 26)	EAI (n = 23)	χ^2	t	p
Gender			0.13[a]		.76
Male	8	6			
Female	18	17			
Ethnicity			0.18[a]		.67
Caucasian	25	20			
African American	1	2			
Asian	0	1			
Years teaching special education				1.62	.11
M	12.21	9.07			
Median	11.50	8.00			
SD	6.86	6.68			
Range	2-27	1-28			
Highest degree earned			0.06[a]		.99
BA, BS	6	6			
MA, MS	20	17			
	Students				
	BAU (n = 176)	EAI (n = 159)	χ^2		p
Gender			2.04[a]		.15
Boys	111	112			
Girls	65	47			
Grade			4.25[a]		.05
6	2	7			
7	92	60			
8	82	92			
Ethnicity			5.43		.07
Caucasian	137	128			
African American	28	26			
Latino	6	0			
Native American	3	0			
Biracial and other	2	2			
Disability/service area			1.93		.75
MMD	68	59			
OHI	63	50			
SLD	24	25			
Autism	12	15			
EBD	9	10			
Subsidized lunch[b]	130	119	0.19[a]		.70

Note. BAU = business as usual; EAI = enhanced anchored instruction; MMD = mild mental disability; OHI = other health impaired; SLD = specific learning disability; EBD = emotional/behavioral disability. Less populated categories combined for Chi-square test: Teacher ethnicity reduced to Caucasian and Other; grade reduced to 6-7 and 8; student ethnicity reduced to Caucasian, African American, and Other; disability reduced to MMD, OHI, SLD, autism, and EBD/other.
[a]Fisher's Exact Test.
[b]Six students did not have a subsidized lunch status.

began the study, with 16 schools randomly assigned to EAI and 16 schools randomly assigned to BAU. However, a participating teacher in one of the EAI schools resigned for medical reasons before administering the pretests. This left a total of 15 EAI schools (23 teachers, 33 resource rooms) and 16 BAU schools (26 teachers, 31 resource rooms) completing the study. Teachers were comparable across instructional conditions in gender, ethnicity, education level, and teaching experience. Most teachers were White, female, well educated, and experienced.

Instruction in both conditions took place in special education resource rooms. Students received all their math instruction in these pullout settings because the Admissions and Release Committee (ARC; Kentucky Department of Education, 2008) at each school thought their skills were too low to benefit from instruction in general education math classes. The class size of the EAI resource rooms (M = 4.82, SD = 1.36) and BAU resource rooms

Problem solving poses special challenges for low-performing students because problems are typically embedded in text that is difficult to decipher for students with comorbid difficulties in reading and math

(M = 5.68, SD = 2.67) did not differ, $t(49.41)$ = 1.87, p = .07. Most students were receiving special education services for a disability in one of three categories: Mild Mental Disability (MMD), Other Health Impairment (OHI), or Specific Learning Disability (SLD). A small number of students were receiving services for Autism or Emotional Behavioral Disabilities (EBD). Students were comparable across groups in gender, ethnicity, subsidized lunch, and disability area.

Description of EAI

The five EAI units included computer-based interactive lessons, video-based anchored problems, and hands-on applied projects. Together, the units address several of the CCSS-M (2010), especially Ratios and Proportional Relationships, Number System, Statistics and Probability, and Geometry. Teachers were provided with daily lesson plans that included the lesson objectives, materials, warm-up exercises, and detailed descriptions of how to implement each lesson (see "Professional Development"). Based on recommendations of the NMAP (Gersten et al., 2009) and meta-analyses (e.g., Gersten et al., 2009), the lessons included a mix of explicit instruction and problem-solving activities. The mean number of instructional days was 94.1 and 93.7 for EAI and BAU classes, respectively. Most class sessions in both conditions ranged from 45 to 60 min with a few classes meeting in 90-min blocks. The total possible number of instructional minutes varied by less than .01% between EAI and BAU.

Fractions at Work (FAW). A series of computer-based modules and concretized manipulatives called FAW was used to develop students' understanding and procedural competence with rational numbers. The instructional package contains seven chapters divided into self-contained units to help students understand concepts (e.g., fraction equivalence) and develop their computational skills (e.g., adding and subtracting simple and mixed fractions). In combination with the explicit explanations provided in each unit, the lesson plans called for using concrete materials such as fraction strips to help students understand equivalence. Representing fractional components this way follows the recommendations in the Institute of Education Sciences' Practice Guide, What Works Clearinghouse (2011), *Developing Effective Fractions Instruction for Kindergarten Through 8th Grade* (Siegler et al., 2010). FAW took students an average of 20 instructional days to complete.

Fraction of the Cost (FOC). The problem-solving anchor, FOC, is an 8-min video on CD that tells the story of three friends who wish to build a skateboard ramp. Available in English and Spanish, the video stars three middle school students who study a schematic plan of a skateboard ramp and then discuss how they can afford to build it with the materials and money they have available. The solution path is not straightforward. To solve the problem, students (a) calculated percent of money in a savings account and sales tax on a purchase, (b) read a tape measure, (c) converted feet to inches, (d) deciphered building plans, (e) constructed a table of materials, (f) computed mixed fractions, (g) estimated and computed combinations, and (h) calculated total cost. The CD contains several interactive tools to help students solve the problem. For example, students can construct the ramp in multimedia space by dragging individual boards onto a template, which helps them keep track of which boards they have already accounted for. Students completed the unit in an average of 17 instructional days.

