Check for updates

Article

Hammill Institute on Disabilities

Assessment for Effective Intervention
2014, Vol. 39(2) 99–111
© Hammill Institute on Disabilities 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1534508413511848
aei.sagepub.com
SAGE

Examining Teachers' Use of Evidence-Based Practices During Core Mathematics Instruction

Christian T. Doabler, PhD[1], Nancy J. Nelson, PhD[1], Derek B. Kosty, BS[2], Hank Fien, PhD[1], Scott K. Baker, PhD[1,3], Keith Smolkowski, PhD[2], and Ben Clarke, PhD[1]

Abstract

The extent to which teachers implement evidence-based practices, such as explicit instruction, is critical for improving students' mathematics achievement. The purpose of this study was to examine the effect of the kindergarten *Early Learning in Mathematics* (ELM) curriculum on teachers' use of explicit mathematics instruction in core educational settings. Observation data for the study were collected during a randomized controlled trial designed to investigate the efficacy of the ELM curriculum. A multifaceted observation system was used to examine teachers' provision of high-quality and intensive instructional interactions during core mathematics instruction. A total of 379 observations were conducted in 129 classrooms (68 treatment and 61 comparison), involving approximately 2,700 students from 46 schools in Oregon and Texas. Results indicate that ELM classroom teachers delivered significantly higher rates of practice opportunities for individuals and groups of students compared with comparison classroom teachers who implemented standard district mathematics instruction. Implications for instruction are discussed.

Keywords

evidence-based practices, core mathematics instruction, explicit and systematic instruction, instructional interactions, treatment intensity, observation systems

In the current zeitgeist of bringing evidence-based practices into classrooms (Chard, 2004; Clements, Agodini, & Harris, 2013; Cook & Cook, 2011; R. M. Gersten & Dimino, 2001; Landrum & Tankersley, 2004; Vaughn & Dammann, 2001), efforts to support changes in mathematics instruction have begun to materialize. A major focus of these change efforts have concentrated on the prevention of mathematics problems in the early grades, an emphasis supported by a growing body of evidence documenting that students who struggle early with the foundations of mathematics are far more likely than other students to experience persistent mathematics difficulties (Bodovski & Farkas, 2007; Morgan, Farkas, & Wu, 2009). Collectively, this research suggests an urgent need to prevent mathematics difficulties as early as kindergarten. In the absence of effective mathematics instruction, many students will experience early and persistent difficulties in mathematics and thus will struggle to acquire mathematical proficiency. A suggested solution for allowing all students, including those at risk for difficulties in mathematics, to reach their mathematical potential is to ensure high-quality implementation of evidence-based practices, such as explicit mathematics instruction, in classrooms as early as kindergarten.

Explicit Mathematics Instruction

Explicit instruction is defined as a systematic and structured instructional approach for effectively and efficiently teaching foundational concepts and skills (Carnine, Silbert, Kame'enui, & Tarver, 2004). At its core, explicit instruction has a strong focus on learning for mastery and clear delineation of roles for teachers and students during instruction (Archer & Hughes, 2010; Doabler & Fien, 2013; Hudson & Miller, 2006). In light of empirical evidence generated by recent meta-analyses on the attributes of effective mathematics interventions (Baker, Gersten, & Lee, 2002; R. Gersten et al., 2009), explicit instruction has been recommended to help teachers deliver high-quality mathematics instruction when teaching struggling learners. While the

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

Corresponding Author:
Christian T. Doabler, Center on Teaching and Learning, University of Oregon, 1600 Millrace Drive, Suite 108, Eugene, OR 97403-1995, USA.
Email: cdoabler@uoregon.edu

literature base clearly suggests the value of an explicit instructional approach in the context of small-group interventions, explicit instruction also appears to have promising effects in educational settings that deliver core mathematics instruction (Agodini & Harris, 2010; Clarke et al., 2011; Clements et al., 2013). Here we define core instruction as mathematics instruction that focuses on the range of mathematical standards students are expected to learn and know at each grade level (e.g., the Common Core State Standards for Mathematics, 2010). This definition recognizes that core instruction takes place in general education settings and is commonly delivered by teachers using commercially available core programs. Because core mathematics instruction represents the primary source of mathematics for many students, including students struggling with mathematics (Fuchs & Vaughn, 2012), the implementation of an explicit, systematic core mathematics curriculum is critical for preventing mathematics difficulties (MD).

Empirical Support for Explicit Core Mathematics Curriculums

A core mathematics curriculum represents an instructional foundation for students' early mathematical learning because it determines when and how well they will progress through content. A well-designed program will ensure appropriate pacing, include teacher modeling of new concepts, provide ample opportunities for guided and independent practice, and integrate academic feedback, among other critical features (Doabler, Fien, Nelson-Walker, & Baker, 2012). Moreover, it will form a vital link between foundational and more advanced concepts (Schmidt, Houang, & Cogan, 2002). For example, a kindergarten core curriculum will devote strong attention at the start of the academic year to the critical aspects of number, such as magnitude comparison, and then as students acquire initial proficiency progress to more advanced topics, such as place value.

An emerging line of research on the impact of core mathematics programs demonstrates promising effects of core mathematics curriculums that incorporate explicit instruction on student mathematics achievement. For example, Agodini and Harris (2010) investigated the effectiveness of four, commercially available first-grade elementary mathematics curricula in more than 100 schools and found that the mathematics achievement of students in schools that were randomly assigned to the two teacher-directed curriculums (i.e., *Saxon Math* and *Math Expressions*) was significantly greater than the math achievement of students in schools that used the student-centered curriculums (i.e., *Scott Foresman* and *Investigations*).

A major finding of our own work on core mathematics instruction is that an explicit, core mathematics curriculum is critical for many students, particularly for students with or at risk for MD. In a recent efficacy trial funded through the Institute of Education Sciences (Clarke et al., 2011), our research team investigated the efficacy of the Early Learning in Mathematics (ELM) curriculum. ELM is a yearlong, core mathematics program that uses an explicit instructional framework and focuses on the kindergarten standards identified in the Common Core State Standards for Mathematics (2010). In the efficacy trial, 65 kindergarten classrooms were randomly assigned to either treatment (ELM) or comparison (standard district practices) conditions. Findings suggest that at-risk kindergartners in the treatment classrooms made significant gains in mathematics achievement across the year relative to their at-risk peers in comparison classrooms, while typically achieving students in ELM classrooms made gains on par with their typically achieving peers in comparison classrooms. In ELM classrooms, at-risk students also demonstrated greater gains than their typical achieving peers and thus reduced the achievement gap. In the current study, we examine the extent to which ELM supports teachers in delivering explicit mathematics instruction.

