Clarke, B., Doabler, C. T., Smolkowski, K., Baker, S. K., Fien, H., & Strand Cary, M. (2016). Examining the efficacy of a tier 2 kindergarten mathematics intervention. *Journal of Learning Disabilities*, *49*, 152–165. doi: 10.1177/0022219414538514

Grant/Contract Number: R324A120304

This research was supported by the ROOTS Project, Grant No. R324A120304, funded by the U.S. Department of Education, Institute of Education Sciences. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2016, Vol. 49(2) 152–165

DOI: 10.1177/0022219414538514
journaloflearningdisabilities.sagepub.com
SAGE

Examining the Efficacy of a Tier 2 Kindergarten Mathematics Intervention

Ben Clarke, PhD[1], Christian T. Doabler, PhD[1], Keith Smolkowski, PhD[2], Scott K. Baker, PhD[3], Hank Fien, PhD[1], and Mari Strand Cary, PhD[1]

Abstract

This study examined the efficacy of a Tier 2 kindergarten mathematics intervention program, ROOTS, focused on developing whole number understanding for students at risk in mathematics. A total of 29 classrooms were randomly assigned to treatment (ROOTS) or control (standard district practices) conditions. Measures of mathematics achievement were collected at pretest and posttest. Treatment and control students did not differ on mathematics assessments at pretest. Gain scores of at-risk intervention students were significantly greater than those of control peers, and the gains of at-risk treatment students were greater than the gains of peers not at risk, effectively reducing the achievement gap. Implications for Tier 2 mathematics instruction in a response to intervention (RtI) model are discussed.

Keywords

intervention, mathematics, response to intervention

In the past decade, educators and policy makers have become increasingly concerned with the low level of mathematics performance of U.S. students in relation to international comparisons (Trends in International Mathematics and Science Study [TIMSS]; Olson, Martin, & Mullis, 2008) and national standards (National Mathematics Advisory Panel [NMAP], 2008; National Research Council [NRC], 2001; Schmidt, Houang, & Cogan, 2002). Results from international comparisons show consistently low levels of performance and relative rank of American students compared to peers from other developed countries (TIMSS; Olson et al., 2008). The National Assessment of Educational Progress indicates that only 40% of fourth graders are at or above proficient in mathematics and almost a fifth (18%) are classified as below basic. Despite the low number of students meeting proficiency, the results indicate progress at most achievement levels. The average scores at the 25th, 50th, 75th, and 90th percentiles all were higher in 2011 than 2009. However, for students at the highest risk (i.e., 10th percentile), there was no increase in their average score from 2009 to 2011 (National Center for Education Statistics, 2011b). In an increasingly globalized economy where job growth in science, technology, engineering, and mathematics fields is expected to outpace overall job growth at roughly a 3:1 ratio (National Science Board, 2008), concern about mathematics achievement crosses the political aisle.

Signs of long-term difficulty in mathematics appear early with significant differences in student knowledge apparent at school entry on a range of concepts and skills from counting principles and number knowledge to more complex understandings of quantities, operations, and problem solving (Griffin, Case, & Siegler, 1994; Jordan, Kaplan, Locuniak, & Ramineni, 2007). A number of longitudinal research studies have begun to document that students who perform poorly in mathematics at the end of kindergarten are likely to continue to struggle throughout elementary school (Bodovski & Farkas, 2007; Duncan et al., 2007; Hanich, Jordan, Kaplan, & Dick, 2001; Morgan, Farkas, & Wu, 2009). For example, using a nationally representative sample of students from the Early Childhood Longitudinal Study–Kindergarten Cohort, Morgan et al. (2009) found that students who entered and exited kindergarten below the 10th percentile (considered an indicator of a mathematics learning disability; MLD) had a 70% chance of remaining in the lowest 10th percentile five years later. Their overall mathematics achievement in fifth grade remained 2 standard deviations below students who did not demonstrate a MLD profile in kindergarten. Kindergarten represents a logical starting point to begin and address the issue of preventing learning disabilities in mathematics. Without

[1]University of Oregon, Eugene, USA
[2]Oregon Research Institute, Eugene, USA
[3]Southern Methodist University, Dallas, TX, USA

Corresponding Author:
Ben Clarke, Center on Teaching and Learning, University of Oregon, 1600 Millrace Dr., Suite 208, Eugene, OR 97403, USA.
Email: clarkeb@uoregon.edu

targeted efforts in kindergarten, the learning gap is likely to persist and become more difficult to remediate over time as students exit kindergarten without a solid foundation for understanding increasingly complex mathematics (Geary, 1993; Jordan, Kaplan, & Hanich, 2002; Morgan et al., 2009).

One potential schoolwide system to increase student mathematics achievement is response to intervention (RtI). The legislative intention of RtI is to provide a means to identify students with a specific learning disability based on their lack of response to research-based instruction (D. Fuchs, Mock, Morgan, & Young, 2003). Implementation of the most common RtI models in schools attempts to provide research-based instruction to all students through a continuum of supports based on the student's level of need (Burns & Vanderheyden, 2006; L. S. Fuchs, Fuchs, & Zumeta, 2008). The varying degrees of support are typically classified as tiers with the first tier (Tier 1) being delivered to all students in a general education setting. Subsequent tiers provide greater instructional intensity (Tiers 2 and 3) for students who fail to demonstrate adequate growth. Although research on the use of RtI systems as a whole lack support (Baker, Fien, & Baker, 2010; Kovaleski & Black, 2010), there is evidence that individual elements of RtI models including specific intervention programs, instructional strategies, and assessment practices are effective (Gersten, Beckmann, et al., 2009).

Given the critical importance of a successful early start in mathematics, it would be reasonable to expect that an array of research-based instructional programs would be available to schools to help them meet the needs of at-risk students in the early elementary grades. Unfortunately, this is not the case. A review of mathematics programs aligned with the principles of RtI (screening to determine risk, delivery of a research-based intervention and progress monitoring) found a limited number of studies (Newman-Gonchar, Clarke, & Gersten, 2009). In addition, a recent overview of research on early mathematics intervention programs detailed the lack of studies using rigorous experimental designs (Dyson, Jordan, & Glutting, 2013). The few programs that have been empirically tested using rigorous experimental designs have tended to be intervention programs for at-risk students (e.g., Bryant et al., 2011; Dyson et al., 2013; L. S. Fuchs et al., 2005), rather than core curriculum programs for the whole class (e.g., Agodini et al., 2009) and have focused on developing an understanding of whole number concepts (Gersten, Beckmann, et al., 2009). The NMAP (2008) noted the dearth of research-based intervention programs as a significant shortcoming that required immediate attention to advance the quality of mathematics instruction provided by schools.