Hovercraft Project (HOV). The third unit in the series consists of a hands-on problem called the HOV, which asks students to design and construct a rollover cage out of PVC pipe for a hovercraft they ride when the project is successfully completed. A leaf blower inserted into a plywood platform lifts the hovercraft off the floor by blowing air through holes in a plastic sheet fastened to the underside of the base. In small groups, students designed their own HOV rollover cages by drawing the side and top views to scale on graph paper with 1/4" equal to 2" of pipe. The cage must meet several design specifications (e.g., 4 feet long, 3 feet wide) and include knee, side, and head protection. During the last part of the unit, students used their scale model plans and straw cage to build the full size rollover cage. On the final day of the activity, students attached their rollover cage to the hovercraft and took turns riding on them up and down the hallways of the school. This unit took an average of 29 instructional days to complete.

Kim's Komet (KK). KK is one episode in a series called *The New Adventures of Jasper Woodbury* (Learning and Technology Center at Vanderbilt University, 1997). The video anchor involves two girls who compete in pentathlon events. The first challenge asks students to identify the three fastest qualifiers in three regional races where times and distances are known but the distances vary. For example, students explained whether a car that travels 15 feet in .9 s is faster or slower than a car that travels 20 feet in 1.3 s. The second challenge required students to construct a graph and line of best fit to predict the speed of cars at the end of a straightaway when released from any height on the ramp. The purpose of KK is to help students develop their informal understanding of pre-algebraic concepts such as linear function, line of best fit, variables, rate of change (slope), reliability, and measurement error. Students completed KK in an average of 19 instructional days.

Grand Pentathlon (GP). After students solved the problems in KK, they competed in their own pentathlon competition with a full-sized ramp and events like the ones shown in the video. In the GP, students made their own cars out of blocks of wood and timed them from several release points on the ramp. An infrared detector measured the time of the cars from the beginning to end of the straightaway. After students made their graphs and plotted their times, the teacher revealed to students the range of speeds within which their cars should be traveling to negotiate each event. Students had to construct their own graphs and lines of best fit to predict the speed of their cars at the end of the straightaway for every release point on the ramp. This was the shortest unit taking an average of nine instructional days.

Professional Development

Previous studies with EAI revealed the need to provide teachers with high-quality professional development, especially in the appropriate use of the multimedia-based units (i.e., FOC, FAW, KK), the hands-on projects (i.e., HOV, GP), and some of the more complex math concepts (e.g., reliability and measurement error, drawing schematic plans to scale). The main activity was a 2-day summer workshop conducted by a middle school math teacher who had taught with EAI for several years. He modeled teaching strategies and gave teachers opportunities to play the part of "students" during sample lessons. For example, teachers worked in small groups to design, draw to scale, and construct their own rollover cages for hovercrafts they rode up and down the hallways. They also had ample opportunity to discuss issues related to lesson study, assessment techniques, and technology use. The training session was video recorded and placed on a secure computer server, which made the training segments available for EAI teachers to review during the school year.

Description of BAU Instruction

Teachers assigned to the BAU condition followed their regular school math curriculum, which most schools had developed to align with the principles contained in the state's Department of Education Combined Curriculum Document. Data retrieved from teachers' self-reports, lesson plans, and classroom observers showed the learner objectives in BAU classrooms paralleled those of the EAI units. BAU objectives focused on analyzing proportional relationships, solving real-life problems with whole numbers and fractions, graphing of ordered pairs, working on concepts related to geometry, and basic pre-algebraic expressions. Most teachers taught with textbooks that were used in the general education math classrooms. BAU teachers also used interactive whiteboards along with manipulatives. Calendar Math (e.g., Gillespie & Kanter, 2000), which is aligned with the NCTM curriculum standards, was a key part of the daily math instruction in several BAU classrooms.

We matched the schedules of BAU instruction in schools to those of the EAI schools over the course of the school year. Teachers were asked to keep track of interruptions to instruction (e.g., statewide exams, school-wide assemblies, snow days) in their Teacher Logs. Some lessons focused on data representation and making predictions using data and graphs. For example, one teacher displayed data from a table on a line graph and asked students to make predictions. He displayed a "graph" (in the form of a coordinate plane) and defined the term "axes," which he drew and "labeled with my data, my information." The "labels" he referenced were intervals of a number line that he placed on each axis. Next, he stated "all graphs need two things, they need titles and they need labels." This graph represented time versus number of successful free throws displayed in a table. After the teacher plotted several points, he used a straight edge to show that the points "made a straight line" and that they could be used to predict other points. The entire exercise was displayed on the board and students simply copied the work to their notebooks as the teacher explained the procedures and concepts. Next the teacher led the whole group through a problem displaying the age of a puppy versus its weight for which the solution was not displayed. The teacher led students through drawing a graph, constructing number lines on each axis, and assigning labels and a title. He plotted the points and demonstrated how to use a protractor to fit a line to the existing points and then predict its weight.

Other lessons paired instruction of fractions with word problems. For example, students placed five strips of paper (1 whole, 1/2, 1/4, 1/8, 1/16) on their desks and then the teacher played a video (i.e., animated cartoon) that described procedures for rounding numbers. At the end of the video, the teacher projected "Word Problems" on the screen that the students had been taught the previous day. The teacher asked volunteers to read the problems out loud and to share their problem solving strategies.

