Article

Exceptional Children
2017, Vol. 83(4) 340–358

DOI: 10.1177/0014402917691017
journals.sagepub.com/home/ecx
SAGE

Investigating a Tier 1 Intervention Focused on Proportional Reasoning: A Follow-Up Study

Asha K. Jitendra[1], Michael R. Harwell[1], Stacy R. Karl[1], Gregory R. Simonson[1], and Susan C. Slater[1]

Abstract

This randomized controlled study investigated the efficacy of a Tier 1 intervention—schema-based instruction—designed to help students with and without mathematics difficulties (MD) develop proportional reasoning. Twenty seventh-grade teachers/classrooms were randomly assigned to a treatment condition (schema-based instruction) or control condition (business as usual). Participants included 373 students, of whom 253 demonstrated MD. A measure of proportional problem solving (PPS) was administered at pre- and posttesting and at 11 weeks following treatment, along with a general mathematical problem-solving measure at pre- and posttesting. For the full sample, posttest differences favoring the treatment group were statistically significant for all measures. For students with MD, posttest differences favored the treatment group for the PPS posttest and PPS delayed posttest but not for general problem-solving posttest.

The present Tier 1 study represents a follow-up and extension of a study conducted by Jitendra et al. (2015) that provided evidence of the positive impact of a research-based intervention—schema-based instruction (SBI)—in improving middle school students' proportional reasoning. The present study resembles a replication, with its focus on consistency of findings, but sample size limitations suggest that the study represents a follow-up to Jitendra et al. (2015) that permits limited generalizability rather than a traditional replication.

Several authors have pointed out that replications are a staple in methods textbooks and are badly needed (Makel & Plucker, 2014) but rarely appear in practice (Duncan, Engel, Claessens, & Dowsett, 2015). The argument for replication in educational research is simple: Replication is central to understanding and improving scientific research in that it decreases research bias by providing evidence supporting (validating) positive findings or invalidating false-positive findings (Cook, 2014; Coyne, Cook, & Therrien, 2016; Gottfredson et al., 2015; Schmidt, 2009; Valentine et al., 2011).

We conducted a follow-up study of Jitendra et al. (2015), whose features were consistent with a conceptual replication (see Doabler et al., 2016), which alters "specific aspects of a previous study" (Cook, 2014, p. 234) to determine the extent to which findings generalize across participants, settings, and conditions (Schmidt, 2009). The follow-up study used a smaller but more demographically and geographically diverse sample. Jitendra et al. (2015) investigated the efficacy of the SBI intervention in mostly rural schools in one state in the upper Midwest using a sample of students and teachers who were predominantly

[1]University of Minnesota

Corresponding Author:
Asha K. Jitendra, Department of Educational Psychology, University of Minnesota, Minneapolis, MN 55455.
E-mail: jiten001@umn.edu

White. In contrast, the follow-up study was conducted in urban and suburban schools from a metropolitan area in the Southeast with a sample that was more racially diverse and included larger percentages of English language learners (ELLs) and students from economically disadvantaged backgrounds. Students in the present study also received instruction on ratios and proportional relationships in late October, which was about 2 months earlier than the original study (early January).

Proportionality . . . involves the concept of ratio and is central to topics in mathematics, such as linear functions, scale drawings, similarity, trigonometry, and probability.

An overarching principle in research conducted within multitiered systems of support (MTSS) or a response-to-intervention framework is that a Tier 1 intervention must be effective for students with mathematics disabilities/difficulties (MD) as well as for their peers without MD because students with MD often receive instruction in traditional mathematics classrooms (Fuchs, Fuchs, & Compton, 2012). This focus is consistent with national educational policy embodied in the Common Core State Standards in mathematics (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA & CCSSO], 2010) requiring all students to connect the standards of mathematical practice (e.g., making sense of problems, reasoning abstractly and quantitatively, looking for and making use of structure, modeling with mathematics, attending to precision) to the standards for mathematical content, in order to develop proficiency in mathematics. Meeting these standards creates important challenges for students with MD and those with persistently low achievement in mathematics. For example, the 2015 National Assessment of Educational Progress scores indicated that 68% of eighth-grade students with disabilities performed below basic, as compared with 23% of students without disabilities in mathematics (National Center for Education Statistics, n.d.). Thus, an important focus of research, including ours, is to identify programs that improve these students' mathematics achievement.

One mathematics topic of particular importance in the middle grades is proportionality, which involves the concept of ratio and is central to topics in mathematics, such as linear functions, scale drawings, similarity, trigonometry, and probability. Ratio and proportional relationships—along with the interrelated topics of fractions, decimals, and percent—provide an essential foundation for algebra (National Mathematics Advisory Panel, 2008). Students in the middle grades are expected to "develop understanding of proportionality to solve single and multi-step problems . . . solve a wide variety of percent problems, including those involving discounts, interest, taxes, tips, and percent increase or decrease" (NGA & CCSSO, 2010, p. 46). Proportional reasoning is challenging for many children and adolescents; it requires not only understanding of the concept of ratios and that two or more ratios are equal but also the ability to extract relevant information to develop a representation of the problem situation (Ozgun-Koca & Altay, 2009). As such, it seems reasonable to provide effective core (Tier 1) classroom instruction to all students before providing more intensive (Tier 2 and Tier 3) interventions to students who do not respond to classroom instruction (see Center on Response to Intervention, http://www.rti4success.org/). Determining whether students who do not respond to Tier 1 instruction need Tier 2 services or just better Tier 1 instruction is an important challenge (Fuchs, Fuchs, & Vaughn, 2008).

Tier 1 Instruction in Proportional Reasoning

In four randomized controlled Tier 1 studies, Jitendra et al. (2009; Jitendra et al., 2015; Jitendra, Star, Dupuis, & Rodriguez, 2013; Jitendra, Star, Rodriguez, Lindell, & Someki, 2011) examined the effectiveness of the SBI intervention in improving students' proportional reasoning, three of which reported results separately for students with and without MD. Two of the three studies found evidence of effective proportional problem-solving (PPS) instruction for students without MD that was also effective for students with MD.

In Jitendra et al. (2009) and Jitendra and Star (2012), eight seventh-grade classrooms were randomly assigned to SBI or a "business as usual" control condition and received instruction five times a week for 45 min over a 2-week period on the topics of ratio and proportion or percent. Jitendra et al. (2009) found that students in the SBI condition, on average, scored higher than students in the control condition on the PPS test and maintained the effects on a 4-month retention test. However, the results also showed that scores of low-achieving students in the SBI group were comparable to those of low-achieving students in the control group. Similar findings were reported by Jitendra and Star (2012). In a follow-up study, Jitendra et al. (2013; see also Jitendra, Dupuis, Star, & Rodriguez, 2016) included more classrooms and found that posttest and retention test (6 weeks later) differences favoring the SBI group were significantly different for students with and without MD. On a transfer measure that included items not directly aligned with the taught content, there were no intervention effects.