Distinguishing Features of Explicit Mathematics Curriculums

A common feature of effective, core mathematics programs, such as Saxon Math, Math Expressions, and ELM, is incorporation of *explicit* instructional design principles that have been empirically validated to improve student mathematics achievement (Bryant et al., 2008; Doabler et al., 2012). An explicit curriculum offers opportunities for teachers to (a) facilitate high-quality practice opportunities for students, (b) deliver clear and consistent demonstrations of new and complex mathematical concepts and skills, and (c) provide timely academic feedback to address student errors and misconceptions (Doabler et al., 2012). For example, an explicit math curriculum will provide specific teaching directions on how to demonstrate and explain key math concepts, offer students opportunities to practice taught concepts, and deliver immediate feedback.

When well designed, a core curriculum represents an optimal platform for teachers to deliver high-quality instruction that results in students' construction of deep and robust mathematical knowledge. It also serves as a mechanism to increase the intensity of instruction for struggling students. Features of instructional intensity in the context of core mathematics instruction include the active ingredients that are hypothesized to promote student achievement (Warren, Fey, & Yoder, 2007). In this study, such ingredients are operationalized as student–teacher interactions that are nested within kindergarten mathematics instruction. We believe that an explicit, core mathematics curriculum can increase instructional intensity by providing opportunities for teachers to facilitate frequent, high-quality instructional interactions.

In the early grades, instructional interactions commonly entail students verbalizing their mathematical thinking and understanding (Cirillo, 2013; Clements et al., 2013). Such verbalizations are essential in the early development of mathematical learning because they allow students to use productive mathematical discourse at a time when other modes of responding, such as independent written exercises, are not yet instructionally appropriate (Clements et al., 2013; Doabler, Baker et al., in press). For example, a first-grade student is more apt to verbalize his or her solution method for solving a word problem than to write the specific steps that lead to the problem's answer.

Math verbalizations occur through two mediums: group responses and individual responses. Group responses offer a way for teachers to keep all students engaged in instruction and check students' understanding of specific content (Blackwell & McLaughlin, 2005). For example, a group response might entail 25 kindergartners stating in unison how the identity property applies when adding zero to a whole number (e.g., "the number stays the same"). Individual responses entail one student verbalizing his or her mathematical knowledge. When interspersed with group response opportunities, individual responses can be particularly effective for reducing potential misconceptions and determining whether specific individuals understand the target concepts.

Purpose of the Study

The purpose of the current study is to extend our efficacy research on the ELM kindergarten curriculum. In our initial efficacy study, Clarke et al. (2011) found a positive impact of ELM on the mathematics achievement of kindergarten students with and without MD. The current study continues this line of efficacy research but with a specific focus on the impact of ELM on teaching behavior. In this study, it was hypothesized that the ELM curriculum would help teachers implement explicit, core mathematics instruction. Specifically, we expected teachers in classrooms that implemented ELM to deliver more intensive and higher quality explicit instruction than teachers in classrooms that delivered standard district instructional practices.

Our hypothesis is anchored to the growing knowledge base of effective mathematics instruction (R. Gersten et al., 2009; National Mathematics Advisory Panel, 2008; National Research Council [NRC], 2001) and the findings of a recently conducted efficacy trial on a multi-tiered, first-grade literacy intervention, *Enhancing Core Reading Instruction* (ECRI; Nelson-Walker et al., 2013). Nelson-Walker and colleagues investigated the efficacy of the ECRI intervention for changing teachers' instruction. A key finding from the Nelson-Walker et al. study is that ECRI increased teachers' use of explicit instructional techniques, suggesting that an explicit, systematic intervention can enhance the quality and intensity of core reading instruction that students receive in general education classrooms. We believe that this finding may generalize to the current study because many of the architectural features of the ECRI intervention are analogous to those of the ELM curriculum. Both use an explicit instructional approach and offer frequent occasions for teachers to provide high-quality demonstrations of key concepts and skills. ELM and ECRI also facilitate opportunities for teachers to increase the intensity of reading and mathematics instruction, respectively, by fostering structured practice for students, including opportunities for students to verbalize their understanding of reading and mathematics content.

In this study, we conceptualized *instructional intensity* as the frequency with which explicit instructional interactions occurred during core mathematics instruction. Comprising such interactions were teachers' demonstrations of mathematical content, practice opportunities for students, and teacher-provided academic feedback. We also included student errors as part of these instructional interactions because they serve as a proxy indicator of instructional effectiveness and student accuracy. We anchored this conceptualization of instructional intensity to a framework proposed by Warren et al. (2007). Their framework coupled with findings from recent research reported by Clements et al. (2013) highlights the importance of defining the active ingredients of classroom instruction, such as the instructional interactions that occur between teachers and students. It also suggests the use of a frequency-based observation instrument as a measurement approach for precisely capturing indicators of treatment intensity.

We conceptualized *instructional quality* as the clarity and timeliness of instructional interactions, effective classroom management techniques, and dimensions of a supportive learning environment. This conceptualization of instructional quality has its roots in several programs of observation research, including the work of Pianta and Hamre (2009), and Englert and colleagues (cf. Englert, Tarrant, & Mariage, 1992). Because a defining characteristic of instructional intensity and quality is their observable nature, we used a multifaceted observation system to investigate the efficacy of the ELM curriculum for increasing teachers' use of explicit instructional interactions during core mathematics instruction.

Two research questions were addressed in this study:

Research Question 1: Do ELM teachers deliver more intensive explicit instruction than teachers in classrooms that implemented standard district practices?

Research Question 2: Do ELM teachers demonstrate higher quality explicit instruction than teachers in classrooms that implemented standard district practices?

Method

Research Design

The current study was part of a larger efficacy trial funded by the Institute of Education Sciences that was designed to investigate the efficacy of the ELM kindergarten program when implemented under rigorous experimental conditions. The ELM study used a randomized controlled design, nesting 2,681 students within 129 kindergarten classrooms. Blocking on schools, classrooms were randomly assigned to either a treatment or comparison condition. Sixty-eight treatment classrooms implemented the ELM program for the entire school year, whereas 61 comparison classrooms implemented business-as-usual teaching practices. Because the current study focuses on teacher behavior as the outcome, the primary unit of analysis is the teacher.