Recognizing that mathematics trajectories are established early in school, kindergarten represents a critical transition from informal to formal mathematics and the lack of research-based instructional programs, our research group developed ROOTS, a 50-lesson (Tier 2) kindergarten math intervention. Previously we had developed and evaluated the *Early Learning in Mathematics* (ELM) kindergarten core curriculum (Clarke et al., 2011; Davis & Jungjohann, 2009). ELM consists of 120 lessons and focuses on four key mathematics strands: (a) Number and Operations, (b) Geometry, (c) Measurement, and (d) Vocabulary. The first three of these map directly onto the three content domains contained in the National Council of Teachers of Mathematics (NCTM, 2006) Curriculum Focal Points, and the fourth (vocabulary) is addressed in the NCTM (2000) Process Standards. We tested the efficacy of ELM in a randomized controlled trial (RCT), randomly assigning 66 kindergartens classrooms to treatment and control conditions. Students in the treatment condition, ELM, outperformed students in control classrooms on two distal measures of math proficiency: the *Test of Early Mathematics Ability* (TEMA; $t = 2.41$, $p = .02$, Hedges's $g = .15$) and *Early Numeracy Curriculum-Based Measurement* (EN-CBM; $t = 1.99$, $p = .05$, $g = .13$; Clarke et al., 2011). In Condition × Risk Status analyses, at-risk students (defined as performing below the 40th percentile on the TEMA at pretest) demonstrated the greatest treatment benefit. At-risk treatment students significantly outperformed at-risk control students on both the TEMA ($t = 3.29$, $p = .0017$, $g = .24$) and EN-CBM total score ($t = 2.54$, $p = .0138$, $g = .22$).

The pattern of findings is important for two reasons. First, differential impact favoring the at-risk students was aligned with our theoretical framework. In developing ELM our objective was to create a core mathematics programs (i.e., Tier 1) that would address the needs of the average- and high-performing students (in the analysis, students above the 40th percentile at pretest performed the same at posttest in both ELM and control conditions) and substantially increase the mathematics achievement of students at risk for math difficulties (which occurred). Second, although ELM was beneficial to at-risk students in particular, the performance of at-risk students at the end of the year did not match the performance of students *not* at risk. In other words, although there was differential impact by risk status, the ELM program did not fully eliminate the gap between at-risk and average-achieving students.

Our findings that ELM helped reduce but not fully eliminate the achievement gap between at-risk students and their on-track peers were not unexpected. Given the deficits that at-risk students enter kindergarten with and the limitations of a core program targeting the learning needs of all students to fully address the learning needs of at-risk students, we developed ROOTS as a Tier 2 intervention, to be used in conjunction with ELM. The ROOTS intervention was designed to focus exclusively on Number and Operations because an in-depth understanding of the whole number system is a critical step in achieving proficiency in more sophisticated mathematics, such as rational numbers and algebra (Gersten, Beckmann, et al., 2009; NCTM, 2006; NRC, 2001, 2009). In kindergarten, interventions should target the building of conceptual understanding of whole number, operations, and relations (Gersten & Chard, 1999; NRC, 2009, Wu, 2005). For example, authors of the Institute of Education Sciences practice guide on effective mathematics instruction and intervention for at-risk students observed that "individuals knowledgeable in instruction and mathematics look for [and develop] interventions that focus on whole numbers extensively in kindergarten through grade 5" (Gersten, Beckmann, et al., 2009, p. 18). Concurring, the Common Core State Standards for Mathematics (CCSS, 2010) recommend extensive coverage of whole number concepts and skills during kindergarten (i.e., counting and cardinality, operations and algebraic thinking, and number and operations in base 10). In kindergarten, this often entails a strong focus on developing number sense and related skills (Dyson et al., 2013).

Although the definition of number sense varies among experts in the field (Berch, 2005), there is general agreement that it permits a child to flexibly manipulate numbers, use computation methods, and understand strategies for solving real-world and mathematical problems (Dehaene, 1997; Gersten & Chard, 1999; Griffin, 2004). For example, students with early number sense can use efficient counting strategies, compose and decompose numbers, make magnitude comparisons, and understand mathematical relationships, such as the connection between numbers and quantities. Because a growing body of evidence indicates that number sense and related skills can be taught and that instruction can increase the probability that at-risk learners will acquire deep numerical knowledge and reach proficient levels of mathematics (Bryant et al., 2011; Chard et al., 2008; Dyson et al., 2013; Griffin et al., 1994; Locuniak & Jordan, 2008), an optimal window of opportunity to target this critical construct is in kindergarten. Therefore, a central aim of the ROOTS intervention is to support students' early development of number sense. Many children get off to a good start in mathematics by developing early number sense through informal learning experiences at home with parents and in preschool settings. These children enter kindergarten prepared to acquire a richer understanding of number and numeration. For example, they are ready to learn how to make quantitative comparisons and use the word-object correspondence principle (Dougherty, Flores, Louis, & Sophian, 2010; Jordan et al., 2007; Lago & DiPerna, 2010). Conversely, children who lack these learning opportunities are typically at an elevated risk for mathematics difficulties and disabilities. As a result, they struggle to grasp concepts such numerical equivalence and understand the relationship between quantities (Gersten, Beckmann, et al., 2009; Griffin, 2004).

Purpose of the Study

The study employed random assignment of classrooms to condition and teacher selection of students to participate in the study to test the impact of a kindergarten intervention program, ROOTS, on the achievement of students at risk in mathematics. Specifically, we had one primary research question:

1. What is the impact of the ROOTS program on the mathematics achievement of at-risk students?

We had one secondary research question:

2. Do ROOTS students reduce the achievement gap with their non-at-risk peers by making greater gains than their non-at-risk peers?

This study has the potential to contribute to the growing yet still underdeveloped research base on effective mathematics instruction for student's at risk for MLD by examining a Tier 2 intervention program with a focus on critical whole number content and effective instructional design and delivery principles. In addition, because ROOTS is delivered within the context of a research-based Tier 1 core program (ELM), the study has the potential to allow for an examination of how to construct effective RtI models to deliver services to at-risk students across multiple tiers of instruction.

Method

Design

This study involved full-day kindergarten teachers who had participated in the ELM whole-classroom study (Clarke et al., 2011). Classrooms were assigned randomly to condition, and then within classrooms teachers selected students whom teachers expected would most benefit from small-group instruction. Specifically, classrooms were randomly assigned to treatment or control conditions, blocking on teachers' prior experience with ELM. That is, we randomly assigned teachers with 1 year of ELM experience to ROOTS or control and then randomly assigned teachers new to ELM implementation to ROOTS or control. In schools with multiple classrooms, we also assigned classrooms to condition within school. Blocking, also called stratification, on ELM experience and school experimentally controls for biases that might stem from systematic differences between conditions (e.g., more ROOTS teachers with no prior ELM experience). A total of 29 classrooms were included: 14 in the treatment condition (ELM + ROOTS) and 15 in the control condition (ELM only). Teachers were asked to nominate the five lowest performing students or those who would most benefit from a small-group math intervention. Teachers nominated 140 students as eligible for small-group instruction, with 67 students in intervention classrooms and 73 students in control classrooms. Twenty-three teachers identified five students, one teacher identified three, four teachers identified four students, and one identified six, with the deviations from five students split similarly across conditions.

Classroom teachers in both conditions provided whole class ELM instruction throughout the year for all students, and treatment and control classrooms provided the same amount of daily mathematics instruction. In intervention classrooms (ELM + ROOTS), the "ROOTS students" received all of the whole-class ELM instruction. On 3 days per week, however, instead of practicing that day's ELM topics independently at the end of the lesson (i.e., math practice worksheets), they received ROOTS instruction. Given that ROOTS was not offered in control classrooms, nominated control students participated in whole class ELM instruction, 5 days per week, including all of the individualized math practice. We controlled for time by delivering the ROOTS instruction during the individual, worksheet-based math practice portion of ELM in treatment classrooms. ROOTS instruction began in January and continued until the end of May. Trained instructional assistants provided ROOTS instruction.