A recent randomized study (Jitendra et al., 2015; see also Jitendra, Harwell, Dupuis, & Karl, in press) made methodological improvements by randomly assigning 82 teachers to the SBI and control conditions and then randomly selecting one classroom per teacher to participate in the study, meaning that each teacher taught a single SBI or control classroom. Findings comparing the SBI and control conditions in Jitendra et al. (2015) were reported for all students (i.e., students with and without MD) and for MD students only (Jitendra et al., in press). In both instances, SBI classrooms outperformed control classrooms on the posttest and maintained the effects 9 weeks following the end of the study. Multilevel standardized effect sizes (ESs) for all students and for students with MD on the posttest were 0.46 and 0.32; on the retention test administered 8 weeks later, the ESs were 0.32 and 0.25. Results for the full sample indicated that students in SBI classrooms showed more growth on PPS than students in control classrooms; this finding was not statistically significant for students with MD. There were no intervention effects on a standardized test that assessed overall mathematical problem solving involving multiple content areas (e.g., algebra, geometry).

The Present Study

The present study focused on increasing confidence and generalizability of findings to support the impact of the SBI intervention when implemented in a geographically diverse location with a more demographically diverse sample. The research questions were as follows:

Research Question 1a: What are the effects of the SBI intervention as compared with business-as-usual instruction on students' PPS at immediate posttest and after 11 weeks?

Research Question 1b: Does SBI moderate students' learning trajectory over time?

Research Question 1c: What are the effects of the SBI intervention as compared with business-as-usual instruction on students' general mathematical problem solving?

Research Question 2a: What are the effects of the SBI intervention as compared with business-as-usual instruction for students with MD on PPS at immediate posttest and after 11 weeks?

Research Question 2b: Does SBI moderate the learning trajectory of students with MD?

Research Question 2c: What are the effects of the SBI intervention as compared with business-as-usual instruction for students with MD on general mathematical problem solving?

Method

Following Institutional Review Board approval, the proposed research study was presented to middle school principals, curriculum specialists, and math coaches at a meeting in a Southeast U.S. district in the spring prior to the academic year in which the study occurred. Middle school math teachers received information about the study via presentations and a newsletter from the principal and the director of secondary curriculum and instruction. To meet study inclusion criteria, teachers were required to be responsible for teaching at least one seventh-grade math class, agree to randomization, administer mathematics measures to assess student learning, and allow videotaping to assess fidelity of treatment implementation. Once teachers provided written consent, one class for each teacher was selected at random to participate in the study. The 22 teachers who consented and their participating classes were randomly assigned to the SBI condition or a control condition that received business-as-usual instruction. Following random assignment, two teachers assigned to the control group were eliminated from the study because they did not teach a typical seventh-grade math class (i.e., advanced seventh-grade content), resulting in unequal numbers of teachers in the treatment (j = 11) and control (j = 9) conditions.

School Context and Participants

The 10 middle schools participating in the study were part of a district in a metropolitan area with an approximate student enrollment of 193,000. Participating schools ranged in size from 907 to 1,807 students. Most schools in the sample (90%) were located in urban settings, whereas one school was located in a suburban setting.

Students. The initial student sample consisted of 429 seventh graders but was reduced to 373 (87% of the original sample) because 56 students had missing demographic data. The remaining 373 students constituted the sample used in our analyses and represented a more diverse sample than that used in prior evaluations of SBI (Jitendra et al., 2015). Among students, the racial group sample sizes were, in order of representation, Hispanic (143 of 373 = 38.3%), Black (31.4%), White (23.6%), Asian (3.8%), and multiracial (2.9%). The mean age of the students was 12.7 years (SD = 0.55). Approximately 73% of the sample was eligible for a free or reduced-price lunch (FRL); 6.2% received special education services; and 23.9% were ELLs.

In addition, 253 students (67.8% of the sample) were classified as having MD based on end-of-the-year sixth-grade state mathematics achievement assessment. Precise reasons for the large number of students with MD in the sample are unknown but may be related to characteristics of the state mathematics test used to categorize students as having MD, as well as the fact that all sampled schools received Title I assistance (see Bottge et al., 2014). Achievement levels on this test were 1 (*inadequate*), 2 (*below satisfactory*), 3 (*satisfactory*), 4 (*above satisfactory*), or 5 (*mastery*). Students whose achievement level was inadequate (n = 117, 32.2%) or below satisfactory (n = 136, 37.5%) met the criterion for mathematical difficulties (i.e., low average performance on the state math achievement test) and were classified as having MD and those scoring ≥3 as without MD.

Teachers. The 20 participating seventh-grade mathematics teachers' mean years of experience was 7.05 (SD = 5.42; range, 2–19 years). All teachers were certified to teach mathematics; 5% were also certified to teach science, 20% were certified in subjects other than mathematics or science, and 25% were certified in all subjects (generalist). Similarly, all teachers were certified to teach Grades 6–8, 35% to teach Grades 9–12, and 35% to teach Grades K–5. The majority of teachers were White (70%), with three Black, two Asian, and one Hispanic teacher. Seventy percent of teachers were female. One teacher (5%) taught in a large suburban setting and the remaining teachers (95%) in urban settings. Student and teacher demographic characteristics in the SBI and control groups are presented in Table 1.

Study Design

We used a prospective randomized cluster design with longitudinal data (pretest, posttest, delayed posttest) in which teachers/classrooms served as clusters. If properly implemented, this design ensures that estimated treatment effects are unbiased (Bloom, Richburg-Hayes, & Black, 2007). An a priori power and sample size analysis based on the Optimal Design software (Version 3.0; Spybrook et al., 2011) was performed that focused on testing the SBI-versus-control effect for cross-sectional data needed to answer the research questions. The results indicated that 20 clusters, 373 students, and an intraclass correlation (ICC) of .15 (Hedges & Hedberg, 2007) provide a power of 0.74 to detect a standardized effect of 0.55 (a moderate effect following Cohen, 1988) for the SBI-versus-control comparison. This ES is consistent with that reported by Jitendra et al. (2015).

Procedures

Teachers implemented two instructional units (ratio and proportion; percent) in their treatment and control classes using the assigned curricular program in late October, after the pretest had been administered. The two units were implemented 45 min daily for 6 weeks. Control teachers used their district-mandated textbook and typical practices to teach the content (ratio and proportion; percent), whereas treatment teachers used the SBI curriculum to teach the same content over the same period.

Treatment teachers received 2 days (16 hr) of professional development training prior to the start of the study in early October. The training was conducted by a researcher outside the research team who was an expert in problem solving and professional development training and was familiar with the SBI intervention. The primary purpose of the training was to familiarize teachers with the SBI curriculum by explaining the background and principles of the SBI program and discussing all materials as well as the planning and organization of the lessons. The training also covered SBI practices (e.g., recognizing problem types, generating estimates, applying multiple-solution strategies) used to support student learning of ratio, proportion, and percent. There was a significant focus on SBI implementation, which involved not only reading the detailed scripts on when and how to implement SBI practices but also viewing multiple short video clips in which other teachers demonstrated the application of the SBI practices in their classrooms.

Description of the SBI intervention. Several features underlie the SBI intervention, including recognizing the mathematical structure of problems, using visual representations to model the problem, applying problem-solving procedures for a given class of problems, and developing procedural flexibility. Also integral to SBI is monitoring and reflecting on the problem-solving process. These practices parallel the mathematical practices in the CCSS (e.g., look for and make use of structure, model with mathematics) and are consistent with the recommendations articulated in the What Works Clearinghouse research synthesis on improving students' mathematical problem-solving performance (Woodward et al., 2012). In addition, the SBI intervention incorporates several instructional features that are important to promoting problem solving for students with MD, such as explicit and systematic instruction and scaffolding instruction with guided questions to help clarify and refine student thinking (Gersten, Beckmann, et al., 2009; Gersten, Chard, et al., 2009).