Classroom Recruitment in the ELM Efficacy Trial

The principal investigators and key personnel of the ELM project conducted all recruitment efforts for the larger efficacy trial. These efforts entailed contacting district leaders of public, private, and charter school districts in Dallas, Texas, and areas of western Oregon. District leaders were provided information on the study's research aims and activities. Interested district leaders then identified potential schools for participation, namely those that contained large percentages of students in need of intensive instructional support in mathematics. Public schools targeted for recruitment were those that received Title 1 funding. Principals and the kindergarten teachers of schools in all three sectors (i.e., public, private, and charter) were then contacted. All kindergarten teachers in each participating school were eligible to participate in the study.

Participants

Forty-six schools from seven school districts in Oregon and Dallas, Texas, participated in the ELM study. In Oregon, the ELM efficacy trial was conducted during the 2008–2009 school year. Texas schools participated in the efficacy trial during the following school year 2009–2010. Of the 46 schools, 32 were public institutions, 11 were private, and 3 were charter schools. The participating schools were located in urban and suburban areas in each state. From the 46 schools, 130 classrooms were eligible for randomization. Because we blocked on schools, 68 of the 130 classrooms were randomly assigned to the ELM program and 62 were randomly assigned to the comparison condition. One comparison classroom in Oregon was dropped from the analysis because of an inadvertent change in condition at the start of the study. The analytic sample for this study involved 129 kindergarten classrooms (68 ELM; 61 comparison).

Table 1 provides descriptive information about the classrooms and teachers by condition and region. Of the 129 classrooms in the analysis, 112 provided a full-day kindergarten program and 17 provided a half-day program. All half-day classrooms were located in Oregon. While the majority of the classrooms provided instruction 5 days per week, one full-day classroom in Oregon met 4 days per week. All math instruction was provided in English; however, 17 bilingual education classes were part of the sample. Average class size for treatment and comparison classrooms was $M = 21.3$ ($SD = 3.7$) and $M = 20.2$ ($SD = 3.7$), respectively. The 129 participating classrooms were taught by 130 teachers. One comparison classroom had two teachers, each working a half-day schedule. All teachers participated for the duration of the study, and thus the outcomes of the study were not affected by attrition (What Works Clearinghouse [WWC], 2013).

Within the 129 classrooms were 2,681 students, of which 1,448 and 1,233 were in ELM and comparison classrooms, respectively. Student demographic data were only available for those students who attended participating public schools. In the 32 public schools, an average of 76% of the student population qualified for free or reduced price lunch programs. Students in Oregon schools were Hispanic (36%), Black (2%), White (56%), Asian and Pacific Islander (5%), and American Indian (1%). In Texas schools, students were Hispanic (69%), Black (29%), White (1%), Asian and Pacific Islander (<1%), and American Indian (<1%).

ELM Intervention

ELM is a core kindergarten mathematics curriculum designed to promote students' development of mathematical proficiency in the concepts and skills identified in the Common Core State Standards for Mathematics (2010). ELM includes 120 core daily lessons that are approximately 45 min in duration and designed to be delivered in whole-class settings. As a core curriculum, ELM attends to six mathematics domains: (a) counting and cardinality, (b) operations and algebraic thinking, (c) number and operations in base 10, (d) measurement and data, (e) geometry, and (f) precise mathematics vocabulary. To meet the instructional needs of all students, mathematics content is explicitly introduced in each lesson, and systematically reviewed and extended across lessons (Coyne, Kame'enui, & Carnine, 2011). ELM teachers are expected to model and demonstrate what they want students to learn, and provide specific and frequent feedback to students during learning activities. Teachers are also expected to facilitate frequent and deliberate opportunities for students to practice key mathematics concepts, such as opportunities for students to verbalize their mathematical thinking and understanding.

Each ELM lesson incorporates four to five activities, with the first activity typically introducing or reviewing a

Table 1. Descriptive Information for Classrooms and Teachers by Region and Condition.

	Treatment		Comparison		
Classroom and Teacher Characteristics	Oregon	Texas	Oregon	Texas	Total
Number of classrooms	34	34	30	31	129
School type					
Public	34	17	30	16	97
Private	0	11	0	9	20
Charter	0	6	0	6	12
Program structure					
Full-day program	26	34	21	31	112
Half-day program	8	0	9	0	17
Teacher gender					
Female	32	33	31	31	127
Male	2	1	0	0	3
Teacher age (35 years or older)	18	24	16	21	79
Teacher ethnicity					
White	33	17	26	16	92
Hispanic	0	11	2	9	22
African American	0	5	0	6	11
Native American	0	0	1	0	1
Asian American	0	1	0	0	1
Years teaching kindergarten (4 or more years)	16	22	18	17	73
Teacher education					
Master's degree	22	6	14	7	49
Special education	2	3	0	1	6
Completed 3 or more college math courses	12	6	7	5	30
Completed college algebra	17	24	13	14	68
Number of students *M* (*SD*)	22.4 (3.9)	20.2 (3.2)	21.7 (3.4)	18.8 (3.4)	20.8 (3.7)

Note. The 129 participating classrooms were taught by 130 teachers. One classroom had 2 teachers, each working a half-day schedule.

mathematical concept or skill that is central to the lesson's overall objective. For this part of the lesson, the teacher provides concrete examples and makes explicit the focus of the activity's targeted content. For example, an initial activity might have a teacher explain the concept of addition and demonstrate the procedure of adding one to a number. The second and third activities in ELM lessons involve either an extension of the first activity or a review of previously learned material. The fourth activity often targets previously learned material from a different math domain. If, for example, the first three activities focus on teen numbers (i.e., number and operations in base 10), the fourth activity will address material related to geometry or measurement. The last activity entails a paper–pencil review. Facilitated by the teacher, this worksheet activity provides children with a cumulative review of the lesson's content.