Participants

Instructional assistants. A total of 14 instructional assistants (IAs) participated in the study; 13 were female and all identified themselves as White. IAs were included in the study based on time and schedule availability. Four of the IAs had college degrees, of whom 2 held current teacher certifications in elementary education. Of the remaining 2 IAs who were college graduates, one had a degree in business and the other had a degree in education but did not hold a current teaching license. With respect to the remaining 11 IAs, 3 held associate's degree and 7 were high school graduates. Five of the IAs had completed college level coursework in mathematics. In this sample, 5 of the IAs had 10 or more years' experience, 3 had between 4 and 6 years' experience, and 5 had 1 to 3 years' experience. At the time of the present study (2009–2010), one IA was new to teaching. It is important to note that all IAs were employed by the participating school districts.

Students. Students were drawn from two districts in the Pacific Northwest. In District A 46% of students were White, 27% Hispanic, 6% Asian, and 2% Black, and 44% were eligible for free or reduced lunch. In District B, 61% of students were White, 20% Hispanic, 1% Black, and 1% Asian, and 50% were eligible for free or reduced lunch. Within the ROOTS condition, 51% of students were male, 72% were English learners, 15% received special education services, and the average age was 66.0 months (SD = 3.5). Among control participants, 55% were male, 56% were English learners, 16% received special education services, and their average age was 66.5 months (SD = 3.6). Of the 122 students with a TEMA percentile rank, 91% scored at or below the 10th percentile, with 95% of students in ROOTS classrooms and 88% of students in control classrooms falling below the 10th percentile.

The sample also included 538 students who were not eligible for ROOTS, with 253 in intervention classrooms and 285 in control classrooms. Within the intervention classrooms, 54% of students were male, 34% were English learners, and the average age was 66.8 months (SD = 3.5). Among control participants, 46% were male, 36% were English learners, and their average age was 67.2 months (SD = 3.6). All of these students received ELM instruction, and none of these students participated in ROOTS.

Measures

Fidelity of implementation. Online logs completed by the 14 IAs who delivered the ROOTS intervention revealed that groups generally completed all 50 ROOTS lessons during the year. Trained research staff also directly measured implementation fidelity using a standardized observation instrument. The observation instrument was specifically designed to target mathematics activities within each lesson of the ROOTS curriculum. During the observations, observers coded whether IAs taught key design components prescribed within each lesson activity.

All observations were scheduled in advance and observers coded fidelity of implementation data for the duration of the assigned 20-min instructional time periods. Each ROOTS group was observed 3 times over the course of the study, with approximately 4 to 5 weeks separating each observational round. Observers rated implementation fidelity using a 3-point rating scale, where a score of 3 represented *full implementation*, 2 represented *partial implementation*, and 1 indicated an activity was *not taught*. Fidelity scores were computed as the mean across all lesson activities. The mean across the three observations per ROOTS group were used as an overall indicator of implementation fidelity. IAs demonstrated high fidelity scores (M = 2.92, SD = 0.06) for prescribed lesson activities. Interobserver reliability was measured on 20% of all observation occasions. Reliability checks consisted of two observers simultaneously observing ROOTS instruction and coding fidelity of implementation data. Intraclass correlation coefficients (ICCs) were calculated to measure interobserver reliability. The ICC gives the proportion of variance associated with the occasion, opposed to observers. ICCs of .00 to .20 representing slight reliability, .21 to .40 representing fair reliability, .41 to .60 representing moderate reliability, .61 to .80 representing substantial reliability, .81 to 1.00 and nearly perfect (Landis & Koch, 1977).

Interobserver reliability of the implementation fidelity mean across three observations per small group showed substantial agreement between observers (ICCs = .67).

Test of Early Mathematics Ability. TEMA (PRO-ED, 2007) is a norm-referenced individually administered measure of early mathematics for children ages 3 to 8 years 11 months. The TEMA is designed to identify student strengths and weaknesses in specific areas of mathematics. The TEMA measures both formal mathematics and informal mathematics including skills related to counting, number facts and calculations, and related mathematical concepts. The test authors report alternate-form reliability of .97, and test–retest reliability ranges from .82 to .93. Concurrent validity with other criterion measures of mathematics is reported as ranging from .54 to .91.

Early Numeracy Curriculum-Based Measurement. EN-CBM (Clarke & Shinn, 2004) is a set of four measures based on principles of curriculum-based measurement (Shinn, 1989). Each 1-min fluency-based measure assesses an important aspect of early numeracy development including magnitude comparisons and strategic counting. The EN-CBM measures have been validated for use with kindergarten students including established validity with other measures of early mathematics including the *Number Knowledge Test* and the *Stanford Achievement Test* (Chard et al., 2005; Chard et al., 2008). The Oral Counting measure requires students to orally rote count as high as possible without making an error. Concurrent and predictive validities range from 46 to .72. The Number Identification measure requires students to orally identify numbers between 0 and 10 when presented with a set of printed number symbols. Concurrent and predictive validities range from .62 to .65. The Quantity Discrimination measure requires students to name which of two visually presented numbers between 0 and 10 is greater. Concurrent and predictive validities range from .64 to .72. The Missing Number measure requires students to name the missing number from a string of numbers (0–10). Students are given strings of three numbers with the first, middle, or last number of the string missing. Concurrent and predictive validities range from .46 to .63.

Procedures

Data collection. All measures were individually administered to students. Trained staff with extensive experience in collecting educational research for research projects administered all student measures. All data collectors were required to obtain interrater reliability coefficients of .90 prior to collecting data with students. Follow-up trainings were conducted prior to each data collection period to ensure continued reliable data collection. Student assessment protocols were processed using Teleform, a form processing application. Tests of Teleform scoring procedures of assessment protocols from previous research projects reveal high reliability values (i.e., .99) relative to assessor-scored protocols (.95).

ROOTS intervention. ROOTS is a Tier 2 kindergarten intervention program that was designed to be delivered by IAs in small-group instructional formats, 3 times per week, for 16 to 20 weeks during the second half of the school year. In contrast to the control condition (ELM only), ROOTS differs on a number of key variables. First, ROOTS is taught in small groups, whereas ELM is taught to the whole classroom. ELM occurs everyday and contains 120 lessons. ROOTS occurs 3 days per week and contains 50 lessons. ROOTS exclusively focuses on content associated with whole number understanding. In contrast, ELM covers content in whole number understanding, geometry and measurement and thus is broader in content coverage than ROOTS. The goal of ROOTS is to support students' development of procedural fluency with and conceptual understanding of whole number concepts. The specific focus on whole number aligns with the CCSS (2010) and calls from mathematicians and expert panels for more focused and coherent Tier 1 curricula (NCTM, 2006; NMAP, 2008), and intervention programs designed to meet the needs of students at risk for MLD (Gersten, Beckmann, et al., 2009). ROOTS provides in-depth instruction in whole number concepts by linking the informal mathematical knowledge developed prior to school entry with the formal mathematical knowledge developed in kindergarten. The program includes 50 lessons, approximately 20 min in duration. Each lesson consists of 4 to 5 brief math activities that center on three key areas of whole number understanding: (a) Counting and Cardinality, (b) Number Operations, and (c) Base 10/Place Value. Curricular objectives advance students from an initial understanding of whole number through more sophisticated aspects of whole numbers in kindergarten mathematics. For example, the first half of the curriculum addresses counting objects, identifying numbers, and counting on from a given number. In the second half, lessons focus on beginning computational methods, such as adding one to a number, and place value concepts, such as using base 10 models to compose and decompose teen numbers into one 10 and so many ones.