The SBI program materials include a detailed teacher guide and other resources (e.g., visual diagrams, problem-solving checklists, homework answer key) to support teachers in implementing activities to develop important concepts and skills, as well as student materials (i.e., workbook and homework book). The program comprises 21 lessons that can be completed in about 30 days (some lessons take more than a day to implement) with the first unit focusing on ratio/proportion (e.g., equivalent ratios, rates, ratio and proportion word problem solving, scale drawings)

Table 1. Participant Demographic Information by Treatment.

	SBI		Control	
	n	%	*n*	%
Student information: Full sample				
Age, years, *M* (*SD*)	12.73	(0.55)	12.69	(0.55)
Sex				
Female	116	53.7	75	47.8
Male	100	46.3	82	52.2
Race				
Asian	4	1.9	10	6.4
Black	73	33.8	44	28.0
Hispanic	78	36.1	65	41.4
Multiracial	8	3.7	3	1.9
White	53	24.5	35	22.3
FRL				
Yes	157	72.7	114	72.6
No	59	27.3	43	27.4
ELL				
Yes	50	23.1	39	24.8
No	166	76.9	118	75.2
SpEd				
Yes	17	7.9	6	3.8
No	199	92.1	151	96.2
Students with MD				
Age, years, *M (SD)*	12.74	(0.56)	12.79	(0.59)
Sex				
Female	85	54.1	46	47.9
Male	72	45.9	50	52.1
Race				
Asian	2	1.3	2	2.1
Black	60	38.2	34	35.4
Hispanic	57	36.3	38	39.6
Multiracial	7	4.5	2	2.1
White	31	19.7	20	20.8
FRL				
Yes	120	76.4	78	81.2
No	37	23.6	18	18.8
ELL				
Yes	42	26.8	29	30.2
No	115	73.2	67	69.8
SpEd				
Yes	15	9.6	6	6.2
No	142	90.4	90	93.8
Teacher information				
Sex				
Female	7	63.6	7	77.8
Male	4	36.4	2	22.2
Race				
Asian	2	18.2	0	0.0
Black	2	18.2	1	11.1
Hispanic	0	0.0	1	11.1
White	7	63.6	7	77.8
Experience,[a] *M* (*SD*)	6.45	(5.28)	7.78	(5.83)
Math courses, *M* (*SD*)	4.45	(3.98)	7.00	(4.84)

Note. Total students, *N* = 373. SBI = schema-based instruction; FRL = students eligible for free or reduced-price lunch; ELL = English language learner; SpEd = students receiving special education services; MD = mathematics difficulties.
[a]Years experience teaching math.

and the second unit on percent (e.g., percent increase or decrease—including those involving discounts, interest, taxes, tips). Instruction included both whole class instruction as well as partner work with a think/plan-share strategy to enhance students' critical thinking. Each unit includes nine lessons and a culminating lesson that evaluates student knowledge of material covered. These lessons present real-world scenarios (i.e., designing a recording studio, constructing a digital planetarium) and require students to work in small groups to solve several problems involving ratios and proportional relationships. Lesson 21 culminated with a review of the content from both units. Mid- and end-of-unit tests are used to check content understanding and evaluate student progress in the SBI program.

The SBI intervention consists of four instructional practices: problem structure identification, problem representation through schematic diagrams, procedural flexibility, and problem solving and metacognitive strategy knowledge application. Problem solving and metacognitive strategy knowledge application allow students to apply learned content (e.g., ratios/rates, percent) in problem-solving activities: (a) recognize the problem type, (b) connect the problem to what is already known, (c) identify and represent crucial information in the problem through an appropriate diagram, (d) estimate the answer, (e) select a strategy to solve the problem, (f) solve and present the solution within the context of the problem, and (g) check the reasonableness of the solution. Further, metacognitive activities include monitoring and reflecting on the problem-solving process (e.g., when, how, and why to use multiple strategies—equivalent fractions, unit rate, cross multiplication—for a given class of problems). For further details of the SBI program, see Jitendra et al. (2015; see also Jitendra et al., 2009; Jitendra et al., 2011; Jitendra et al., 2013). Instruction in applying SBI practices to support student learning was scaffolded such that teachers initially modeled problem solving by thinking aloud and gradually shifted responsibility to the students, with teachers using explicit prompts to help clarify and refine student thinking.

"Business as usual" control instruction. Students in the control condition received instruction on the same topics and in the same period as the treatment condition. Control teachers used the district-adopted mathematics textbook (*Go Math! Florida*; Adams et al., 2011) to introduce the topics of ratio, proportion, and percent by using various tips to "motivate the lesson" (e.g., "Have you ever measured how far you can walk in 15 minutes?"), "focus on patterns" (e.g., use number patterns to complete a table and use the pattern to calculate distance), and "connect vocabulary" (rates and ratios). In general, instruction in the control classrooms differed on a number of important aspects from the SBI program. Although ratio, proportion, and percent lessons in the control textbook focused on a specific type of problem (e.g., rates, percent of change), opportunities were lacking for teachers to focus on identification of problem structure or emphasize how the problem was similar or different from a previously solved problem. The control textbook presented several visual representations (e.g., tables, graphs, bar diagrams). Worked examples in the textbook focused on modeling the problem-solving steps, and instructions were included to use questioning strategies (e.g., "What characteristic do you look for in the table in order to decide whether the relationship is proportional?"). However, the opportunities to apply these practices were limited to one or two examples. Within the context of problem solving, an emphasis on multiple solution and estimation strategies was nonexistent.

Measures

PPS and general mathematical problem-solving assessments were administered prior to (early October) and immediately following the intervention, with retention of PPS data collected 11 weeks following intervention. We also collected students' sixth-grade scores on a state-mandated mathematics assessment, as well as information on race, sex, special education, ELL status, and eligibility for FRL lunch. In addition, all teachers completed the Mathematical Knowledge for Teaching (MKT) test online prior to the start of the study to establish baseline estimates of teachers' knowledge for teaching proportional relationships. MKT was developed as part of the Learning Mathematics for Teaching Project at the University of Michigan to assess the specific mathematics content knowledge that teachers need to know (Hill, Ball, Blunk, Goffney, & Rowan, 2007). We used the Grade 4–8 proportional reasoning measure, which consists of two equivalent forms (A and B) of 15 and 16 multiple-choice and multistep questions, including applied word problems. The measure takes about 30 min to complete. We report item response theory–based scores for the MKT that take into account the relative difficulty of items.

PPS test. The PPS test (Jitendra et al., 2015) was used to assess PPS. The PPS has been used in previous studies and consists of 22 multiple-choice questions and four short-response items. Multiple-choice items were dichotomously scored (correct, incorrect), whereas short-answer items were scored 0 to 2, based on a rubric that emphasized correct reasoning, by trained research personnel who were unaware if a student was in treatment or control. We estimated average interrater reliability for the short-answer items using an ICC, which produced values of .90, .92, and .90 at pretest, posttest, and delayed posttest, respectively.