EAI and BAU Classroom Observations

Trained observers (graduate assistants, project manager, principal investigator [PI]) collected a total of 324 whole-class classroom observations using FileMaker Pro 10. The data template included space for recording demographic information (e.g., school, instructional condition), level of treatment fidelity (e.g., quality of lesson implementation), and open areas for describing classroom activities (e.g., specific teacher and student behaviors). The basis for computing EAI treatment fidelity was derived from the daily lesson plans that teachers were provided during training. Primary observers conducted 173 EAI whole-class observations that indicated teachers taught the lesson plan activities 95% of the time. A secondary observer was present during 34 of the observed class periods (19.6%). Interobserver agreement was 94% for the main part of the lesson. We also collected a total of 117 primary observations and nine secondary observations of BAU classes. Interobserver agreement was 94%.

Academic Measures

Two researcher-developed tests and two standardized achievement tests measured the effects of EAI on students' computation and problem solving skills. They were administered over 3 consecutive days immediately prior to and following the instructional period. Interrater agreement was calculated on 20% of each of the pretest and posttest measures.

Fractions Computation Test. A 20-item (14 addition, six subtraction), 42-point test measured students' ability to add and subtract simple fractions and mixed numbers with like and unlike denominators. All but one of the items included fractions that could be found on a ruler. The test also included some items that required students to add three fractions. Students were told to reduce their answers to simplest form and to show all their work. Calculator use was not allowed. On 18 of the items, students earned 1 point for correct work and 1 point for the correct answer. On two items with mixed numbers that required renaming prior to subtracting, students were awarded an additional point. Internal consistency estimates for this sample were .81 at pretest and .96 at posttest. Interrater agreement was 97%.

Problem-Solving Test (PS1-PS2). Problem-Solving Test Part 1 (PS1) and Problem-Solving Test Part 2 (PS2) were used to sample the problem solving performances of students and, specifically, their understanding of concepts in the four EAI problem solving units. Each open response item presented a question about a figure, table, or graph, and a box where students showed their work. For most items, students earned partial credit for showing they knew how to solve the problem and full credit for computing the correct answer. Some of the problems were in standard word problem format with the reading level kept at or below the fourth grade. Students could use calculators. Versions of the tests went through cycles of refinement based on student performances in previous research and on suggestions from math and assessment specialists (i.e., math teachers, math researchers, test consultants).

PS1 consists of 12 items (10 open response items and two multiple-choice questions) worth 20 points assessing the concepts closely aligned with the CCSS-M Measurement and Data, Number and Operations–Fractions, and Ratios and Proportional Relationships. One open response item asked students to measure the length of an object with neither endpoint positioned at zero inches, similar to the "toothpick" item on the NAEP tests in 1996, 2000, and 2003. The 2003 results showed that 58% of the students in eighth grade correctly chose the correct answer, which was presented in multiple-choice format. Other items asked students to determine a length given a measurement scale (i.e., one square of building plan equals two inches), interpret a bank statement to calculate 10% of a bank balance (i.e., finding a percent of a quantity as a rate per 100), and show fractions on a number line. Internal consistency estimates were .76 at pretest and .82 at posttest.

PS2 consists of nine items (all open response) worth 15 points that assessed the concepts closely aligned with the CCSSI-M Ratios and Proportional Relationships and Geometry. To figure out the correct answers, students had to understand figures, interpret data from tables and graphs, construct their own tables and graphs, recognize relationships among these data, and make predictions based on their solutions. Some items asked students to solve a word problem with an accompanying figure by computing time when rate and distance were given or represented in a figure. Another item had students solve a word problem by computing rate when distance was given but time was not explicitly stated. Internal consistency estimates were .63 at pretest and .73 at posttest.

Concurrent validity correlation coefficients based on pretest scores of the *Iowa Tests of Basic Skills* (ITBS) Problem Solving and Data Interpretation Subtest (University of Iowa, 2008) and PS1-PS2 ranged from .50 to .52 and sample coefficients ranged from .40 to .44, which appear acceptable given that the range of mathematics concepts sampled by the problem solving test items was more restricted than that sampled by the ITBS. Interrater agreement across both parts of PS1 and PS2 was 99%.

ITBS. The ITBS (Form C, Level 12; University of Iowa, 2008) is a norm-referenced standardized achievement test that was used to complement the researcher-developed tests. Two math subtests of the ITBS were administered according to the directions in the test administration booklet. Students entered answers directly into their test booklet rather than on a separate answer sheet, a procedure identified as an accommodation for students with disabilities.

The 30-item Computation subtest required students to perform one of the four arithmetic operations (addition, subtraction, multiplication, or division) with whole numbers, fractions, and decimals. Ten items assessed whole number operations. Of the 10 fraction computation items, there were four each of addition and subtraction and two of multiplication. Of the 10 decimal items, three were addition, three subtraction, three multiplication, and one division. Students worked each problem and then selected their answer from three choices or "N" if students thought the answer was not one of the choices.

The 28-item Problem-Solving and Data Interpretation subtest consisted of problems in several formats that took one or more steps to solve. Of the items at this test level, 12 were word problems and four required students to use data displays to obtain information and compare quantities. The other items asked students to compute answers from interpreting graphs, charts, and tables.

Per test administration instructions, students were allowed to use calculators for the problem solving test but not for the computation test. Sample KR20 estimates were .72 and .78 for Computation and .61 and .58 for Problem-Solving and Data Interpretation pretests and posttests, respectively. Interrater agreement on each multiple-choice subtest was 100%.

Data Analysis

We employed a three-level hierarchical linear model (HLM; Raudenbush & Bryk, 2002) to evaluate student performance on each of the outcome measures controlling for five variables (gender, grade level, free-reduced lunch, race-ethnicity, and disability status) at the student level and three variables (gender, teaching experience, and graduate degree) at the teacher level. The analyses employed a full information maximum likelihood estimation method, which uses all available data except those missing on a primary outcome measure.