A central feature of the ROOTS program is its *explicit* and *systematic* approach to instruction. Carnine, Silbert, Kame'enui, and Tarver (2004) described explicit and systematic instruction as a method for teaching the "essential skills in the most effective and efficient manner possible" (p. 5). The emerging body of evidence generated by mathematics interventions for students at risk for math difficulties suggests convincingly that explicit and systematic instruction should be an integral approach for teaching students struggling with mathematics (Baker, Gersten, & Lee, 2002; Gersten, Chard, et al., 2009; Kroesbergen & Van Luit, 2003). In small-group instructional formats, like ROOTS, explicit and systematic instruction has the potential capacity to (a) deeply engage students in critical mathematics content, such as whole number concepts and (b) increase intervention intensity through differentiated instruction matched to student need (Baker et al., 2010; Clarke et al., 2011; L. S. Fuchs & Vaughn, 2012). Consequently, it can accelerate learning for kindergartners at risk for MLD (Gersten, Beckmann, et al., 2009).

ROOTS incorporates the principles of instructional delivery that have been empirically validated to improve the mathematics achievement of at-risk learners and students with learning disabilities (Baker et al., 2002; Gersten, Beckmann, et al., 2009; Nelson-Walker et al., 2012). These delivery principles include modeling and demonstrating what students will learn, providing guided practice opportunities, using visual representations of mathematics, and delivering academic feedback. For example, lessons provide interventionists with specific guidelines for demonstrating concepts and skills associated with whole numbers, and providing timely academic feedback to students as they engage in learning activities. The program also provides students with frequent and structured practice opportunities to promote procedural fluency and incorporates a host of visual representations to deepen conceptual understanding. Combined, the program's delivery principles form an instructional base for interventionists to facilitate overt and conspicuous instructional interactions among teachers and students around key math content. An important category of these instructional interactions is mathematical discourse or math verbalizations (Gersten, Beckmann, et al., 2009; NRC, 2001, 2009). Math verbalizations permit students to demonstrate their mathematical thinking and understanding. Such verbalizations are critical for gauging mathematics proficiency, particularly when students have yet to develop the math skills necessary to work independent of teacher support (Doabler et al., in press). In ROOTS, interventionists elicit math verbalizations from individual students as well as the group at large. Group verbalizations offer a way for interventionists to engage all students in the lesson's content. Verbalizations directed to individuals allow one student to practice with math content on his or her own. These opportunities also serve as effective means for differentiating instruction for students struggling with math. For example, in ROOTS, an interventionist might provide an easier response opportunity for a particular student who has difficulty grasping base 10 ideas, such as the grouping-by-10s concept.

Systematic instruction is the "behind-the-scenes" design activities (Kame'enui & Simmons, 1999; Simmons et al., 2007) that attend to the architectural features of a curriculum. Principles of systematic instruction include prioritizing instruction around critical content, connecting new content with students' background knowledge, selecting and sequencing instructional examples, and scaffolding instruction. ROOTS focuses intensely on the whole number standards identified in the CCSS (2010). When introducing students to new and difficult mathematics concepts and skills, for example, the program initiates instruction with simpler teaching examples. Once students demonstrate initial proficiency with targeted math content, instructional scaffolds are systematically withdrawn to promote learner independence. Finally, the program incorporates positive teaching examples along with a select number of nonexamples to promote students' discrimination skills (Coyne, Kame'enui, & Carnine, 2011).

Professional development. Participating IAs attended three professional development (PD) workshops focused on the ROOTS curriculum. The initial PD workshop targeted the instructional objectives of Lessons 1 to 25, the critical content of kindergarten mathematics (CCSS, 2010), small-group management techniques, and the instructional practices that have been empirically validated to increase student math achievement (e.g., teacher provided academic feedback; Gersten, Chard, et al., 2009). In the second and third workshops a similar format was followed, except that the focus was on the second half of the ROOTS curriculum, Lessons 26 to 50. Workshops were 4 hr in length and were organized around three principles: (a) active participation, (b) content focused, and (c) coherence. On at least three occasions, IAs also received in classroom coaching from two expert teachers to increase implementation fidelity. Implementation research shows that ongoing coaching enhances teachers' sustained use of new instructional practices (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). Two former educators, who were knowledgeable in the science of early mathematics development and instruction, served as coaches during the study. Typical coaching visits included direct observation and postobservation feedback focusing on instructional delivery and implementation fidelity. Some IAs received more than three coaching visits if they or the coach felt more support was warranted (e.g., when there were particularly pervasive student behavior problems or the IA struggled with lesson implementation).

Statistical Analysis

We assessed intervention effects on each of the primary outcomes with a mixed model (multilevel) time by condition analysis (Murray, 1998) to account for students nested within classrooms. Primary analyses included the students in each classroom identified as at risk for math difficulties by the classroom teacher. Because each classroom included only one small group, the classroom and small group are considered the same unit for analysis purposes. The analysis tests differences between conditions on change in outcomes from the fall of kindergarten (T_1) to the spring (T_2) clustered within classroom. The specific model tests time,

T, coded 0 at T_1 and 1 at T_2, condition, C, coded 0 for control and 1 for ROOTS, and the interaction between the two with the following hierarchical model.

$$Y_{ij} = \pi_{0j} + \pi_{1j}T_{ij} + e_{ij}e_{ij} \sim N\left(0, \sigma^2\right)$$

$$\pi_{0j} = \beta_{00} + \beta_{01}C_j + r_{0j}r_{0j} \sim N\left(0, \tau^2\right)$$

$$\pi_{1j} = \beta_{10} + \beta_{11}C_j + r_{1j}r_{1j} \sim N\left(0, \tau^2\right)$$

In the Level 1 model, the first equation, Y_{ij} represents a score for individual i within classroom j, and the model includes time, T_{ij}, as a predictor. The two Level 2 equations predict the intercept, π_{0j}, and slope, π_{1j}, from the Level 1 model with condition, Cj. The composite model, however, created by substituting the Level 2 equations into the Level 1 equation better corresponds to the presentation of results:

$$Y_{ij} = \beta_{00} + \beta_{01}C_j + \beta_{10}T_{ij} + \beta_{11}T_{ij}C_j + \left(r_{0j} + T_{ij}r_{1j} + e_{ij}\right)$$

Given the coding of C and T, the model includes the pretest intercept for the control condition, β_{00}, the difference between conditions at pretest, β_{01}, the estimate of gains for the control condition, β_{10}, and the difference in gains between conditions, β_{11}. The model also includes three error variances for the classroom-level intercept, r_{0j}, the classroom-level gains, $T_{ij}r_{1j}$, and the residual, e_{ij}. The models included a fourth variance term to represent the student-level covariation between pretest and posttest assessments, consistent with standard gain-score analyses.