Student scores on the PPS were calculated by computing the sum of the total points earned (i.e., 30), meaning that the short-response and multiple-choice items contributed unequally to the overall score. We assessed the internal consistency of the PPS using the jMetrik software (Version 2.1.0; Meyer, 2011) and fit a congeneric model assuming that a single continuous latent factor underlies the dichotomously and trichotomously scored PPS items (McDonald, 1999). The coefficient omega values (Dunn, Baguley, & Brunsden, 2013) for the PPS pretest, posttest, and delayed posttest of .76, .82, and .80, respectively, represent reliabilities estimated as the ratio of true score variance to observed score variance (Dunn et al., 2013).

Group Mathematics Assessment and Diagnostic Evaluation. The Process and Applications subtest of the Group Mathematics Assessment and Diagnostic Evaluation (GMADE; Pearson Education, 2004) is a group-administered untimed assessment of general problem solving. This measure consists of 30 multiple-choice questions, including multiple-step problems and process problems, and it assesses students' ability to comprehend mathematical language and concepts and apply relevant operations to solve word problems across multiple content areas (e.g., algebra, geometry, number, and operations). Each question was awarded 1 point if correct. The coefficient omega reliabilities for our sample were .74 for the pretest and .75 for the posttest.

Fidelity of Implementation

We collected data on proportion problem-solving instruction in the treatment and control classes by videotaping each teacher's class and then coding the extent to which the SBI treatment was implemented with fidelity and whether SBI instructional practices were implemented in control classes (i.e., identifies the problem type, connects the new problem to previously solved problems, represents crucial information in the problem text through an appropriate diagram, generates an estimate prior to solving the problem, discusses multiple solution strategies, presents the solution within the context of the problem, checks the solution). Research team members utilized a rubric from prior studies (e.g., Jitendra et al., 2015) to rate the classroom instruction via a 0- to 3-point scale (0 = *did not implement*, 3 = *high level of implementation*). Fidelity for each classroom was independently assessed by two of the four coders, producing 40 ratings (two per classroom). Disagreements in ratings were resolved through discussion and review of the videotapes. We estimated interrater reliability by computing ICCs for the ratings, which averaged .94 (range, .87–.99).

We conducted *t* tests to test group differences on the fidelity of implementation data and used the Dunn-Bonferroni correction to control for compounding of Type I error. Results indicated statistically significant and fairly substantial differences between the treatment (M = 14.45, SD = 4.13) and control (M = 9.00, SD = 4.18) groups on the total fidelity score, $t(18)$ = 2.92, p = .009, ES = 1.31, and on Item 4 (i.e., generates an estimate prior to solving the problem), $t(18)$ = 3.92, p < .001, ES = 1.76, with treatment teachers (M = 1.91, SD = 1.14) implementing this SBI element more than control teachers (M = 0.22, SD = 0.67).

Data Analysis

To assess differences between treatment and control conditions, we fitted a series of multilevel models (i.e., two-level, students within classrooms) with covariates at both levels, using the HLM 6 software (Raudenbush, Bryk, & Congdon, 2004). Adjusting the outcomes with control variables can account for variation that may improve statistical power (Bloom et al., 2007). The outcome variables were the PPS posttest, PPS delayed posttest, and GMADE posttest, which were analyzed separately. We also performed an analysis of the PPS longitudinal data to explore student change over time and whether change was moderated by the treatment.

For each outcome, the Level 1 (student) model contained the following covariates: pretest score, sex (0 = male, 1 = female), race (Black = 1, Hispanic = 1, Asian = 1, multiracial = 1, and White = 0, so the latter served as the reference group), and math achievement level (inadequate = 1, below satisfactory = 2, satisfactory and above = 3; inadequate and below satisfactory math achievement levels were used to select students with MD). All Level 1 covariates were grand mean centered. Level 2 covariates included the treatment variable (1 = SBI, 0 = control), two teacher variables (years of teaching experience in mathematics and MKT proportional reasoning item response theory score), and variables capturing the percentage of the following: ELLs, students eligible for FRL, and students receiving special education services per classroom or teacher. We focused on classroom ELL, FRL, and special education because these variables sometimes showed little or no variation in a classroom at the student level. Because the classroom distributions of the percentages were ragged and discontinuous, we rescaled the ELL and FRL variables to quartiles and the special education variable into a trichotomy and used the rescaled versions as Level 2 covariates.

Slopes capturing the impact of student variables (e.g., sex) on the outcome variables did not vary significantly across classrooms (i.e., variance = 0). Thus, models examining the impact of treatment on Level 1 relationships were not fitted, and in what follows, the results are based on intercepts-only models. We also examined the data for evidence that model assumptions were satisfied, and we found no major violations. To control for compounding of Type I error rates, several methods are available (What Works Clearinghouse, 2014). We used the Dunn-Bonferroni correction, in which an overall experiment-wise Type I error rate (α = .15) is divided among all statistical tests linked to each outcome variable, with no requirement that the error rate be divided equally. Accordingly, we assigned .05 to the test of the treatment effect because this was the most important effect in the model and because this is consistent with our focus on identifying a program that can potentially improve the mathematics achievement of students. The remaining .10 was divided equally among tests of the remaining fixed effects, producing α = .10/13 = .008.

Impact of Missing Data

The percentages of missing data for the PPS pretest, posttest, delayed posttest, and GMADE pretest and posttest among the students (N = 373) were 2.5%, 4.2%, 9.7%, 3.9%, and 5.1%, respectively. The only other missing student data occurred for math achievement level (2.8%). Importantly, the percentage of missing data was approximately equal in the SBI and control conditions. However, the potential biasing effect of missing data (What Works Clearinghouse, 2014) prompted us to assess its impact.

Because students with missing data in the multilevel analyses were omitted by the HLM 6 software, it is important to explore the sensitivity of the findings to missing data. For a sensitivity analysis, we fitted each of the final multilevel models to the student sample that provided complete data (309 of 373, 83%) and compared the results to those obtained with all available data (some missing data). The results of these models did not differ in any significant way, suggesting that our models were relatively insensitive to the presence of missing data. As such, for the analyses, we used all available data for our sample of 373 students, meaning that student sample sizes varied across analyses. The sample of teachers produced complete data.

Results

We initially performed a series of descriptive analyses that included examining the correlations among all measures and exploring preexisting differences between the SBI and control students. Results of the correlational analyses showed that correlations between the PPS pretest and posttest, pretest and delayed posttest, and posttest and delayed posttest were .58, .57, and .67, respectively. The correlation between the GMADE pretest and posttest was .46, and correlations between the PPS and the GMADE tests ranged from .41 to .61 across time points.

Table 2 reports means and *SD*s for the treatment and control groups for each measure for the sample of 373 students. Differences between the treatment and control groups on the PPS and GMADE pretests were statistically significant ($d = -0.38$ *SD*, $d = -0.35$ *SD*). For the MD sample in Table 2, the treatment group scored lower than the control group on the PPS pretest ($d = -0.41$) but not the GMADE pretest. It is not clear what the source of this difference is, but the inclusion of the pretest variables as covariates in the multilevel analyses

Table 2. Descriptive Statistics for Measures by Treatment.