The Level 1 model situated at the student level included the control variables together with the pretest and posttest scores. Using the ITBS Computation Test Score as an example with ITBS_C1 as the pretest score and ITBS_C2 as the posttest score, the Level 1 model was:

$$ITBS\_C2_{ijk} = \beta_{0jk} + \beta_{1jk} ITBS\_C1_{ijk} + \sum_{n=1}^{9} \beta_{(n+1)jk} SC_{nijk} + \varepsilon_{ijk}, \qquad (1)$$

where ITBS_C2$_{ijk}$ is the computation score after the treatment (posttest) for student *i* with teacher *j* in school *k*, ITBS_C1$_{ijk}$ is the computation score before the treatment (pretest) for the same student, and ε_{ijk} is an error term unique to each student, assumed $\varepsilon_{ijk} \sim N(0, \sigma^2)$. Student characteristic is indicated with SC_{nijk} (*n* is an index for student characteristics, $n = 1, 2, \ldots, 9$) as control variables. The average computation posttest score of students is represented by β_{0jk} for teacher *j* adjusted for student characteristics and computation pretest scores of students for that teacher, β_{1jk} is the regression coefficient for the computation score before the treatment (pretest), and $\beta_{(n+1)jk}$ is the regression coefficient for each of the respective student control variables.

The Level 2 model was situated at the teacher level and included the control variables and treatment condition (one dummy variable with BAU as the baseline against which EAI was compared) with β_{0jk} modeled as the dependent variable. The Level 2 model was:

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} EAI_{jk} + \sum_{m=1}^{3} \gamma_{0(m+1)k} TC_{mjk} + u_{0jk}, \qquad (2)$$

where γ_{00k} is the average posttest computation score for school *k* adjusted for the treatment condition (EAI vs. BAU) and teacher characteristics TC_{mjk} (*m* is an index for teacher characteristic, $m = 1, 2, 3$) as control variables; γ_{01k} is the regression coefficient for the treatment condition, which represents the treatment effect; $\gamma_{0(m+1)k}$ is the regression coefficient for each of the respective teacher control variables, and u_{0jk} is an error term unique to each teacher, assumed $u_{0jk} \sim N(0, \tau_\pi^2)$.

Finally, the Level 3 model was situated at the school level:

$$\begin{aligned} \gamma_{00k} &= \theta_{000} + \nu_{00k}, \\ \gamma_{01k} &= \theta_{010}, \\ \gamma_{0(m+1)k} &= \theta_{0(m+1)0}, m = 1, 2, 3, \end{aligned} \qquad (3)$$

where θ_{000} is the average posttest computation score and ν_{00k} is the school-level error, assuming $\nu_{00k} \sim N(0, \tau_\beta^2)$. Because randomization was conducted at the school level to control for school characteristics, the school model was left empty without any adjustment. Its main function was to account for the data hierarchy of teachers nested within schools. The regression coefficient for the treatment condition, γ_{01k}, was specified as the same across schools at Level 3 (i.e., specifying EAI as fixed effects at the teacher level), which is a statistically efficient, parsimonious way to examine the treatment effects (e.g., Spybrook, Raudenbush, Liu, & Congdon, 2006). The regression coefficient for each of the respective teacher control variables, $\gamma_{0(m+1)k}$, was also specified as the same across schools at Level 3. All statistical significance tests were performed at the alpha level of .05. Hedge's *g* was used as an effect size (ES) measure for treatment effects.

The interaction between treatment and pretest is often an issue in experimental design. To examine this interaction, we first set the coefficient of the pretest variable, which represents the effect of pretest on posttest as random at the student level. If this effect does not show statistically significant variance across teachers, the relationship of pretest to posttest is the same across teachers, implying that this relationship is not moderated by treatment (i.e., lack of treatment by pretest interaction). If this effect does vary across teachers, the treatment condition is used to model this variation. If the treatment condition is not statistically significant, the relationship of pretest to posttest is not moderated by treatment (i.e., lack of treatment by pretest interaction). We performed this procedure for our 24 outcome measures. Three of them (Mixed Number Addition, Unlike Denominator of the Fractions Computation Test, Whole Number Multiplication of the ITBS Computation) showed statistically significant variation in the relationship of pretest to posttest at the teacher level. The treatment condition, however, was not statistically significant in any of the three cases. We concluded that there was no concern about the interaction between treatment condition and pretest scores.

RESULTS

The overall effects of EAI are reported on each of the cognitive measures followed by an in-depth analysis of the test subscales. Table 2 shows descriptive summaries of the pretests and posttests for each measure. Three standard deviations on posttest results under the EAI condition were much larger than the rest (i.e., Fractions Computation Test as a whole and its subscales of Overall Addition and Unlike Denominator). This indicates that EAI students performed quite differently on posttests of these three outcome measures—namely, some EAI students performed far ahead of other EAI students. Table 3 displays results of the HLM analyses estimating effects of EAI on total and subscale scores. Part of these HLM analyses yielded what is often referred to as ICC (intraclass correlation), which is related to the partition of the total variance in a certain outcome measure into variance components at different levels (student, teacher, and school). Because of space restrictions, ICCs for each level are not reported but are available from the first author.