Professional development. Across the ELM study, treatment teachers received 4 days of professional development related to program implementation and kindergarten mathematics instruction. Each curriculum workshop lasted 6 hr and was conducted by the program's lead author. The first curriculum workshop took place prior to the start of the school year and focused on three key elements: (a) the research-based principles of math instruction, (b) the instructional design and delivery features of the ELM curriculum, and (c) an overview of Lessons 1 to 30. During each workshop, participating teachers were provided opportunities to deliver sample lessons and receive feedback on their teaching from members of the professional development team. ELM teachers received three follow-up workshops distributed across the school year. These workshops reviewed instructional design elements shared in previous sessions and provided treatment teachers with an in-depth overview of the remaining lessons of the ELM curriculum.

Comparison Classrooms

Mathematics instruction in the comparison condition consisted of standard district practices. Teachers in comparison classrooms used a variety of instructional materials, including teacher-developed activities and a number of commercially available curriculums. Surveys administered in the larger efficacy trial (Clarke et al., 2011) indicated that mathematics materials used in comparison classrooms varied within participating districts and schools. The most widely used curriculums were *Everyday Mathematics, Houghton Mifflin, Scott Foresman*, and *Bridges in Mathematics*. The instructional focus in comparison classrooms also varied. Some teachers emphasized whole number concepts, while others focused primarily on patterning and particular aspects of geometry and measurement. A variety of different mediums were used to deliver instruction in the comparison classrooms, including whole-class and center-based activities.

Observation Measures

A multifaceted observation system, which was comprised of four observation instruments, was used to measure the efficacy of the ELM curriculum for increasing the intensity and quality of explicit mathematics instruction. The first instrument in the observation system was the Classroom Observations of Student–Teacher Interactions–Mathematics (COSTI-M; Doabler, Baker, et al., in press), an observational tool designed to document the frequency of explicit instructional interactions that occur between teachers and their students. The COSTI-M was used during observations in all participating kindergarten classrooms, regardless of condition, to measure the intensity of explicit instruction.

To measure the quality of explicit instruction, we designed two moderate-inference observation instruments: Quality of Classroom Instruction (QCI; see Doabler, Baker, et al., in press) and Ratings of Classroom Management and Instructional Support (RCMIS; Doabler & Nelson-Walker, 2009). The QCI and RCMIS were administered in separate geographical regions in the current study: The QCI was used alongside the COSTI-M in all Oregon classrooms, whereas the RCMIS was used alongside the COSTI-M in all Texas classrooms. As described earlier, the Texas classrooms participated in the ELM study 1 year after their Oregon counterparts. Following the first year of the ELM efficacy trial, the research team elected to revise the QCI to serve as a more global instrument of instructional quality, capturing features of classroom management, the delivery of instruction, and the learning environment. This revision resulted in a new, refined instrument named the RCMIS. Thus, the ELM research team used the RCMIS in the Texas classrooms because it permits a more comprehensive focus on instructional quality than can be examined using the QCI. A fourth instrument named the ELM Fidelity of Implementation instrument was used in all classrooms to measure implementation fidelity and potential treatment diffusion.

COSTI-M. The COSTI-M is a modified version of a direct observation instrument designed by Smolkowski and Gunn (2012) for use during observations of early literacy instruction. This instrument was modified to document the frequency of explicit instructional interactions during kindergarten mathematics instruction. Observers used the COSTI-M to collect data on (a) teacher demonstrations, (b) teacher-provided academic feedback, (c) group responses, (d) individual responses, (e) student errors, and (f) other forms of student responses. Teacher demonstrations were defined as a teacher's explanations, verbalizations of thought processes, or physical demonstrations of mathematics content. For example, observers coded a teacher model if a teacher used a "think-aloud" to explain how to solve an addition word problem. Academic feedback was defined as a teacher's error correction or a response affirmation to a preceding student response. For example, an observer coded academic feedback if a teacher corrected a student's mistake by restating the steps involved in a mathematical procedure. Observers also coded academic feedback if a teacher affirmed a correct response by a group of students.

Group responses were defined as two or more students verbalizing their mathematical thinking or understanding in unison. For instance, observers coded a group response if an entire class stated the answer to a basic number combination. An individual response was defined as one student verbalizing or physically demonstrating the answer to a mathematical problem. Group and individual responses were only coded if elicited by the teacher. This way, observers avoided coding extraneous conversation (e.g., student "call-outs") and captured interactions prompted by the teacher. Observers also documented "other" practice opportunities, including group written exercises, use of math representations by multiple students, and peer–partner learning. An example of an "other" response is an entire class of students writing the numeral five on individual dry-erase boards. Finally, observers coded errors made during group and individual responses. For example, if a student verbally miscounted a set of objects during a rational counting activity, observers coded an individual or group response (depending on the number of students responding), followed by a student error.

Individual response opportunities and academic feedback were modestly stable over time (intraclass correlation coefficients [ICCs] = .34 and .35, respectively), suggesting that three observations per teacher-provided reasonable estimates of these behaviors (see Doabler, Baker, et al., in press). Stability ICCs for other COSTI-M behaviors range from .13 to .19, suggesting less stable behaviors that may require more than three observations per teacher to obtain reasonable estimates of the behavior. The COSTI-M is reported to have preliminary evidence of predictive validity with the *Test of Early Mathematics Ability–3rd Edition* (TEMA-3), a broad, standardized measure of mathematics achievement (p = .004, pseudo-R^2 = .08), and a battery of early mathematics curriculum based measures (p = .017, pseudo-R^2 = .05; see Doabler, Baker, et al., in press).

The frequency of instructional interactions documented by the COSTI-M served as indicators of *instructional* *intensity*. Instructional intensity data are average rates per minute of teacher demonstrations, group responses, individual responses, and other forms of student responses. Mean rates of these four COSTI-M behaviors were calculated by dividing the frequency of each behavior in an observed lesson by the duration of the observation in minutes. We also computed the proportion of student responses in which students provided incorrect answers (i.e., accurate vs. inaccurate responding) and the proportion of incorrect student responses followed by academic feedback (i.e., teacher corrective feedback). These two conditional probabilities, along with the mean rates for each of the four instructional interactions, were among the study's dependent variables.