	SBI			Control			Total		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Full sample									
PPS pretest	210	9.08	4.13	154	10.74	4.72	364	9.78	4.46
PPS posttest	208	12.53	5.60	150	11.74	4.77	358	12.20	5.27
PPS delayed	197	11.99	5.00	143	11.96	4.95	340	11.98	4.97
GMADE pretest	207	10.26	4.16	152	11.72	4.28	359	10.87	4.27
GMADE posttest	207	12.28	4.45	148	12.43	4.37	355	12.34	4.41
Students with MD									
PPS pretest	153	7.80	3.51	94	9.21	3.31	247	8.34	3.50
PPS posttest	151	10.84	4.79	92	9.89	3.53	243	10.48	4.38
PPS delayed	143	10.49	4.20	88	9.99	3.55	231	10.30	3.97
GMADE pretest	149	9.49	3.80	93	10.43	3.62	242	9.85	3.76
GMADE posttest	150	11.15	3.77	91	10.81	3.88	241	11.02	3.81

Note. All test statistics are based on the total number of items correct. SBI = schema-based instruction; PPS = proportional problem solving; GMADE = Group Mathematics Assessment and Diagnostic Evaluation (Pearson Education, 2004); MD = mathematics difficulties.

means that outcomes will be adjusted for these differences.

Question 1: Tests of Treatment Effects on Outcomes for All Students

We examine the first research question, involving the effects of SBI for the following outcomes: PPS at immediate posttest and after 11 weeks, students' learning trajectory over time, and general mathematical problem solving. To estimate the ICC, we fitted unconditional two-level models (students-within teachers/classrooms) separately for the PPS posttest, PPS delayed posttest, and GMADE posttest. The ICC was .36 ($p < .001$) for the PPS posttest and .41 ($p < .001$) for the PPS delayed posttest, indicating that 36% and 41% of the variance of these variables were between classrooms. For the GMADE posttest, the ICC was .30. We then fitted a model with student background variables and a pretest at Level 1 and with teacher covariates plus the SBI treatment at Level 2, separately to each outcome variable.

Tables 3–5 present the multilevel results for the PPS posttest, PPS delayed posttest, and GMADE posttest outcomes. Table 3 shows that SBI was a significant predictor of PPS posttest and that SBI students on average scored 2.25 points higher than control students (conditional on the model). Equivalently, the SBI-versus-control difference was 0.63 *SD*, or 74% of SBI students scored above the mean of control

Table 3. Hierarchical Linear Modeling Results for PPS Posttest.

Fixed effects	Full sample					Students with MD				
	B	*SE*	*t*	*df*	*p*	*B*	*SE*	*t*	*df*	*p*
Between-student model										
Sex	0.40	0.397	1.01	326	.313	0.13	0.478	0.27	224	.784
Asian	–0.51	1.079	–0.47	326	.636	–0.57	1.884	–0.30	224	.763
Black	–0.64	0.568	–1.12	326	.264	–1.06	0.687	–1.54	224	.125
Hispanic	–0.84	0.526	–1.59	326	.112	–0.97	0.669	–1.45	224	.148
Multiracial	–0.25	1.163	–0.22	326	.830	0.29	1.305	0.23	224	.822
Math achievement level	1.88	0.296	6.33	326	<.001	1.69	0.490	3.44	224	.001
Pretest	0.45	0.055	8.21	326	<.001	0.39	0.076	5.15	224	<.001
Between-classroom model										
Intercept	12.78	0.641				11.24	0.880			
Treatment	2.25	0.444	5.07	13	<.001	1.82	0.573	3.17	12	.009
Years experience	–0.02	0.047	–0.47	13	.647	–0.02	0.065	–0.26	12	.802
MKT IRT estimate	1.37	0.424	3.23	13	.007	1.39	0.564	2.46	12	.030
Special education	–0.13	0.318	–0.39	13	.699	0.04	0.396	0.11	12	.916
English language learner	0.25	0.319	0.78	13	.449	0.08	0.413	0.18	12	.857
FRL	–1.35	0.405	–3.35	13	.006	–1.17	0.530	–2.20	12	.048
Random effects	VC	*SD*	χ^2	*df*	*p*	VC	*SD*	χ^2	*df*	*p*
Classroom	0.04	0.194	21.31	13	.067	0.20	0.451	21.32	12	.046
Student	12.81	3.579				12.44	3.527			

Note. Here α = .05 for the test of the SBI effect and α = .10/13 = .008 for tests of the remaining fixed effects. PPS = proportional problem solving; MD = mathematics difficulties; MKT = Mathematical Knowledge for Teaching; IRT = item response theory; FRL = eligible for free or reduced-price lunch; VC = variance component.

students. Teacher MKT score was also a significant predictor, as was FRL, with the latter interpreted to mean that each quartile increase in FRL rates was associated with a 1.35-point decline on the PPS posttest. Among student covariates, math achievement level was significant, as was the PPS pretest, with the former interpreted to mean that each one-unit increase in math achievement level was associated with an expected gain of 1.88 points on the posttest.

Table 4 shows that SBI was a significant predictor of the delayed posttest and that, on average, SBI students scored 1.14 points higher on the delayed PPS posttest than control students. Equivalently, the SBI-versus-control difference was 0.33 *SD*, or 63% of SBI students scored above the mean of control students. FRL was also a significant predictor. At the student level, the significant predictors were pretest and math achievement level.

To examine the impact of SBI on students' growth, we fitted a two-level model (repeated measures within students) to the PPS data, assuming heterogeneity of variance over time. Based on the exploratory nature of this analysis, the only predictor in this model beyond time was treatment. The average linear slope over time was significant ($B = 0.63$, $t = 3.41$, $p < .001$), meaning that student scores on average increased over time. The results also indicated that the treatment variable was a statistically significant predictor of student linear growth

Table 4. Hierarchical Linear Modeling Results for PPS Delayed Posttest.

Fixed effects	Full sample					Students with MD				
	B	*SE*	*t*	*df*	*p*	*B*	*SE*	*t*	*df*	*p*
Between-student model										
Sex	0.18	0.393	0.46	310	.644	0.03	0.456	0.07	213	.947
Asian	−0.22	1.058	−0.21	310	.835	−1.03	1.757	−0.58	213	.560
Black	−1.40	0.567	−2.47	310	.014	−1.67	0.648	−2.57	213	.011
Hispanic	−1.05	0.518	−2.02	310	.044	−1.37	0.626	−2.19	213	.029
Multiracial	−2.29	1.131	−2.03	310	.043	−2.87	1.217	−2.35	213	.020
Math achievement level	1.57	0.298	5.29	310	<.001	1.47	0.467	3.16	213	.002
Pretest	0.37	0.055	6.71	310	<.001	0.27	0.070	3.85	213	<.001
Between-classroom model										
Intercept	14.06	0.733				12.58	0.775			
Treatment	1.14	0.524	2.17	13	.049	1.17	0.511	2.29	12	.041
Years experience	−0.13	0.055	−2.29	13	.039	−0.12	0.057	−2.07	12	.061
MKT IRT estimate	0.83	0.503	1.65	13	.122	0.98	0.501	1.96	12	.074
Special education	−0.16	0.381	−0.41	13	.688	−0.19	0.346	−0.55	12	.595
English language learner	0.34	0.374	0.92	13	.377	0.38	0.363	1.04	12	.320
FRL	−1.51	0.472	−3.19	13	.008	−1.51	0.465	−3.26	12	.007
Random effects	VC	*SD*	χ^2	*df*	*p*	VC	*SD*	χ^2	*df*	*p*
Classroom	0.36	0.605	30.04	13	.005	0.02	0.127	17.69	12	.125
Student	11.88	3.446				10.91	3.303			

Note. Here α = .05 for the test of the SBI effect and α = .10/13 = .008 for tests of the remaining fixed effects. PPS = proportional problem solving; MD = mathematics difficulties; MKT = Mathematical Knowledge for Teaching; IRT = item response theory; FRL = eligible for free or reduced-price lunch; VC = variance component.