The coefficient estimate (γ_{010}) in Table 3 represents the estimated difference between the means of EAI and BAU posttest scores adjusted for pretest scores on the respective outcome and student characteristics at the student level and teacher characteristics at the teacher level. This coefficient can also be understood as representing the adjusted difference between EAI and BAU in gain scores from pretest to posttest. Table 3 contains results directly relevant to our research questions. In general, EAI students scored higher than BAU students on three of the four achievement measures. Also, student and teacher characteristics were used as control variables in data analysis and a couple of them were statistically significant. Because of space restrictions, information on those statistically significant student and teacher covariates is not reported but is available from the first author.

Computation–Research Question 1

The Fractions Computation Test showed across-the-board improvement of EAI students over BAU students on all 10 subscales. EAI students gained about one standard deviation more than BAU students on each of the three addition subscales and close to one standard deviation more on the subtraction subscales. This improvement was consistent across fractions with like and unlike denominators and on subscales where rewriting or not rewriting answers was possible.

On the ITBS Computation Test, we grouped items by objective and analyzed them as subscales. EAI students showed statistically significant larger gains than BAU students in their ability to both add and subtract fractions. Treatment effects were one half or more of a standard deviation and supplement the results of the researcher-developed Fractions Computation Test. The overall treatment effect on the ITBS Computation Test was not due to an improvement on all subscales but rather to strong treatment effects on three of the six subscales that involved computation with fractions. As was the case with the Fractions Computation Test, students' most frequent mathematical error was Combining in the fraction subscale (including addition and subtraction) and reducing this error was associated with statistically significant improvement from pretest to posttest in both instructional groups (ES = .56, $p < .01$).

TABLE 2

Descriptive Statistics of Pretest and Posttest Scores on Subscales of Cognitive Measures

	BAU				*EAI*			
	Pretest		*Posttest*		*Pretest*		*Posttest*	
Subscales	M	SD	M	SD	M	SD	M	SD
Fractions Computation Test (42)	3.04	4.59	4.74	6.65	2.34	3.38	17.36	13.70
Addition								
Overall (28)	1.93	3.34	3.01	4.76	1.27	2.06	12.44	10.10
Simple (16)	1.78	2.53	2.65	3.32	1.25	1.91	8.24	5.65
Mixed (12)	.15	1.14	.36	1.78	.03	.33	4.20	4.82
Subtraction								
Overall (14)	1.11	1.76	1.74	2.39	1.07	1.80	4.92	4.22
Simple (4)	.69	.99	.97	1.15	.60	.94	2.21	1.58
Mixed (10)	.42	.93	.77	1.50	.47	.96	2.71	2.95
Denominator								
Like (8)	2.59	3.01	3.50	3.18	2.19	3.04	5.84	2.77
Unlike (34)	.45	2.90	1.25	4.96	.15	1.09	11.52	12.03
Rewriting Answer								
No rewrite (14)	1.09	2.25	1.78	3.10	.71	1.42	8.28	7.07
Rewrite (28)	1.95	2.63	2.96	4.59	1.63	2.28	9.07	6.82
Problem Solving Test (PS1 and PS2, 35)	5.32	3.51	7.40	4.51	6.16	3.57	9.52	4.30
Ratios and Proportional Relationships (8)	1.10	1.46	2.29	2.27	1.34	1.74	3.72	2.59
Measurement and Data (10)	3.78	3.28	4.87	3.57	4.65	3.41	5.78	3.36
Number and Operations–Fractions (8)	2.33	1.81	3.18	2.41	2.52	1.83	3.69	2.32
Geometry–Graphing (9)	3.23	2.54	4.16	2.54	3.38	2.38	5.55	2.62
ITBS Computation (30)	10.27	3.79	10.73	4.33	11.47	5.18	13.62	5.48
Fraction								
Addition (4)	1.46	.88	1.51	.96	1.46	.89	2.22	1.23
Subtraction (4)	1.87	1.10	2.13	1.10	1.89	1.10	2.64	1.02
Whole Numbers/Decimals								
Addition (5)	1.80	1.11	1.86	1.33	2.32	1.35	2.54	1.38
Subtraction (5)	1.47	1.31	1.71	1.40	1.96	1.57	2.32	1.72
Multiplication (6)	1.51	1.23	1.38	1.32	1.67	1.63	1.90	1.74
Division (4)	1.34	.99	1.19	.98	1.22	1.05	1.27	1.01
ITBS Problem Solving and Data Interpretation (28)	10.28	3.69	11.47	3.62	9.90	3.68	11.84	3.59

Note. BAU = business as usual; EAI = enhanced anchored instruction. Numbers in parentheses indicate number of points possible for subscales. Each of the measures has been aligned to the Common Core State Standards for Mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Space limitations prohibited inclusion of the table. Copies of the measures and the alignment can be obtained from the first author.

Problem Solving–Research Question 2

A significant effect was found in favor of EAI on the combined parts of the Problem Solving PS1-PS2 Test with the ES approaching moderate (.39). Much of the improvement can be traced to students' performance on PS2 that assessed ratios and proportional relationships (ES = .61) and geometry (ES =.57). No statistically significant difference was found on the ITBS Problem Solving and Data Interpretation Test.