QCI observation instrument. The QCI was designed to measure the quality of explicit instruction in Oregon classrooms. Observers used the QCI alongside the COSTI-M to collect data on eight aspects of explicit instruction: (a) teacher modeling, (b) instructional pacing, (c) response time, (d) transitions between activities, (e) student engagement, (f) learning success, (g) checks of student understanding, and (h) academic feedback. Observers completed the QCI at the conclusion of each observation session, using a 3-point rating scale to rate each instructional aspect. On the 3-point scale, a rating of 1 represented the lowest score and a rating of 3 represented the highest score. The internal consistency of the QCI items was high, with Cronbach's alpha equal to .94. The QCI is reported to capture fairly stable ratings of instructional quality (ICC = .35), suggesting that three observations per teacher provided reasonable estimates of the construct. The QCI also demonstrates preliminary evidence of predictive validity with TEMA-3 (p = .014, pseudo-R^2 = .14; see Doabler, Baker, et al., in press). Reported QCI scores represent the average item score across three observation occasions. This mean item score was used as the dependent variable of *instructional quality* for Oregon classrooms in the study's analyses.

RCMIS observation instrument. The RCMIS was designed as a broad measure of instructional quality and used in tandem with the COSTI-M in Texas classrooms. The RCMIS is comprised of 11 items that target general features of mathematics instructional quality, including classroom management techniques, delivery of instruction, and the learning environment. To rate the quality of each item, observers used a 4-point rating scale, with scores of 1–2 representing the lower quality range and 3–4 representing the upper quality range. Observers relied on a detailed scoring rubric to differentiate between scores. Observers in the Texas classrooms completed the RCMIS at the conclusion of each classroom observation. Internal consistency of the RCMIS items was high, with Cronbach's alpha equal to .92. The RCMIS is reported to capture fairly stable ratings of instructional quality (ICC = .33), suggesting that three observations per teacher provided reasonable estimates of the construct. The RCMIS also demonstrates preliminary evidence of predictive validity with the TEMA-3 (p = .039, pseudo-R^2 = .05; see Doabler, Baker, et al., in press). Reported RCMIS scores represent the average item score across three observation occasions. This mean item score was used as the dependent variable of *instructional quality* for Texas classrooms in the study's analyses.

ELM fidelity of implementation instrument. For each activity within an ELM lesson, teachers' adherence to the curriculum was rated on a scale ranging from 0 (*did not implement*), 0.5 (*partial implementation*), to 1.0 (*full implementation*). Implementation fidelity in ELM classrooms was monitored in the fall, winter, and spring during each experimental year. A total of 179 curriculum-specific fidelity checks were conducted in the treatment classrooms. Results indicate that, on average, ELM teachers implemented the curriculum with moderate levels of implementation fidelity across each observation time point: fall (M = .86, SD = .13), winter (M = .87, SD = .15), and spring (M = .87, SD = .14). To protect against contamination of comparison classrooms by the ELM curriculum, the observation instrument included one item that measured whether comparison teachers implemented ELM. No evidence of contamination between ELM and comparison classrooms was observed.

Observation Procedures

Classroom observations. Observations of the 129 participating classrooms were conducted in the fall, winter, and spring of each school year. Each observation round lasted approximately 2 weeks, with 6 weeks separating each round. Because study classrooms were recruited in two separate waves, observers conducted observations in Oregon classrooms during the 2008–2009 school year. Texas classrooms were observed during the 2009–2010 school year. Together, a total of 379 observations were conducted. All classrooms were observed 3 times during the school year, with eight observations missed due to scheduling conflicts or teacher absences.

All classroom observations were scheduled in advance and conducted during the entire core mathematics instruction time period. The average observation length by condition in minutes was 38.0 (SD = 8.5) ELM and 38.7 (SD = 12.5) comparison. Observations were not scheduled according to the specific content planned for instruction (e.g., number and operations in base 10), or a particular day in the instructional sequence (e.g., introductory lesson).

Observations were conducted using the COSTI-M, the ELM fidelity measure, and one of two measures of instructional quality, depending on the region. Oregon observers used

Table 2. Effects of Intervention Condition on Instructional Intensity and Quality of Explicit Instruction.

Observer Ratings	ELM *M* (*SD*)	Comparison *M* (*SD*)	*t*	*p*	Hedges's *g*
Instructional intensity[a]					
Rate of teacher models	0.6 (0.3)	0.5 (0.3)	1.09	.278	0.20
Rate of group responses	1.3 (0.6)	0.8 (0.6)	5.09	<.001	0.91
Rate of individual responses	0.7 (0.4)	0.5 (0.3)	3.30	.001	0.57
Proportion of practice with errors	0.1 (0.0)	0.1 (0.1)	−0.73	.467	−0.22
Proportion of practice with feedback	0.4 (0.1)	0.5 (0.1)	−1.30	.195	−0.24
Quality of explicit instruction[b]					
Oregon	1.4 (0.4)	1.2 (0.5)	1.62	.110	0.41
Texas	3.0 (0.4)	2.9 (0.4)	1.42	.160	0.37

Note. ELM = early learning mathematics.
[a]Instructional intensity was measured using the Classroom Observation of Student–Teacher Interactions (COSTI). [b]Items were rated from 1 (low) to 3 (high) in Oregon and from 1 (low) to 4 (high) in Texas. Items and summary scores were averaged across three observation occasions. Tests of condition effects on instructional intensity included 61 comparison classrooms and 68 ELM classrooms (127 degrees of freedom). Tests of condition effects on quality of explicit instruction included 30 comparison classrooms and 34 ELM classrooms in Oregon (62 degrees of freedom) and 31 comparison classrooms and 34 ELM classrooms in Texas (63 degrees of freedom).

the QCI to rate instructional quality, while observers in Texas used the RCMIS. Observers coded instructional interactions using the COSTI-M throughout the entire instructional period and completed the fidelity measure, and the RCMIS or QCI at the conclusion of each observation. In Oregon, the COSTI-M was used in the fall, winter, and spring observation rounds. In Texas classrooms, however, the COSTI-M was used in the winter and spring observation rounds only.

Observation team and training. A total of 18 trained observers conducted observations in ELM and comparison classrooms. Across each school year, observers received approximately 14 hr of training from the project observation coordinator. Observation training sessions were scheduled just prior to each observation round to maximize interobserver reliability and minimize observer drift. As part of the observation training and a requirement before observing in study classrooms, observers coded a 5-min video of kindergarten math instruction and received feedback on the alignment between their codes and those of the observation coordinator. Once observers completed the video portion of the training, they conducted a paired observation with the observation coordinator in a participating classroom. All observers met the minimum interobserver agreement level of .85 for both coding checkouts.