With 29 classrooms, tests of time by condition used 27 degrees of freedom. The test of net differences, opposed to covariate-adjusted outcomes, provides an unbiased and straightforward interpretation of the results (Cribbie & Jamieson, 2000). The nested time by condition analysis also accounts the intraclass correlation associated with multiple students nested within the same classrooms. We also included all students in each classroom for additional analyses to test gains among ROOTS students compared to their typically achieving peers. These analyses entered ROOTS eligibility as a moderator.

Model estimation. We fit models to our data with SAS PROC MIXED version 9.2 (SAS Institute, 2009) using restricted maximum likelihood, generally recommended for multilevel models (Hox, 2002). From each model, we estimated fixed effects and variance components. This analysis approach included all available data, whether or not students' scores were present at both time points. Maximum likelihood estimation for the time by condition analysis uses of all available data to provide potentially unbiased results even in the face of substantial attrition, provided the missing data were missing at random (Schafer & Graham, 2002). In the present study, we did not believe that attrition or other missing data represented a meaningful departure from the missing at random assumption, meaning that missing data did not likely depend on unobserved determinants of the outcomes of interest (Little & Rubin, 2002). The majority of missing data involved students who were absent on the day of assessment (e.g., due to illness) or transferred to a new school (e.g., due to their family moving).

The models assume independent and normally distributed observations. We addressed the first, more important assumption (van Belle, 2008) by explicitly modeling the multilevel nature of the data. The data in the present study also do not markedly deviate from normality; skewness and kurtosis fell below an absolute value of 0.65 for both the TEMA standard score and CBM at pretest and posttest. Nonetheless, multilevel regression methods have also been found quite robust to violations of normality and outliers have a limited influence on the results in a variety of contexts For example, Hannan and Murray (1996) showed that group-randomized trials do not typically suffer from violations of normality at the individual level for samples with at least 10 clusters per condition. Murray and colleagues (2006) showed that violations of normality at either or both the individual and group levels do not bias results as long as the study is balanced at the group level.

Effect sizes. To ease interpretation, we computed an effect size, Hedges's g (Hedges, 1981), for each fixed effect. Hedges's g, recommended by the What Works Clearinghouse (WWC, 2011), represents an individual-level effect size comparable to Cohen's d (Cohen, 1988; Rosenthal & Rosnow, 2008).

Results

Table 1 presents means, standard deviations, and sample sizes for the TEMA standard score, the TEMA percentile, and the EN-CBM by assessment time and condition. Below we present results for tests of attrition effects, ROOTS intervention impact, and ROOTS students compared to peers who did not qualify for ROOTS. We also examined whether transitions between risk categories differed between students in ROOTS compared to their peers.

Attrition

Student attrition was defined as students with data at T_1 but missing data at T_2, and we examined attrition with respect to the ROOTS-eligible sample of 140 students. For the TEMA standard score, we experienced 10.7% attrition at T_2, with 11 of 73 students missing T_2 data in control classrooms and 4 of 67 students missing T_2 data in ROOTS classrooms, $\chi^2(1) = 3.02$, $p = .0821$. EN-CBM scores were

Table 1. Descriptive Statistics for Mathematics Measures by Condition and Assessment Time.

		Students eligible for ROOTS				Students ineligible for ROOTS			
		Intervention		Control		Intervention		Control	
Measure	Statistic	T_1	T_2	T_1	T_2	T_1	T_2	T_1	T_2
TEMA standard score	*M* (*SD*)	67.3 (9.58)	85.3 (9.54)	70.3 (9.89)	84.2 (12.05)	86.1 (14.51)	100.4 (12.32)	85.7 (15.42)	99.2 (12.37)
TEMA percentile	*M* (*SD*)	4.0 (9.17)	20.3 (15.18)	5.1 (6.52)	20.5 (18.42)	25.1 (24.48)	51.6 (24.70)	24.6 (25.77)	48.2 (25.60)
EN-CBM	*M* (*SD*)	14.9 (15.78)	106.3 (38.99)	21.4 (22.46)	100.0 (40.76)	61.3 (44.17)	160.0 (46.91)	60.9 (48.25)	153.1 (45.62)
Sample size	*n*	58	63	67	63	228	236	260	255

Note. The sample sizes represent the maximum students available across measures for each assessment period. Minimum sample sizes included 3 fewer students among students eligible for ROOTS (T_1 TEMA for controls) and 5 fewer students among students ineligible for ROOTS (T_1 TEMA for intervention sample). EN-CBM = *Early Numeracy Curriculum-Based Measurement*; TEMA = *Test of Early Mathematics Ability*.

Table 2. Fixed Effect and Variance Component Estimates From the Test of Condition on Mathematics Outcomes, With Hedges's g Values for the Time × Condition Effect.

Effect or statistic	TEMA standard score	EN-CBM
Fixed effects		
Intercept	69.93**** (1.60)	20.88*** (5.41)
Time	14.00**** (1.30)	78.89**** (6.15)
Condition	−2.84 (2.31)	−5.98 (7.86)
Time × Condition	4.07* (1.85)	12.01 (8.87)
Variance components		
Residual	53.66**** (8.62)	615.83**** (95.00)
Student	37.37*** (10.77)	133.65 (82.63)
Classroom intercept	17.90* (9.03)	131.36 (92.20)
Classroom gains	−0.84 (4.03)	135.03 (84.38)
Hedges's *g*		
Time × Condition	.375	.301

Note. Table entries show parameter estimates with standard errors in parentheses. Time is coded 0 for T_1 and 1 for T_2. Condition is coded 0 for control and 1 for ROOTS. All test fixed effects used 27 *df*. EN-CBM = *Early Numeracy Curriculum-Based Measurement*; TEMA = *Test of Early Mathematics Ability*.
*$p < .05$. ***$p < .001$. ****$p < .0001$.

missing for 10.0% of students at T_2, with 10 of 73 students missing T_2 data in control classrooms and 4 of 67 students missing T_2 data in ROOTS classrooms, $\chi^2(1) = 2.32$, $p = .1278$. Thus, attrition rates did not differ between conditions.

Although differential rates of attrition are undesirable, differential scores on math tests present a far greater threat to validity, so we conducted an analysis to test whether student math scores were differentially affected by attrition across conditions. We examined the effects of condition, attrition status, and their interaction on pretest scores of TEMA and EN-CBM within a mixed-model analysis of variance (Murray, 1998), which nests students' T_1 scores within classrooms and condition. We did not find statistically significant interactions between attrition and condition for either the TEMA standard score ($p = .6877$) or the EN-CBM ($p = .6927$). That is, we found no evidence of differential attrition for our two dependent variables.

Intervention Effects for ROOTS

Pretest differences. First, we tested whether students differed between conditions at pretest. No differences were found on either the TEMA standard score ($t = -1.23$, $df = 27$, $p = .2285$) or the EN-CBM ($t = -0.76$, $df = 27$, $p = .4536$). This suggests but cannot demonstrate that students were similar at pretest because the study has sufficient power (.80) to detect only medium differences between conditions ($g \approx .45$ to .60, depending on measures).