To help understand the nature of student performances on the Problem Solving PS1-PS2 Test, two examples are provided. One test item addressing the

TABLE 3

HLM Results Estimating Effects of Treatment Condition on Fractions Computation, Problem Solving (PS1-PS2), ITBS Computation, and ITBS Problem-Solving and Data Interpretation Tests Adjusted for Respective Pretest Scores, Student Characteristics at Level 1, and Teacher Characteristics at Level 2

Outcomes	γ_{010}	SE	ES	R^2
Fractions Computation Test (n = 319)	12.55***	2.10	1.00	.38
Addition				
Overall	9.25***	1.49	1.00	.39
Simple Fractions	5.44***	.80	1.00	.39
Mixed Numbers	3.71***	.76	.90	.31
Subtraction				
Overall	3.22***	.59	.85	.28
Simple Fractions	1.32***	.22	.87	.26
Mixed Numbers	1.97***	.43	.77	.25
Denominator				
Like	2.31***	.48	.73	.26
Unlike	10.11***	1.73	.95	.36
Rewriting				
No rewrite	6.41***	1.09	.99	.39
Rewrite	6.05***	1.01	.96	.35
Problem Solving Test (PS1-PS2) (n = 324)	1.76*	.79	.39	.37
Ratios and Proportional Relationships	1.56**	.51	.61	.22
Measurement and Data	.56	.52	.16	.30
Number and Operations–Fractions	.34	.36	.15	.20
Geometry (Graphing)	1.53**	.46	.57	.24
ITBS Computation (n = 317)	2.22*	.82	.44	.24
Fractions				
Addition	.70**	.17	.60	.22
Subtraction	.45**	.13	.41	.14
Whole Numbers/Decimals				
Addition	.56*	.24	.40	.19
Subtraction	.41	.26	.26	.20
Multiplication	.38	.22	.25	.19
Division	.09	.11	.09	.06
ITBS Problem Solving and Data Interpretation (n = 327)	–.15	.57	.04	.15

Note. Because HLM does not allow any missing value on outcome variable, data analysis is based on a different number of students (indicated by n) for a cognitive measure who have valid scores on the posttest. R^2 = estimated proportion of overall variance in outcome scores explained by the model including pretest scores and student characteristics at Level 1 as well as treatment condition and teacher characteristics at Level 2.

$$ES = \frac{\gamma_{010}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1+n_2-2)}}}$$

, where γ_{010} represents the difference in posttest scores between EAI and BAU groups adjusted for pretest scores on the respective outcome, student characteristics at Level 1, and teacher characteristics at Level 2; n_1 and n_2 are EAI and BAU sample sizes, respectively; S_1 and S_2 are the student-level unadjusted posttest standard deviations for EAI and BAU groups, respectively.

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

ratios and proportional relationships standard asked students to identify relevant information from a table and then to apply proportional operations to obtain the answer. The percentage of EAI and BAU students answering this item correctly on the pretest was only 15% and 13%, respectively. At posttest, the proportion of EAI students who supplied the correct answer rose to 44% compared to 26% for BAU students. One of the most difficult items required students to interpret a table, plot the coordinates, and draw a line of best fit. Over half of students in the EAI group (53%) received full credit on the posttest compared to 41% of students in the BAU group.

Portions of Table 3 are devoted to measures of the adequacy of the HLMs by means of R^2 estimates. The range of R^2 estimates for HLMs that demonstrated statistically significant treatment effects was from .14 to .39. In this range, two HLMs had R^2 estimates .14 and .19, eight HLMs had R^2 estimates between .20 and .30, and eight HLMs had R^2 estimates above .30. Using Gaur and Gaur's (2006) standard, all of the HLMs were adequate in capturing variance in posttest scores.

DISCUSSION

The purpose of this study was to test the effects of blending explicit and anchored instructional strategies on the math performances of middle school students whose math skills were so low that they received their instruction in special education resource rooms. About one third of students in the EAI group had a cognitive disability and many of them made substantive progress in learning relatively complex concepts. Results showed that the students taught with EAI outscored their BAU counterparts on three of four math measures. On the curriculum-based Fractions Computation Test, EAI students made significantly larger gains than BAU students on total and subscale scores (i.e., adding and subtracting simple fractions/mixed numbers with like/unlike denominators, rewriting answers). EAI students also outscored BAU students on the standardized ITBS Computation Test with much of the difference coming from improved performance in adding and subtracting fractions.

Positive effects were also found in favor of EAI students on the curriculum-based problem solving measure. Test scores showed modest improvement among BAU and EAI students on the items that assessed the CCSS-M standards Measurement & Data and Number & Operations, but EAI students had statistically significant higher scores than BAU students on Proportional Relationships (i.e., ratio and rate) and Geometry (i.e., graphing a line of best fit). These two standards were addressed mainly in the KK and GP units. The complexity of operations for computing rate, scaling the *x*- and *y*-axes, plotting data points, drawing a line of best fit, and identifying coordinates of points demanded a sophisticated level of problem interpretation and problem solving.

Previous versions of this intervention resulted in similar improvements on the problem solving measures (e.g., Bottge et al., 2010), but this is the first time that students made such large gains in fraction computation. In earlier studies, fraction instruction was incorporated into the anchored problem solving units, which turned out to be unsuccessful with low-achieving students. In this study, EAI teachers taught computation skills in a direct way using a separate unit (FAW) prior to students using them with the anchored problems. Helping students develop a sound conceptual understanding of fractions, not simply procedural, was especially important. Of course, facilitating deep conceptual understanding while building procedural skills is not a new idea. Almost a century ago, Thorndike (1922) made the observation that mathematics pedagogy made artificial distinctions between computation and reasoning, and that it "separated too sharply the 'understanding of principles' by reasoning from the 'mechanical' work of computation . . . remembering facts and the like, done by 'mere' habit and memory" (p. 190). In short, students need to know both the how and why of computing.

In short, students need to know both the how and why of computing.