Interobserver reliability. Interobserver reliability data were collected on 74 occasions across treatment and comparison classrooms and represented through ICCs. Reported interobserver reliability ICCs for the COSTI-M were .67 for teacher models, .92 for group responses, .95 for individual responses, .91 for other forms of responses, .84 for errors, and .90 for feedback. These ICCs are considered substantial to nearly perfect interobserver reliability (Landis & Koch, 1977). For the QCI, RCMIS, and ELM fidelity instrument, moderate to high interobserver reliability was found, with ICCs of .72, .61, and .83, respectively.

Statistical Analysis

The overall effects of intervention condition on observer ratings of the quality of explicit instruction and instructional intensity were assessed within a series of independent-sample *t* tests comparing ELM classrooms and comparison classrooms. Non-nested analyses were appropriate for this study given that classrooms were the units of randomization and analysis.

To ease the interpretation of results, we computed Hedges's *g* (Hedges, 1981) to describe the effect of the ELM condition on observer ratings of quality of explicit instruction and instructional intensity. Hedges's *g*, recommended by the WWC (2013), represents an effect size comparable with Cohen's *d* (Cohen, 1988), except that Cohen's *d* uses the sample standard deviation while Hedges's *g* uses the population standard deviation (Rosenthal & Rosnow, 2008).

Results

Table 2 presents classroom-level means and standard deviations of observer ratings of instructional intensity and quality of explicit instruction by intervention condition. This study aimed to test the hypotheses that teachers in the ELM condition would outperform teachers in the comparison condition on levels of instructional intensity and quality of explicit instruction. As described above, hypotheses were tested within a series of independent-sample *t* tests. Table 2 also summarizes the results of the *t* tests and includes Hedges's *g* effect sizes for each outcome.

In a preliminary analysis, we tested whether region of study (i.e., Oregon and Texas) moderated the impact of intervention condition on measures of instructional intensity in a series of two-way ANOVAs. These models included the main effect of condition, the main effect of region, and the interaction between condition and region. None of the interaction terms were statistically significant ($p > .05$ for all tests).

Contrasted with comparison classrooms, ELM classrooms provided higher rates of group practice opportunities ($t = 5.09$, $p < .001$, $g = 0.91$) and individual practice opportunities ($t = 3.30$, $p = .001$, $g = 0.57$). These results correspond to large and moderate effect sizes, respectively. We found no effects of condition on the rate of teacher demonstrations ($p = .278$), proportion of practice with errors ($p = .467$), proportion of practice followed by teacher corrective feedback ($p = .195$), or quality of explicit instruction in Oregon ($p = .110$) or Texas ($p = .160$).

Discussion

The ELM project sought to investigate the efficacy of a well-designed, explicit core mathematics curriculum for improving kindergarten mathematics achievement. In an earlier efficacy trial (Clarke et al., 2011), we examined the impact of the ELM curriculum on student mathematics achievement and observed significant effects for at-risk learners. The primary purpose of this study was to extend prior work by investigating the efficacy of the ELM kindergarten mathematics curriculum for increasing the quality and intensity of explicit instruction provided by teachers in Tier 1 settings. In this section, we summarize the study's findings and discuss implications for increasing the intensity and quality of core mathematics instruction. We also make suggestions for improving observation systems to better study relationships between evidence-based teaching practices and student outcomes.

Instructional Intensity

Instructional intensity data documented by the COSTI-M revealed significant differences between ELM and comparison classrooms in terms of teachers' provision of opportunities for students to practice with foundational mathematics content. Practice is a distinguishing characteristic of effective mathematics instruction and findings from a number of studies on mathematics interventions have demonstrated its particular importance for supporting students' development of mathematical proficiency (Clements et al., 2013; Doabler, Baker, et al., in press; Fuchs et al., 2010; R. Gersten et al., 2009). In this study, we operationalized practice as the instructional interactions between students and teachers around mathematical content. In ELM classrooms, teachers were trained to facilitate a high rate of instructional interactions to deeply engage students in foundational ideas of kindergarten mathematics. A primary medium of such interactions was mathematics verbalizations, which provided students opportunities to communicate their mathematical understanding, and explain and justify their methods for solving mathematics problems (Clements et al., 2013).

Instructional intensity data from our study suggest two important findings. First, ELM treatment classroom teachers delivered significantly higher rates of individual responses compared with comparison classroom teachers who implemented standard district mathematics instruction ($g = 0.57$). Previous research has found that frequent practice for individual students facilitates mathematical learning and improves student achievement (Doabler, Baker, et al., in press; Sutherland & Wehby, 2001). Individual responses served as a mechanism for teachers to increase the intensity of core mathematics instruction for targeted students. For example, in ELM classrooms, teachers elicited individual practice opportunities to determine whether targeted students understood the lesson objectives. The implication for this finding is that individual responses are a common and feasible way for teachers to facilitate student practice during core instruction. While there was substantial variability in how teachers in ELM and comparison classrooms delivered individual responses, attempts to increase such practice opportunities, particularly in comparison classrooms, would be a reasonable and valuable professional development objective. For instance, professional development opportunities could support teachers in interspersing individual responses within whole-class discussions and judiciously distributing these types of practice opportunities to students with more intensive instructional needs. Collectively, results from our ELM studies suggest that if teachers provided higher rates of individual responses during core instruction it could improve the outcomes of all students, including those at risk for MD.

A second important finding with respect to instructional intensity is that ELM classrooms provided significantly higher rates of group responses compared with comparison classrooms ($g = 0.91$). Group responses are an integral component of explicit and systematic instruction (Archer & Hughes, 2010; Hudson & Miller, 2006) and a defining feature of mathematics learning (Cirillo, 2013; NRC, 2001). Moreover, they represent an efficient teaching technique for increasing instructional intensity and the response rates of all students. When used effectively during core instruction, these types of practice opportunities can promote student participation and engagement. However, as evidenced by our findings, eliciting high rates of group responses can be quite challenging without an explicit and systematic curriculum. Use of the COSTI-M documented low rates and high variability in how comparison teachers facilitated group responses, which suggests that group responses are a difficult teaching technique to facilitate and manage in whole-class settings. This may be more evident in classrooms using curriculums that fail to use an explicit and systematic approach. Our data suggest that teachers benefit from having access to a core mathematics curriculum that integrates a structure for them to systematically provide a high rate of group practice opportunities for students.