(B = 0.73, t = 3.02, p = .003), meaning that SBI students improved at a faster rate than control students.

For the GMADE posttest, Table 5 shows that SBI was a significant predictor, with an SBI-versus-control difference of 0.32 *SD*, or 63% of SBI students scored above the mean of control students. The only significant student predictors were pretest and math achievement level.

Question 2: Tests of Treatment Effects on Outcomes for Students With MD

We also conducted subgroup analyses that focused on the effects of SBI for a sample of students with MD. Preliminary analyses of the PPS posttest, PPS delayed posted, and GMADE posttest outcomes separately produced ICCs of .18, .19, and .19, respectively (all p < .001). The number of classrooms for these analyses involving only MD students was 19; one classroom had no students with MD and was omitted.

The MD-only results in Tables 3, 4, and 5 follow the same pattern as those already described. Results for the PPS posttest in Table 3 show that SBI was a significant predictor, with an SBI-versus-control difference of 0.51 *SD*, or 70% of SBI students scored above the mean of control students. Math achievement level and pretest were both significant student predictors.

Table 5. Hierarchical Linear Modeling Results for GMADE Posttest.

Fixed effects	Full sample					Students with MD				
	B	*SE*	*t*	*df*	*p*	*B*	*SE*	*t*	*df*	*p*
Between-student model										
Sex	0.81	0.369	2.19	320	.029	1.06	0.440	2.42	217	.017
Asian	1.15	0.993	1.15	320	.250	2.97	1.711	1.73	217	.084
Black	–0.34	0.533	–0.64	320	.522	–0.19	0.645	–0.30	217	.767
Hispanic	0.16	0.488	0.33	320	.742	0.41	0.617	0.66	217	.510
Multiracial	–0.19	1.117	–0.17	320	.865	0.42	1.244	0.34	217	.736
Math achievement level	1.77	0.264	6.72	320	<.001	1.97	0.441	4.46	217	<.001
Pretest	0.25	0.050	5.00	320	<.001	0.23	0.062	3.74	217	<.001
Between-classroom model										
Intercept	13.31	0.661				12.62	0.880			
Treatment	1.06	0.466	2.28	13	.040	0.88	0.568	1.55	12	.147
Years experience	–0.02	0.049	–0.43	13	.671	–0.05	0.064	–0.75	12	.466
MKT IRT estimate	0.69	0.451	1.53	13	.149	0.93	0.557	1.67	12	.120
Special education	–0.61	0.343	–1.79	13	.097	–0.48	0.403	–1.18	12	.260
English language learner	0.15	0.340	0.43	13	.673	0.15	0.410	0.37	12	.718
FRL	–0.72	0.434	–1.66	13	.120	–0.89	0.536	–1.67	12	.121
Random effects	VC	*SD*	χ^2	*df*	*p*	VC	*SD*	χ^2	*df*	*p*
Classroom	0.23	0.481	28.14	13	.009	0.37	0.610	27.18	12	.007
Student	10.75	3.278				10.14	3.185			

Note. Here α = .05 for the test of the SBI effect and α = .10/13 = .008 for tests of the remaining fixed effects. GMADE = Group Mathematics Assessment and Diagnostic Evaluation (Pearson Education, 2004); MD = mathematics difficulties; MKT = Mathematical Knowledge for Teaching; IRT = item response theory; FRL = eligible for free or reduced-price lunch; VC = variance component.

The results for the PPS delayed posttest in Table 4 indicate that SBI was again a statistically significant predictor, with an SBI-versus-control difference of 0.35 *SD*, or 64% of SBI students scored above the mean of control students. Pretest and math achievement level were also significant predictors. The longitudinal analysis for the MD sample showed a somewhat different pattern from the full sample. The average linear slope was not statistically significant ($B = 0.42$, $t = 1.86$, $p = .062$; there was still significant variability among student linear slopes), but treatment was a significant moderator of change over time, with SBI students improving at a faster rate than control students ($B = 0.93$, $t = 3.23$, $p = .002$). Treatment was not a significant predictor for the GMADE posttest in Table 5, but once again pretest and math achievement level were significant predictors.

Discussion

Previous research demonstrated the benefits of SBI for middle school students who were predominantly White and enrolled in schools located mostly in rural settings in the Midwest (Jitendra et al., 2015). This study followed up and extended previous results regarding the effects of SBI in schools using a sample from a metropolitan area in the Southeast that was more racially diverse and included larger percentages of ELLs and students from economically disadvantaged backgrounds. Also, the onset of the intervention in the present study occurred earlier in the school year (late October) as compared with the original study (early January).

Overall, we again found that SBI is effective in enhancing proportional learning, as found in previous studies (e.g., Jitendra et al., 2009; Jitendra et al., 2011; Jitendra et al., 2013; Jitendra et al., 2015). Our first research question asked if SBI had a positive effect on PPS and general mathematical problem-solving outcomes for all students. The results showed that the SBI condition outperformed the control condition on PPS and that the effect was still present at the 11-week follow-up for PPS, providing evidence that SBI improves long-term retention of PPS skills. The ESs for the PPS posttest and delayed posttest were $g = 0.63$ *SD* and $g = 0.33$ *SD*. Furthermore, the learning trajectories of the SBI and control conditions were significantly different in that the rate of improvement for SBI students was faster than that of control students. This finding is similar to that found in the Jitendra et al. (2015) study and is noteworthy because the control condition covered approximately the same instructional topics as SBI and thus had the same advantage on the PPS assessment, which assessed proportional reasoning. Therefore, these results provide strong evidence of the superiority of SBI over instruction in the control condition in supporting student learning of ratios and proportional relationships. In both the present study and the original Jitendra et al. (2015) study, students in SBI classrooms learned the content more effectively than control students, which can be attributed to SBI practices such as using visual representations to highlight the underlying problem structure, engaging in problem-solving and metacognitive activities, and developing procedural flexibility.

These results provide strong evidence of the superiority of SBI over instruction in the control condition in supporting student learning of ratios and proportional relationships.