Based on these test results and more than 300 classroom observations, we offer several explanations for these findings. First, the hands-on problems immersed students directly in content-rich, relatively complex, and stimulating learning activities. That students attend to the learning activities they find most meaningful has been a consistent theme in education (Bruner, 1960; Dewey, 1938). In their view, good problems are not determined so much by appearance but rather on their capacity to evoke genuine student interest. In contrast to the EAI approach, the usual method of helping low-achieving students solve word problems is to teach them how to identify keywords that trigger number operations. This approach imitates drill-and-practice, builds routine procedures, and can actually discourage understanding. Lave (1993) described traditional word problems as a school activity having "no intuitive connections with everyday experience" (p. 89).

Second, the design of the EAI curriculum included sufficient cognitive supports. The need for providing these supports is based on cognitive load theory (Chandler & Sweller, 1991; Sweller, 1988), which in its simplest form suggests all learners have limited working memory and learning tasks should be structured so as not to overload it. One important consideration relates to what is called *element interactivity*, the extent to which relevant elements of content embedded in curricular materials interact. If constructed properly, adding story contexts, visual representations, and multimedia applications to materials at this level of complexity can be important for reducing cognitive overload (Wouters, Paas, & van Merriënboer, 2008).

Finally, a striking difference between EAI and BAU was instructional pacing. Although the BAU teachers were comparable to EAI teachers in managing student behaviors and teaching lessons, they often moved from one concept or skill to another in a span of 2 or 3 days or, in some cases, within a single class period. This practice appeared to fragment the curriculum and seemed especially confusing to students who needed more time to grasp the skill or concept. The EAI teachers stayed on one set of concepts for extended times as prescribed in the lesson plans. This afforded the EAI students opportunities to gain a more complete understanding of math concepts before moving on to the next unit. EAI students seemed to understand that their teachers were giving them the time necessary to learn the skill or concept.

Limitations

The results favored the use of EAI for improving both computation and problem solving skills, but several factors temper our findings. First, the duration of the study spanned 18 weeks, which is 6 weeks longer than is suggested by some best evidence syntheses (e.g., Slavin, 2008). However, the length of the intervention was relatively brief, especially considering the severity of the students' disabilities. The research design did not allow teachers enough time for practice that some of these students required. Nevertheless, teachers in the EAI group were impressed with what their students had learned in such a relatively short time.

Second, the paper-based tests may not have adequately captured the problem solving skills of these students. Recent reviews (e.g., Gijbels, Dochy, Van den Bossche, & Segers, 2005) suggest that the method of assessment carries with it important implications about how to judge the effectiveness of problem solving instruction. Classroom observers described instances of rich problem solving efforts by EAI students, but they suggested that some of the test items fell short of capturing them. Computer-based assessments may be useful in tracing students' movements through an anchored problem (e.g., Bottge, Rueda, Kwon, Grant, & LaRoque, 2009).

Third, a maintenance measure was not administered. However, months after conclusion of the study some of the EAI teachers emailed unsolicited descriptions of what their students had retained. For example, one teacher wrote how students scored on tests the next fall semester:

> The dept head of the sp ed dept here at [high school] just shared some test scores with me. Of the five students that I had in the resource math class last year at [middle school], 1 had a distinguished score, 3 had proficient and 2 were novice. 1 of the two novices entered into the class mid-year. Just thought I would share with you.

CONCLUSIONS

These findings suggest that some of the lowest achieving students in schools can make important gains in their math skills when explicit instruction and anchored problems are used appropriately in tandem to build both procedural and problem solving skills. One of the five attributes associated with math proficiency is productive disposition (NRC, 2001), an attitude that develops when instruction conveys to students that math is worthwhile. To stimulate this feeling, instruction should help them understand the connection between the math skills they are learning and the reasons for learning them. We suspect that the unique characteristics of the blended curriculum helped many of the students in the EAI resource rooms make this connection. A study now underway will investigate whether this holds true for students with disabilities taught in inclusive settings.

REFERENCES

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education*, *38*, 115–131.

Bottge, B. A. (2001). Reconceptualizing math problem solving for low-achieving students. *Remedial and Special Education*, *22*, 102–112. doi:10.1177/074193250102200505

Bottge, B. A., Rueda, E., Grant, T. S., Stephens, A. C., & LaRoque, P. T. (2010). Anchoring problem-solving and computation instruction in context-rich learning environments. *Exceptional Children*, *76*, 417–437.

Bottge, B. A., Rueda, E., Kwon, J. M., Grant, T., & LaRoque, P. (2009). Assessing and tracking students' problem solving performances in anchored learning environments. *Education Technology Research & Development*, *57*, 529–552. doi:10.1007/511423-007-9069-y

Briars, D., Asturias, H., Foster, D., & Gale, M. (2012). *Common Core Mathematics in a PLC at Work, Grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.

Brookings Institution. (2004). *The 2004 Brown Center Report on American Education: How well are American Students Learning?* Washington, DC: Author. Retrieved from http://www.brookings.edu/reports/2004/11education_loveless.aspx

Bruner, J. S. (1960). *The process of education*. New York, NY: Random House.

Butler, F. M., Miller, S. P., Crehan, K., Babbitt, B., & Pierce, T. (2003). Fraction instruction for students with mathematics disabilities: Comparing two teaching sequences. *Learning Disabilities Research & Practice*, *18*, 99–111. doi:10.1111/1540-5826.00066

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332. doi:10.1207/s1532690xci0804_2

Cho, S.- J., Bottge, B. A., Cohen, A. S., & Kim, S.- H. (2011). Detecting cognitive change in the math skills of low-achieving adolescents. *Journal of Special Education*, *45*, 67–76. doi:10.1177/0022466909351579

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, NJ: Erlbaum.