Gains across kindergarten. Among those students identified as eligible for ROOTS, we found statistically significant gains among students provided with ROOTS over those in control classrooms on the TEMA standard scores ($t = 2.19$, $df = 27$, $p = .0371$), but not the EN-CBM total score ($t = 1.35$, $df = 27$, $p = .1870$). The nested time by condition model estimated differences in gains between intervention conditions of 4.1 for the TEMA standard scores and 12.0 for EN-CBM. These corresponded to Hedges's *g* effect sizes of .38 for the TEMA standard score and .30 for the EN-CBM. Complete model results can be found in Table 2.

Closing the Gap

We hypothesized that students provided with the ROOTS intervention would make greater gains than students in the same intervention classrooms who did not receive ROOTS. That is, ROOTS was designed to close the gap between lower performing students and higher performing students. We tested this model with a second mixed model time by condition analysis. This time the model included the complete sample of students with an additional predictor, ROOTS participation, and its interaction with condition. We then examined a specific test of math gains between students who received ROOTS versus those who did not in intervention classrooms. The gains made by ROOTS intervention students exceeded gains by students who did not receive ROOTS by 3.5 on the TEMA standard score (t = 2.33, df = 27, p = .0273) but not on the EN-CBM (t = −1.45, df = 27, p = .1579). The differential gain was statistically significant on the TEMA but was not on the EN-CBM.

It is interesting that students selected for small-group instruction in control classrooms made statistically significantly smaller gains on the EN-CBM (−14.3) than students not selected in control classrooms (t = −2.55, df = 27, p = .0169). Although we had not hypothesized this result, it is generally consistent with the hypothesis that the ROOTS intervention may help close the gap, or in this case keep it from widening, between lower performing students and their more typically achieving peers.

Transitions Between Risk Categories

To address the practical implications of the effects of the ROOTS curriculum, we created a cross-tabulation of transitions between risk categories from pretest to posttest for students deemed eligible for ROOTS. The boundary between high-risk and no risk was set at the 10th percentile. The 10th percentile was selected because it corresponds roughly to the percentage of the student population that is eventually classified as learning disabled in mathematics (Geary, 2004) and because all but 10 students in our at-risk sample fell below this marker at pretest. This analysis included the 111 students with TEMA data at both T_1 and T_2. In both conditions, students migrated out of the high-risk category at a statistically significant rate (control, McNemar's S = 18.61, p < .0001; intervention, McNemar's S = 28.00, p < .0001). In the control condition, 50 of 57 students (87.7%) were identified as high risk at pretest, as were 51 of 54 intervention students (94.4%). Of the high-risk students at pretest, 24 (48.0%) in control classrooms and 28 (54.9%) in ROOTS classrooms shifted out of the high-risk category by the end of the school year. Although 6.9% more students who received ROOTS transitioned out of the high-risk category than control students, this difference was not statistically significant, $\chi^2(1) = 0.48$, p = .4877.

Discussion

We examined the impact of a kindergarten mathematics intervention program, ROOTS, on the achievement of at-risk students. We hypothesized that students in the ROOTS condition would demonstrate greater gains than their control peers and reduce the achievement gap between themselves and their non-at-risk peers. Both hypotheses were partially supported by our findings. ROOTS students demonstrated significant gains compared to their controls on one of two distal measures and on both measures demonstrated substantively important positive effects (WWC, 2011). In addition, students in ROOTS reduced the achievement gap between themselves and their non-at-risk peers. In contrast, it should be noted that the achievement gap in the control classrooms widened. This finding in the control classrooms is consistent with the "Matthew effect," in which the gap between low- and high-performing students increases over time. This effect is most commonly associated with low- and high-performing student samples in reading, but it is entirely consistent theoretically with similar predictions that could be made in early mathematics, particularly in the absence of effective intervention.

An important consideration when evaluating the findings from the study is the context in which the research took place. In comparison to other studies where an intervention program is compared to business as usual condition or practice that has no evidence of effectiveness, control students in our study received instruction in a program, ELM, that has been shown to be efficacious with at-risk students (Clarke et al., 2011). Second, because ROOTS was delivered during ELM practice time, there was no difference in the overall amount of time spent on mathematics instruction between treatment and control students.

Limitations

The findings in this study should be interpreted with awareness of limitations to the study. The primary limitation of the study is the potential for selection bias. Since students were not randomly assigned to condition, we are limited in our ability to attribute cause. The nonrandom selection of students introduces the possibility of selection bias. That is, the potential exists that teachers, because they were aware of the condition of their classroom, may have selected students in a different manner than if they were not aware of condition. This process may have resulted in different groups of students being selected in treatment and control classrooms and differences between treatment and control students on either observed (i.e., pretest mathematics achievement scores) or unobserved variables potentially confounding the interpretation of the study. In addition, we utilized a selection method that relied on teachers selecting students based on classroom data and their observation of student behavior and performance on a daily basis. Although the selection model chosen mirrors practice and thus emphasizes external validity, it sacrifices internal validity where the same criteria (e.g., cut score on a screening measure) would be used across classrooms.

The demographics of the sample in the present study differed from national demographics and thus as with any study conducted with a limited sample caution should be exercised when interpreting results. The importance of replication in education research is garnering increased attention (Duncan, Engel, Claessens, & Dowsett, 2013) given the complexities of social science research and the need to expose the null hypothesis regarding an intervention's hypothesis to repeated attempts at falsification. Replication using an array of experimental designs including RCTs and rigorous quasi-experimental designs across geographically and demographically diverse sites with different samples would increase confidence in evaluating the impact of ROOTS. Currently replication studies of ROOTS are in progress or planned in three different states.

The effect sizes reported in the study may also represent the upper bound of potential treatment effects. If teachers were able to accurately select the student most likely to benefit from the treatment, then the results for that group would be the maximum possible for the intervention. In part, this represents the goal of a tiered instructional model where students at risk are deliberately selected to receive an intervention that would benefit them more than a not at risk group of students. Although as a group students in the study "responded" to the intervention, without better control of the selection process it is difficult to determine if the selected group in this study represents a typical response to the intervention. Research on Tier 2 interventions is beginning to examine patterns of students who gain or fail to gain the full benefit of a Tier 2 intervention and in the case of evaluating a student for an MLD identifying students who don't respond to research-based Tier 1 and Tier 2 instruction (L. S. Fuchs, Fuchs, & Compton, 2012). Such investigations will further our understanding of estimates of treatment effect for the range of students who receive intervention services. Further exploration of this issue is warranted and needed in particular as we attempt to construct the strongest possible instructional experience for students with learning disabilities (L. S. Fuchs & Vaughn, 2012). In addition, data for this study were collected only at pre- and posttest. Thus we are not able to explore whether or not growth patterns differed across time between control and treatment groups.

Educational Implications and Future Research

Encouragingly, the findings from this study fit within a pattern of promising results generated by other intervention programs focused on the development of early mathematical knowledge. Commonalities across these programs are significant and offer guidance both for future development of and research into effective programs. First, there is a concentrated effort to reduce the amount of content coverage to provide a more in-depth and focused concentration on developing whole number understanding. For example, L. S. Fuchs et al. (2005) focused on 17 key first grade topics in the development of whole number understanding. Similarly, Dyson et al. (2013) focused on critical kindergarten whole number competencies designed to build beginning number sense. Second, programs utilized a more systematic and explicit approach to instruction often through the use of suggested or scripted delivery of content. In part, this may help to ensure higher degrees of fidelity of implementation to the program and to key instructional practices (e.g., teacher models). The design of these programs to focus on whole number content and to employ a systematic and explicit instructional approach corresponds highly with approaches advocated for by a number of prominent documents (CCSS, 2010; Gersten, Beckmann, et al., 2009; NMAP, 2008) and national organizations (NCTM, 2006). Consistent findings and common features across these programs suggest strongly that school should consider evaluating potential intervention programs with these parameters in mind.