We also examined the overall mathematical problem-solving performance of the SBI and control conditions. Contrary to findings in prior SBI studies (Jitendra et al., 2009; Jitendra et al., 2011; Jitendra et al., 2013; Jitendra et al., 2015), we found that SBI outperformed the control condition on the GMADE assessment ($g = 0.32$) even though only one third of the GMADE items reflect concepts taught in the study. One explanation for this effect could be Wagner's (2006) theory of transfer in pieces, which highlights the importance of multiple examples necessary for "the incremental growth, systematization, and organization of knowledge resources that only gradually extend the span of situations in which a concept is perceived as applicable" (p. 10). The SBI program provides multiple examples and emphasizes the essential features of the various problem types to effect transfer of learning. At the same time, it is not known what attributes of SBI made a difference and positively affected the transfer of student knowledge in this study and not in previous studies. Nor do we know whether the benefits of SBI are enhanced when the intervention is delivered earlier in the school year rather than later. Furthermore, the improved performance on the GMADE could be largely due to those items reflecting concepts taught in the study, but because item-level responses are unavailable, it is impossible to be sure.

Our second research question focused on the effects of SBI on problem-solving outcomes for students with MD. The subgroup analysis showed a pattern of short- and long-term advantages for students with MD in SBI classrooms similar to that for the full sample on the PPS. The ESs for the PPS posttest and delayed posttest were $g = 0.51$ and $g = 0.35$. The pattern of immediate and delayed benefits also corresponds to the short-term and retention effects observed in prior SBI studies of students with MD (Jitendra, Dupuis, et al., 2016; Jitendra et al., in press). Although there was substantial variability among student linear slopes, contrary to the finding of Jitendra et al. (in press), our results showed that SBI students demonstrated significantly more growth on the PPS relative to control students.

However, similar to the Jitendra et al. (in press) study, there was no effect on the GMADE; however, the ES of $g = 0.26$ is nonnegligible (see What Works Clearinghouse, 2014), and the lack of statistical significance may be attributable to the modest number of classrooms. Substantively, it is possible that SBI did not enable students with MD to adequately transfer the knowledge that they had acquired to solve novel problems and that these students need more time and support to show gains in flexible knowledge of procedures for solving a range of problems. This study provides preliminary evidence that ambitious mathematics practices need to occur over a longer period for students with MD than for their general education peers, within the context of the current Common Core framework of high standards, to effectively influence transfer of knowledge to a new domain. Previous research (Jitendra & Star, 2012) suggests that this explanation is plausible.

Overall, the findings suggest that teacher time dedicated to PPS did not impede general mathematical problem-solving achievement of students. In fact, mathematical problem solving was enhanced for the whole sample. This is an important finding in that we found positive effects on the GMADE that contrast with the findings of Jitendra et al. (2015) for all students. Also notable is that the effects sizes on the PPS posttest for all students and for students with MD exceeded the effects ($g = 0.63$ vs. 0.46, $g = 0.51$ vs. 0.32) obtained in Jitendra et al. (2015, 2017). Interestingly, ESs for all students and for students with MD on the delayed posttest were comparable, even though the retention test was completed 11 weeks after the posttest in this study rather than 8 weeks later, as in Jitendra et al. (2015; Jitendra et al., in press; $g = 0.33$ vs. 0.32, $g = 0.35$ vs. 0.25). It is possible that the benefits of the intervention are enhanced when delivered earlier in the school year, but further controlled studies are needed to replicate the present results to determine whether PPS should be taught earlier in the school curriculum.

Our findings also highlight the importance of disaggregating research findings when evaluating the effects of instructional programs. While SBI generally improved PPS for all students in the sample, the effect for overall mathematical problem solving was not found for students with MD. It is possible that students with MD may have different instructional needs (beyond what is provided in Tier 1) than other students in overall mathematical problem solving. However, further investigations must determine whether more intensive instruction (e.g., more time, small group instructional arrangement) would be advantageous and, if so, how long the intervention should last and what the size of the instructional group should be.

Implications for Practice

Findings from the current study and prior SBI studies provide strong evidence that the SBI curriculum can be used within the MTSS framework in a preventative fashion to meet the needs of all students, including students who struggle to develop mathematical proficiency. With the increased implementation of MTSS, there is a need for empirically validated interventions in mathematics, especially Tier 1 interventions, to meet the instructional needs of a range of learners. Specifically, addressing the needs of struggling students is paramount, because these students may need subsequent Tier 2 or Tier 3 instruction and it would be inappropriate to place students in Tier 2 if there is no evidence to suggest that they had the opportunity to learn from well-designed Tier 1 instruction.

It would be inappropriate to place students in Tier 2 if there is no evidence to suggest that they had the opportunity to learn from well-designed Tier 1 instruction.

Ensuring that all students, including students at risk for MD, meet the Common Core State Standards has important implications for policy makers and teachers. For example, this implies that not only is it important to have strong Tier 1 programs to help students meet the expectations of the CCSS, but teachers also need a deep understanding of ratios and proportions and problem solving, as well as an understanding of what makes it difficult for students, and how they can scaffold appropriate strategies to improve learning. However, many teachers are not well prepared to teach mathematics (Krauss et al., 2008), especially problem solving, in that they often fail to discuss “what strategies students used to solve the problems or whether the solutions can be justified” (Woodward et al., 2012, p. 6). Teachers in this study were provided guidance in teaching PPS (e.g., selecting a strategy based on the numbers in the problem, checking whether the answer is reasonable) to reach a range of students. Consequently, it is important that teacher training translates to teachers helping struggling students make sense of their reasoning related to proportions, which is crucial for students to succeed in algebra and meet the demands of an increasingly competitive workplace, where the demand for mathematics-intensive science and engineering jobs are outpacing overall job growth three to one (National Mathematics Advisory Panel, 2008).

Limitations and Future Research

One limitation of the current study is the modest numbers of classrooms and students, as well as the large number of students with MD, which speaks to the need to further replicate SBI results with a larger sample in a controlled study. A forthcoming study will provide additional evidence of generalizability with large samples obtained from different geographic locations, which will accurately reflect the prevalence of students with MD in schools. Another potential limitation in this study is that fidelity was addressed by evaluating one videotaped lesson on ratio problem solving. Based on the relatively brief period of the intervention (6 weeks), one video-recorded observation may provide a representative sample of participant functioning (Breitenstein et al., 2010). However, teachers’ fidelity of implementation in this study may have been affected by the selection of the lesson, which was the second lesson on word problem solving in Unit 1 on ratio and proportion. Using new instructional materials initially may also have been challenging for the treatment teachers. For example, the majority of the teachers did not emphasize how the problem was similar or different from a previously solved problem.

A third limitation is a plausible novelty effect in that SBI teachers in this study and the Jitendra et al. (2015) study may have been motivated by the new approach, resulting in changes in their teaching practices that positively influenced student outcomes. However, the novelty effect may be less of an issue given the similar results across the two studies. Furthermore, we demonstrated in another study (Jitendra, Harwell, et al., 2016) the attenuation of the novelty effect based on the sustainability of SBI when implemented by SBI-experienced teachers as compared with SBI-novice teachers. In this study, SBI-experienced teachers implemented SBI with fidelity in subsequent years without additional professional development, and the effects for their students on the mathematics outcomes were comparable. Another limitation is the significant difference between the treatment and control conditions on the GMADE and PPS pretests despite randomization. However, the inclusion of pretest as a covariate in the modeling should eliminate or minimize bias resulting from preexisting differences when treatment and control conditions are compared for cross-sectional and longitudinal data.