Dewey, J. (1938). *Experience and education*. New York, NY: Macmillan.

Gallico, R., Burns, T. J., & Grob, C. S. (1991). *Emotional and behavioral problems in children with learning disabilities*. San Diego, CA: Singular.

Gaur, A. S., & Gaur, S. S. (2006). *Statistical methods for practice and research: A guide to data analysis using SPSS*. Thousand Oaks, CA: SAGE.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*, 4–15. doi:10.1177/00222194040370010201

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*, 1202–1242. doi:10.3102/0034654309334431

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, *75*, 27–61. doi:10.3102/00346543075001027

Gillespie, J., & Kanter, P. F. (2000). *Every day counts calendar math*. Wilmington, MA: Great Source Education.

Kauffman, J. M. (2001). *Characteristics of emotional and behavioral disorders of children and youth* (7th ed.). Upper Saddle River, NJ: Merrill, Prentice-Hall.

Knopik, V.S., Alarcón, M., & DeFries, J.C. (1997). Comorbidity of mathematics and reading deficits: Evidence for a genetic etiology. *Behavior Genetics*, *27*, 447–453. doi:10.1023/A:1025622400239

Lave, J. (1993). Word problems: A microcosm of theories of learning. In P. Light & G. Butterworth (Eds.), *Context and cognition: Ways of learning and knowing* (pp. 74–92). Mahwah, NJ: Erlbaum.

Learning and Technology Center at Vanderbilt University. (1997). *The new adventures of Jasper Woodbury*. Mahwah, NJ: Erlbaum.

Misquitta, R. (2011). A review of the literature: Fraction instruction for struggling learners in mathematics. *Learning Disabilities Research & Practice*, *26*, 109–119. doi:10.1111/j.1540-5826.2011.00330.x

Montague, M., Bos, C., & Doucette, M. (1991). Affective, cognitive, and metacognitive attributes of eighth-grade mathematical problem solvers. *Learning Disabilities Research & Practice*, *6*, 145–151.

National Center for Education Statistics. (2011). *The Nation's Report Card Mathematics 2011* (NCES 2012-458). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors. Retrieved from http://www.corestandards.org/

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. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Parmar, R. S., Cawley, J. F., & Frazita, R.R. (1996). Word problem-solving by students with and without mild disabilities. *Exceptional Children*, *62*, 415–429.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., & McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning* (NCER 2007–2004). Washington, DC: U.S. Department of Education.

Raudenbush, S. W. (2008). Advancing educational policy by advancing research on instruction. *American Educational Research Journal*, *45*, 206–230. doi:10.3102/0002831207312905

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: Sage.

Scheuermann, A. M., Deshler, D. D., & Schumaker, J. B. (2009). The effects of the explicit inquiry routine on the performance of students with learning disabilities on one-variable equations. *Learning Disability Quarterly*, *32*, 103–120.

Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J. (2010). *Developing effective fractions instruction for kindergarten through 8th grade: A practice guide* (NCEE #2010-4039). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/practiceguide.aspx?sid=15

Silver, E. A., & Kenney, P. A. (2000). *Results from the seventh mathematics assessment of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

Slavin, R. E. (2008). What works? Issues in synthesizing educational program evaluations. *Educational Researcher*, *37*, 5–14. doi:10.3102/0013189x08314117

Sloane, F. C. (2008). Randomized trials in mathematics education: Recalibrating the proposed high watermark. *Educational Researcher*, *37*(9), 624–630. doi:10.3102/0013189x08328879

Spybrook, J., Raudenbush, S. W., Liu, X., & Congdon, R. (2006). *Optimal design for longitudinal and multilevel research: Documentation for the "Optimal Design" software*. Ann Arbor: University of Michigan.

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *100*, 343–379. doi:10.1037/0022-0063.100.2.343

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Thorndike, E. L. (1922). *Psychology of arithmetic*. New York, NY: Macmillan.

University of Iowa. (2008). *The Iowa Tests of Basic Skills* (ITBS, Form C). Itasca, IL: Riverside Publishing.

What Works Clearinghouse. (2011). *WWC procedures and standards handbook* (Version 2.1). Retrieved from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19

Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*, 645–675. doi:10.3102/0034654308320320

ABOUT THE AUTHORS

BRIAN A. BOTTGE, Early Childhood, Special Education, & Rehabilitation Counseling, University of Kentucky. **XIN MA**, Curriculum and Instruction, University of Kentucky. **LINDA GASSAWAY**, Early Childhood, Special Education, & Rehabilitation Counseling, University of Kentucky. **MICHAEL D. TOLAND**, Educational and Counseling Psychology, University of Kentucky. **MARK BUTLER**, Early Childhood, Special Education, & Rehabilitation Counseling, University of Kentucky. **SUN-JOO CHO**, Department of Psychology and Human Development, Peabody College of Vanderbilt University.

Address correspondence concerning this article to Brian A. Bottge, Professor, Early Childhood, Special Education, & Rehabilitation Counseling, University of Kentucky, 222 Taylor Education Building, Lexington, KY 40506 (e-mail: Bbott2@uky.edu).

The research reported in this article was supported by a grant from the U.S. Department of Education, Institute of Education Sciences (PR Number H324A090179). Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the views of the supporting agency.

Manuscript received November 2012; accepted February 2013.