A common goal in education is eliminating the achievement gap. Although this is an oft-stated goal, it is rarely achieved (Starkey & Klein, 2008). In this study, the achievement gap was not fully bridged. The student sample in the study consisted of students who were serve risk (i.e., below the 10th percentile). Thus although serious efforts at both Tier 1 (ELM) and Tier 2 (ROOTS) helped reduce the gap, it may be students who fall below the 10th percentile or those who end up needing Tier 3 services (i.e., in this case students whose response to ELM and ROOTS was inadequate) need more than what is provided in typical intervention programs if they are to fully eliminate the achievement gap. To meet the needs of those students at greatest risk, Warren, Fey, and Yoder (2007) have advocated examining intervention effectiveness through the lens of instructional intensity and have identified a number of variables that impact intensity ranging from increasing teacher–student interactions, lesson length, and the number of lessons. For example, in response to previous iterations of the intervention not significantly impacting student outcomes across a number of studies Bryant and colleagues (2011) have systematically attempted to increase the intensity of their intervention by increasing the number of lessons provided to at-risk students. We have plans to examine a range of implementation variables that may help explain if and how ROOTS may increase student outcomes. Our observation and teacher reporting systems should provide insight on the particular teacher practices (e.g., teacher models) and content focus (e.g., base 10) that may mediate the association between ROOTS implementation and student outcomes.

A unique feature of this study was the delivery of a Tier 2 program in the context of research-based Tier 1 core instruction. A number of studies have examined mathematics programs or components of RtI models in isolation and in some cases have implemented a supplemental Tier 1 program followed by a Tier 2 intervention program for nonresponders (L. S. Fuchs et al., 2008). However, we know of no cases in which the impact of Tier 1 research-based core mathematics program and a Tier 2 intervention program have been studied in conjunction. It should be noted that we did not isolate the impact of either Tier 1 or 2, which would have required a study design including a condition where Tier 1 and Tier 2 consisted of business as usual. However, the inclusion of a research-based Tier 1 core program mirrors what we would consider best practices in school settings. That is, schools approach instruction at a systems level and implement systems of support rather than implementing singular programs in isolation. In reading, attempts to study and answer questions about the effectiveness of systems of support (i.e., RtI models) has been called for as the next step in advancing reading research and instruction (Baker et al., 2010). Even though efforts in the area are just beginning, initial results are promising (Nelson-Walker et al., 2012). And although it may be that the current state of mathematics research is not yet developed to a level that allows the study of mathematics service delivery systems, we see this area as one that warrants increased attention and eventual study. In addition, we see research that includes data collection during the course of the intervention as promising. As we attempt to identify students who don't respond to Tier 1 and Tier 2 instruction, ongoing data collection is vital to researching and building flexible systems of support to address the needs of all learners.

As evidence mounts that a successful start in mathematics is paramount and that at-risk students have serious and significant achievement gaps as early as kindergarten entry, we see a continued need for intervention programs that immediately address the need of kindergarten students at risk for MLD. Although it may not be feasible to fully eliminate achievement gaps with only a Tier 2 intervention (L. S. Fuchs & Vaughn, 2012) given this study's finding that the ROOTS invention program reduced the gap by end of kindergarten, we see the need to examine programs that either provide a more intensive level of support within kindergarten or attempt to build systems of support across grades (e.g., K and 1) or across settings (e.g., prekindergarten to kindergarten). Given the promising findings in each of these age/grade levels (Bryant et al., 2011; Clements & Sarama, 2007; L. S. Fuchs et al., 2005; Klein, Starkey, Clements, Sarama, & Iyer, 2008), future research that examines multiyear interventions seems especially critical if we are to address long-standing achievement gaps. We believe that continued and focused research in the areas of intervention program development, service delivery models, and multiyear interventions shows promise in significantly affecting the long-term mathematics outcomes of at-risk students.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Scott Baker, Ben Clarke, and Hank Fien are eligible to receive a portion of royalties from the University of Oregon's distribution and licensing of certain ROOTS-based works. Potential conflicts of interest are managed through the University of Oregon's Research Compliance Services. An independent external evaluator and coauthor of this publication completed the research analysis described in the article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305K040081 and R305A080699 to the Center on Teaching and Learning at the University of Oregon and Pacific Institutes for Research. The opinions expressed are those of the authors and do not represent views of the institute or the U.S. Department of Education.

References

Agodini, R., Harris, B., Atkins-Burnett, S., Heaviside, S., Novak, T., & Murphy, R. (2009). *Achievement effects of four early elementary school math curricula: Findings from first graders in 39 schools* (Rep. No. NCEE 2009-4052). Washington, DC: National Center for Education Evaluation and Regional Assistance.

Baker, S., Gersten, R., & Lee, D.-S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal*, *103*(1), 51–73.

Baker, S., Fien, H., & Baker, D. (2010). Robust reading instruction in the early grades: Conceptual and practical issues in the integration and evaluation of tier 1 and tier 2 instructional supports. *Focus on Exceptional Children*, *42*(9), 1–20.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities*, *38*, 333–339. doi:10.1177/00222194050380040901

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *Elementary School Journal*, *108*, 115–130. doi:10.1086/525550

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children*, *78*, 7–23.

Burns, M. K., & Vanderheyden, A. (2006). Response to intervention to assess learning disabilities: Introduction to the special series. *Assessment for Effective Instructions*, *32*, 3–5. doi:10.1177/15345084060320010201

Carnine, D., Silbert, J., Kame'enui, E., & Tarver, S. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Pearson.

Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten curriculum. *Learning Disability Quarterly*, *31*(1), 11–20.

Chard, D. J., Clarke, B., Baker, S. K., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention*, *30*(2), 3–14, doi:10.1177/073724770503000202

Clarke, B., Baker, S. K., & Chard, D. J. (2008). Best practices in mathematics assessment and intervention with elementary students. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (5th ed., pp. 453–464). Bethesda, MD: National Association of School Psychologists.

Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, *33*(2), 234–248.

Clarke, B., Smolkowski, K., Baker, S. K., Fien, H., Doabler, C. T., & Chard, D. J. (2011). The impact of a comprehensive tier 1 core kindergarten program on the achievement of students at-risk in mathematics. *Elementary School Journal*, *111*, 561–584.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*, 136–163. doi:10.2307/30034954

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

Common Core State Standards. (2010). *Common Core State Standards Initiative: Designing common state assessment systems*. Retrieved from http://www.nga.org/Files/pdf/1004NGACCSSOASSESSMENTS.PDF

Coyne, M. D., Kame'enui, E. J., & Carnine, D. W. (2011). *Effective teaching strategies that accommodate diverse learners* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Cribbie, R. A., & Jamieson, J. (2000). Structural equation models and the regression bias for measuring correlates of change. *Educational and Psychological Measurement*, *60*, 893–907. doi:10.1177/00131640021970970

Davis, K., & Jungjohann, K. (2009). *Early learning in mathematics: Level K* (Unpublished curriculum). Eugene: University of Oregon, College of Education, Center on Teaching and Learning.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York, NY: Oxford University Press.