Conclusions

The results of the current study support the view that SBI improves PPS for students with and without MD. SBI—with its emphasis on the underlying problem structure and categorization of problem types, visual representations, and instructional strategies (problem solving, metacognition, multiple solution) accompanied by explanations—seems to enhance performance by facilitating transfer of knowledge to solve novel problems. As such, integrating these strategies with mathematics content is important in connecting mathematics practices to mathematical content articulated in the CCSS (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

References

Adams, T. L., McLeod, J. C., Dion, J. K., Leiva, M. A., Larson, M., & Shaw, J. M. (2011). *Go math! Florida*. Orlando, FL: Houghton Mifflin Harcourt School.

Bloom, H. S., Richburg-Hayes, L., & Black, A. R. (2007). Using covariates to improve precision: Empirical guidance for studies that randomize schools to measure the impacts of educational interventions. *Educational Evaluation and Policy Analysis*, *29*, 30–59. doi:10.3102/0162373707299550

Bottge, B. A., Ma, X., Gassaway, L., Toland, M. D., Butler, M., & Cho, S. J. (2014). Effects of blended instructional models on math performance. *Exceptional Children*, *80*, 423–437. doi:10.1177/0014402914527240

Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health*, *33*, 164–173. doi:10.1002/nur.20373

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Cook, B. G. (2014). A call for examining replication and bias in special education research. *Remedial and Special Education*, *35*, 233–246. doi:10.1177/0741932514528995

Coyne, M. D., Cook, B. G., & Therrien, W. J. (2016). Recommendations for replication research in special education: A framework of systematic, conceptual replications. *Remedial and Special Education*, *37*, 244–253. doi:10.1177/0741932516648463

Doabler, C. T., Clarke, B., Kosty, D. B., Kurtz-Nelson, E., Fien, H., Smolkowski, K., & Baker, S. K. (2016). Testing the efficacy of a tier 2 mathematics intervention: A conceptual replication study. *Exceptional Children*, *83*, 92–110. doi:10.1177/0014402916660084

Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2015). *The value of replication for developmental science*. Retrieved from http://sites.uci.edu/gduncan/files/2013/06/Replication-paper-single-spaced.pdf

Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399–412. doi:10.1111/bjop.12046

Fuchs, D., Fuchs, L., & Compton, D. (2012). Smart RTI: A next-generation approach to multi-level prevention. *Exceptional Children*, *78*, 263–279. doi:10.1177/001440291207800301

Fuchs, D., Fuchs, L. S., & Vaughn, S. (Eds.). (2008). *Response to intervention: A framework for reading educators*. Newark, DE: International Reading Association.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (Document No. 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

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

Gottfredson, D. C., Cook, T. D., Gardner, F. E., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science*, *16*, 893–926. doi:10.1007/s11121-015-0555-x

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*, 60–87. doi:10.3102/0162373707299706

Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, *5*, 107–118. doi:10.1080/15366360701487138

Jitendra, A. K., Dupuis, D. N., Star, J. R., & Rodriguez, M. C. (2016). The effects of schema-based instruction on the proportional thinking of students with mathematics difficulties with and without reading difficulties. *Journal of Learning Disabilities*, *49*, 354–367. doi:10.1177/0022219414554228

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., & Karl, S. R. (2017). A randomized trial of the effects of schema-based instruction on proportional problem solving for students with mathematics problem-solving difficulties. *Journal of Learning Disabilities*, *50*, 322–336. doi:10.1177/0022219416629646

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., Karl, S. R., Lein, A. E., Simonson, G., & Slater, S. C. (2015). Effects of a research-based mathematics intervention to improve seventh-grade students' proportional problem solving: A cluster randomized trial. *Journal of Educational Psychology*, *107*, 1019–1034. doi:10.1037/edu0000039

Jitendra, A. K., Harwell, M. R., Karl, S. R., Dupuis, D. N., Simonson, G., Slater, S. C., & Lein, A. E. (2016). Schema-based instruction: The effects of experienced and novice teacher implementers on seventh-grade students' proportional problem solving. *Learning and Instruction*, *44*, 53–64. doi.org/10.1016/j.learninstruc.2016.03.001

Jitendra, A. K., & Star, J. R. (2012). An exploratory study contrasting high- and low-achieving students' percent word problem solving. *Learning and Individual Differences*, *22*, 151–158. doi:10.1016/j.lindif.2011.11.003

Jitendra, A. K., Star, J. R., Dupuis, D. N., & Rodriguez, M. C. (2013). Effectiveness of schema-based instruction for improving seventh-grade students' proportional reasoning: A randomized experiment. *Journal of Research on Educational Effectiveness*, *6*, 114–136. doi:10.1080/19345747.2012.725804

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction*, *21*, 731–745. doi:10.1016/j.learninstruc.2011.04.002

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J., Sood, S., Caskie, G., . . .Mack, T. R. (2009). Improving students' learning of ratio and proportion problem solving: The role of schema-based instruction. *Contemporary Educational Psychology*, *34*, 250–264. doi:10.1016/j.cedpsych.2009.06.001

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, *100*(3), 716. http://dx.doi.org/10.1037/0022-0663.100.3.716

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the educational sciences, *Educational Researcher*, *43*, 304–316. doi:10.3102/0013189X14545513

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum. doi:10.1111/j.2044-8317.1981.tb00621.x

Meyer, J. P. (2011). jMetrik (Version 2.1.0) [Computer software]. Retrieved from ItemAnalysis.com/jMetrik.com

National Center for Education Statistics. (n.d.). *NAEP data explorer*. Washington, DC: Institute of Education Sciences. Retrieved from http://nces.ed.gov/nationsreportcard/naepdata

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for mathematics*. Washington, DC: Authors.

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

Ozgun-Koca, S. A., & Altay, M. K. (2009). An investigation of proportional reasoning skills of middle school students. *Investigations in Mathematics Learning*, *2*(1), 26–48.

Pearson Education. (2004). Level M, Form A. In *Group Mathematics Assessment and Diagnostic Evaluation (GMADE)*. San Antonio, TX: Author.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). HLM 6.02: Hierarchical linear and nonlinear modeling [Computer software]. Lincolnwood, IL: Scientific Software International.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*, 90–100. doi:10.1037/a0015108

Spybrook, L., Bloom, H., Congdon, R., Hill, C., Martinez, A., & Raudenbush, S. (2011). *Optimal design plus empirical evidence: Documentation for the "Optimal Design" software*. Retrieved from http://sitemaker.umich.edu/group-based

Valentine, J. C., Biglan, A., Boruch, R. F., Castro, F. G., Collins, L. M., Flay, B. R., . . .Schinke, S. P. (2011). Replication in prevention science. *Prevention Science*, *12*, 103–117. doi:10.1007/s11121-011-0217-6

Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction*, *24*(1), 1–71. doi:10.1207/s1532690xci2401_1

What Works Clearinghouse. (2014). *WWC procedures and standards handbook (Version 3.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., . . .Ogbuehi, P. (2012). *Improving mathematical problem solving in Grades 4 through 8: A practice guide* (Document No. 2012-4055). Washington, DC: National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=16

Manuscript received November 2016; accepted January 2017.