Another explanation for why higher rates of group responses were observed in ELM classrooms than in comparison classrooms is that ELM teachers received professional development that targeted ways to facilitate whole-class mathematics discourse. During each curriculum workshop, ELM teachers were taught how to elicit group responses prescribed within the ELM curriculum (e.g., how to identify opportunities for additional practice) and how to use precise signaling techniques. Signaling techniques, such as a snap of the fingers or a clap of the hands, reduce confusion among students in terms of how and when to respond. They also minimize impulsive student responding and permit appropriate "thinking time" for students who require extra time to formulate answers (Kameenui & Simmons, 1990). Moreover, the use of a signaling technique offers a uniform way to promote high rates of choral responses (Baker, Fien, & Baker, 2010). Thus, in ELM classrooms, teachers received professional development experiences designed to intentionally increase rates of group responses.

Data documented using the COSTI-M revealed that the ELM intervention did not have a significant effect on the rate of teacher models. ELM and comparison teachers provided teacher models at roughly the same rate during core mathematics instruction (i.e., approximately one model every 2 min). This finding is surprising for two reasons. The first reason is that teacher models are a hallmark of explicit instruction and were infused throughout the ELM curriculum, with lessons containing specific wording for teachers to demonstrate and explain the mathematical content students are expected to learn. An expectation of the ELM intervention was that teachers would follow lesson scripting with high levels of implementation fidelity. A second reason as to why this non-significant is surprising is that ELM teachers received professional development on how to provide step-by-step demonstrations for solving problems and use "think-alouds" to verbalize their solution methods.

This non-significant finding, therefore, may be more representative of how teacher models were conceptualized in the study, and consequently how observers were trained to document these behaviors using the COSTI-M. For instance, observers did not differentiate simple teacher models from more complex ones during coding. For example, if a teacher stated a simple fact (e.g., "3 + 1 = 4"), observers were trained to code a single teacher model because it was a demonstration of an algorithm. Similarly, observers would also have coded a single model if a teacher explained how the concept of adding 1 to a number is the same as saying the next number in the count sequence. Not differentiating the complexity of teacher models is a plausible reason our analysis failed to detect significant differences between ELM and comparison classrooms with respect to rates of teacher models. Because of its scripted nature, we believe that the ELM curriculum offered teachers more opportunities to provide in-depth demonstrations and explanations than the commercially available mathematics curricula used in comparison classrooms. We base this conclusion on findings from previous curricular evaluation research (Bryant et al., 2008; Doabler et al., 2012), which suggest that opportunities for teachers to explicitly demonstrate mathematical content are largely absent from many U.S. market-leading programs.

We also did not detect significant differences between ELM and comparison classrooms in the proportion of practice with errors. The COSTI-M data revealed that approximately 10% of all practice opportunities consisted of incorrect student responses. These findings suggest that the vast majority (90%) of practice opportunities were accurate for students in the ELM and comparison classrooms. Critical to supporting students' acquisition of new mathematical ideas are instructional interactions that promote initial student success rather than initial frustration. Students' motivation to learn mathematics greatly depends on their opportunities for early success. Students who begin successfully are more likely to apply themselves in math tasks and develop intrinsic motivation in mathematics (R. Gersten et al., 2009; Gottfried, Marcoulides, Gottfried, & Oliver, 2013; NRC, 2001). While it appears that ELM and comparison classrooms were successful in providing accurate learning opportunities for students, the COSTI-M data do not capture where in the instructional sequence errors occurred (e.g., during guided or independent practice). Therefore, we cannot examine potential differences in error rates for practice that occurred during initial teaching activities as compared with practice that occurred later in the instructional sequence.

Finally, the ELM curriculum did not have a significant effect on the proportion of practice followed by corrective feedback. Results suggest that approximately half of all inaccurate practice opportunities in ELM and comparison classrooms were followed by academic feedback. Our results are consistent with the findings of Nelson-Walker et al. (2013), who found that the ECRI intervention did not produce significant effects on teachers' provision of academic feedback. One factor that may have contributed to this finding is the degree to which academic feedback is prescribed in the ELM curriculum. Although the curriculum provides opportunities for teachers to deliver academic feedback, these opportunities are often presented in the form of teaching recommendations rather than teaching requirements. For example, the curriculum reminds teachers to monitor students' understanding and address potential misconceptions at the end of each instructional interaction. Anecdotal evidence suggests that ELM teachers may have skipped many of these feedback opportunities because the recommendations were difficult to operationalize during the delivery of instruction.

In summary, we found that the ELM intervention was effective for increasing the rate of individual and group responding in treatment classrooms. These findings are important because they provide initial support for behaviors that teachers can actively manipulate to increase the intensity of mathematical experiences that students receive in core instructional settings.

Quality of Explicit Instruction

Two separate, moderate-inference observation instruments were used to measure teachers' provision of high-quality, explicit mathematics instruction in Oregon and Texas classrooms. Our analysis revealed that use of the ELM curriculum did not result in significant differences between treatment and comparison classrooms in quality ratings of explicit instruction. In fact, across instruments, regions, and conditions, instructional quality ratings were fairly similar. These findings are surprising because we expected ELM teachers to be poised to provide higher quality instruction, given that their training and access to a mathematics curriculum anchored to evidence-based instructional design principles. It is possible that our inability to observe significant differences in the quality of explicit instruction across conditions may be a function of a lack of statistical power. Because we used distinct observation instruments in Oregon and Texas classrooms, tests of conditional effects for instructional quality were conducted with just half of the sample of participating classrooms. Future research is therefore warranted to investigate potential differences in the quality of explicit instruction with a more robust sample. Second, limited differences in the quality of explicit instruction may be attributed to *observer overload* during classroom observations. Observers were expected to complete the COSTI-M, a curriculum-specific fidelity measure, and one of the quality measures (i.e., RCMIS, QCI) for each observation occasion. Thus, it is possible that observers were unable to attend to features of instructional quality while simultaneously coding implementation fidelity and instructional interactions. The implication is that additional training in completing multiple instruments or assigning observers to different observation instruments may allow for more precise measurements of instructional quality.