Doabler, C. T., Baker, S. K., Kosty, D. B., Smolkowski, K., Clarke, B., Miller, S. J., & Fien, H. (in press). Examining the association between explicit mathematics instruction and student mathematics achievement. *Elementary School Journal*.

Dougherty, B. J., Flores, A., Louis, E., & Sophian, S. (2010). *Developing essential understanding number and numeration for teaching mathematics in prekindergarten–grade 2* (Essential understanding series). Reston, VA: National Council of Teachers of Mathematics.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. doi:10.1037/0012-1649.43.6.1428

Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2013). *The value of replication for developmental science*. Manuscript submitted for publication.

Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. *Journal of Learning Disabilities*, *46*, 166–181. doi:10.1177/0022219411410233

Fixsen, D., Naoom, S., Blase, K., Friedman, R., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Pub. No. 231). Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003), Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, *18*, 157–171. doi:10.1111/1540-5826.00072

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*, 493–513. doi:10.1037/0022-0663.97.3.493

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2012). The early prevention of mathematics difficulty: Its power and limitations. *Journal of Learning Disabilities*, *45*, 257–269.

Fuchs, L. S., Fuchs, D., & Zumeta, R. O. (2008). A curricular-sampling approach to progress monitoring. *Assessment for Effective Intervention*, *33*, 225–233. doi:10.1177/1534508407313484

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness to intervention: A decade later. *Journal of Learning Disabilities*, *45*, 195–203.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, *114*, 345–362. doi:10.1037/0033-2909.114.2.345

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

Gersten, R. M., Beckmann, S., Clarke, B., Foegen, A., March, L., Star, J. R., & Witzel, B. (2009). *Assisting struggling with mathematics: Response to intervention (RTI) for elementary and middle schools* (Practice Guide Report No. NCEE 2009-4060). Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance.

Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *Journal of Special Education*, *33*, 18–28. doi:10.1177/002246699903300102

Gersten, R. M., Chard, D., 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

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly*, *19*, 173–180.

Griffin, S., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first learning of arithmetic to students at risk for school failure. In K. McGilly

(Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–49). Cambridge, MA: MIT Press.

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *NASSP Bulletin*, *89*(642), 24–46. doi:10.1177/019263650508964204

Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, *93*, 615–627. doi:10.1037/0022-0663.93.3.615

Hannan, P. J., & Murray, D. M. (1996). Gauss or Bernoulli? A Monte Carlo comparison of the performance of the linear mixed model and the logistic mixed model analyses in simulated community trials with a dichotomous outcome variable at the individual level. *Evaluation Review*, *20*, 338–352.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128. doi:10.3102/10769986006002107

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum.

Jordan, N., Kaplan, D., & Hanich, L. (2002). Achievement growth in children with learning difficulties in mathematics: Findings of a two-year longitudinal study. *Journal of Educational Psychology*, *94*, 586–597. doi:10.1037/0022-0663.94.3.586

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*, *22*, 36–46. doi:10.1111/j.1540-5826.2007.00229.x

Kame'enui, E. J., & Simmons, D. C. (1999). *The architecture of instruction: Towards successful inclusion of students with disabilities Adapting curricular materials for the inclusive classroom* (Vol. 1). Reston, VA: Council for Exceptional Children.

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, *1*, 155–178. doi:10.1080/19345740802114533

Kovaleski, J. F., & Black, L. (2010). Multi-tier service delivery. In T. A. Glover & S. Vaughn (Eds.), *The promise of response to intervention: Evaluating current science and practice* (pp. 23–56). New York, NY: Guilford.

Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial & Special Education*, *24*, 97–114. doi:10.1177/07419325030240020501

Lago, R. M., & DiPerna, J. C. (2010). Number sense in kindergarten: A factor-analytic study of the construct. *School Psychology Review*, *39*, 164–180.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York, NY: John Wiley.

Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*, *41*, 451–459.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, *42*, 306–321. doi:10.1177/0022219408331037

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York, NY: Oxford University Press.

Murray, D. M., Hannan, P. J., Pals, S. P., McCowen, R. G., Baker, W. L., & Blitstein, J. L. (2006). A comparison of permutation and mixed-model regression methods for the analysis of simulated data in the context of a group-randomized trial. *Statistics in Medicine*, *25*, 375–388. doi:10.1002/sim.2233

National Center for Education Statistics. (2011a). *Common core of data: Public elementary/secondary school universe survey 2009–10* (Version 1a) [Data file and documentation]. Retrieved from http://nces.ed.gov/ccd/pubschuniv.asp

National Center for Education Statistics. (2011b). *NAEP 2011 mathematics: Report card for the nation and the states*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

National Council of Teachers of Mathematics. (2000). *Standards 2000 project*. Retrieved from http://standards.nctm.org/document/index.htm

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Retrieved from http://www.nctm.org/standards/focalpoints.aspx?id=282

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education. DOI: 10.3102/0013189X08329195

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: Mathematics Learning Study Committee.

National Research Council. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.

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

Nelson-Walker, N., Fien, H., Smith, J., Baker, S., Smolkowski, K., & Kosty, D. (2012, February). *Systemic intervention in early reading: Impacts for students and teachers*. Paper presented at the National Association of School Psychologists annual conference, Philadelphia, PA.

Newman-Gonchar, R., Clarke, B., & Gersten, R. (2009). *A summary of nine key studies: Multi-tier intervention and response to interventions for students struggling in mathematics*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (2008). *TIMSS 2007 technical report*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.

PRO-ED. (2007). *Test of Early Mathematics Ability* (3rd ed.). Austin, TX: Author.

Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis* (3rd ed.). Boston, MA: McGraw-Hill.

SAS Institute. (2009). *SAS/STAT® 9.2 user's guide* (2nd ed.). Cary, NC: Author. Retrieved from the SAS Product Documentation website: http://support.sas.com/documentation/index.html

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177. doi:10.1037/1082-989X.7.2.147

Schmidt, W., Houang, R., & Cogan, L. (2002). A coherent curriculum: The case of mathematics. *American Educator*, *26*(2), 1–18.

Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York, NY: Guilford.

Simmons, D. C., Kame'enui, E. J., Harn, B., Coyne, M. D., Stoolmiller, M., Santoro, L. E., . . . Kaufman, N. K. (2007). Attributes of effective and efficient kindergarten reading intervention: An examination of instructional time and design specificity. *Journal of Learning Disabilities*, *40*, 331–347. doi:10.1177/00222194070400040401

Starkey, P., & Klein, A. (2008). Sociocultural influences on young children's mathematical knowledge. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives on mathematics in early childhood education* (pp. 253–276). Charlotte, NC: Information Age.

van Belle, G. (2008). *Statistical rules of thumb* (2nd ed.). New York, NY: John Wiley.

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*, 70–77. doi:10.1002/mrdd.20139

What Works Clearinghouse. (2011). *What Works Clearinghouse: Procedures and standards handbook* (Version 2.1). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Wu, H. (2005). *Must content dictate pedagogy in mathematics education*. Unpublished manuscript.