Limitations

Several limitations should be considered when interpreting the results of this study. First, only two observations were conducted in Texas classrooms using the COSTI-M. Additional data points are likely required to obtain reliable estimates of instructional intensity. In addition, we used separate measures of instructional quality in Texas and Oregon classrooms, and it is possible that the reduced sample size attenuated any real differences in the quality of explicit instruction between ELM and comparison classrooms. Moreover, because we did not document the professional development experiences that the comparison teachers received, it is difficult to juxtapose their training with the training provided to ELM teachers. ELM teachers received approximately 24 hr of professional development on program implementation and the use of evidence-based instructional practices each year of the study. Although comparison teachers reported that they received professional development during each year of the study, we do not know the extent of this training, and whether, like ELM professional development, it focused on the provision of intensive and high-quality explicit core mathematics instruction. Finally, rather than investigating how the intensity and quality of instruction provided in ELM and comparison classrooms may have varied by mathematical domain (e.g., number and operations), we examined kindergarten mathematics as one general category. Had we accounted for mathematical domain in our analyses, we might have observed different results for instructional intensity and quality.

Implications for Instruction

We believe that this study draws further attention to the need for the provision of evidence-based practices, such as teachers' use of explicit and systematic instruction, in core mathematics settings. Although our findings indicate that features of explicit and systematic mathematics instruction do occur in classrooms that use widely available core mathematics curriculums, the frequency with which these practices are observed is low and varies substantially. We suggest that professional development efforts be designed in ways that support teachers' regular use of explicit instruction during their core mathematics teaching. For example, training should focus on how to (a) use consistent mathematical language when introducing new and complex mathematical content, (b) re-voice a previous student response in precise mathematical language to reinforce learning, and (c) manage mathematics verbalizations in whole-class settings.

This study also has potential implications for curriculum development efforts (Clarke et al., in press; Doabler, Clarke, et al., in press; Kame'enui & Simmons, 1999). Instructional intensity and instructional quality data generated from standardized observation systems could guide developers in designing more focused and coherent mathematics programs and interventions. Specifically, curriculum developers could use observation data, such as those generated by the COSTI-M to judiciously embed individual and group practice opportunities into commercially available mathematics curricula. Further research is warranted, however, to determine the amount and type of practice students should receive and how that practice might change over time. Although our study did not observe significant differences in the quality of explicit instruction between ELM and comparison classrooms, research suggests that instructional quality is important (Pianta & Hamre, 2009). If studies using the RCMIS or QCI (or other instructional quality instruments) distinguish between classrooms in terms of teacher or student outcomes, observations using these instruments may also be used to inform features integrated into mathematics curriculums.

Conclusion

The recent release of the Common Core State Standards for Mathematics (2010) has sparked a compelling urgency for our nation's schools to promote mathematical proficiency for all students. If schools and teachers are going to meet this daunting challenge, they will need access to high-quality professional development that is aligned with evidence-based instructional practices and curricula. In this study, we found that a well-designed, explicit core kindergarten mathematics curriculum with strong evidence in terms of increasing student mathematics achievement also facilitated teachers' use of evidence-based instructional practices. Our findings suggest that explicit core mathematics curricula have the capacity to promote beneficial impact for all end-users.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this 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 Grant R305A080699 to the Center on Teaching and Learning at the University of Oregon.

References

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness*, *3*, 199–253.

Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. New York, NY: Guilford Press.

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.

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

Blackwell, A. J., & McLaughlin, T. F. (2005). Using guided notes, choral responding, and response cards to increase student performance. *International Journal of Special Education*, *20*, 1–5.

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, B. R., Bryant, D. P., Kethley, C., Kim, S. A., Pool, C., & You-Jin, S. (2008). Preventing mathematics difficulties in the primary grades: The critical features of instruction in textbooks as part of the equation. *Learning Disabilities Quarterly*, *31*, 21–35.

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

Chard, D. J. (2004). Maintaining the relationship between science and special education. *Journal of Learning Disabilities*, *37*, 213–217. doi:10.1177/00222194040370030501

Cirillo, M. (2013). *What are some strategies for facilitating productive classroom discussions?* Reston, VA: National Council of Teachers of Mathematics. Retrieved from http://www.nctm.org/uploadedFiles/Research_News_and_Advocacy/Research/Clips_and_Briefs/research%20brief%2020%20-%20strategies%20of%20discussion.pdf

Clarke, B., Doabler, C. T., Strand Cary, M., Kosty, D., Baker, S. K., Fien, H., & Smolkowski, K. (in press). Development and preliminary evaluation of a Tier-2 intervention targeting whole number concepts in first grade mathematics: Utilizing a theory of change to guide formative evaluation activities. *School Psychology Review*.

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

Clements, D. H., Agodini, R., & Harris, B. (2013, September). *Instructional practices and student math achievement: Correlations from a study of math curricula* (NCEE Evaluation Brief ED-04-CO-0112/0003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

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

Common Core State Standards. (2010). *Common core standards for English language arts & literacy in history/social studies, science, and technical subjects*. Retrieved from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Cook, B. G., & Cook, S. C. (2011). Unraveling evidence-based practices in special education. *The Journal of Special Education*. Advance online publication. doi:10.1177/0022466911420877

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.

Doabler, C. T., Baker, S. K., Kosty, D., 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*.

Doabler, C. T., Clarke, B., Fien, H., Baker, S., Kosty, D., & Strand Cary, M. (in press). The science behind curriculum development and evaluation: Taking a design science approach in the production of a tier 2 mathematics curriculum. *Learning Disability Quarterly*.

Doabler, C. T., & Fien, H. (2013). Explicit mathematics instruction: What teachers can do for teaching students with mathematics difficulties. *Intervention in School and Clinic*, *48*, 276–285.

Doabler, C. T., Fien, H., Nelson-Walker, N. J., & Baker, S. K. (2012). Evaluating three elementary mathematics programs for presence of eight research-based instructional design principles. *Learning Disability Quarterly*, *35*, 200–211.

Doabler, C. T., & Nelson-Walker, N. J. (2009). *Ratings of Classroom Management and Instructional Support* (Unpublished observation instrument). Center on Teaching and Learning, College of Education, University of Oregon, Eugene.

Englert, C. S., Tarrant, K. I., & Mariage, T. V. (1992). Defining and redefining instructional practice in special education: Perspectives on good teaching. *Teacher Education and Special Education*, *15*, 62–86. doi:10.1177/088840649201500203

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Fuchs, D., Hamlett, C. L., Cirino, P. T., & Fletcher, J. M. (2010). A framework for remediating number combination deficits. *Exceptional Children*, *76*, 135–156.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention a decade later. *Journal of Learning Disabilities*, *45*, 195–203. doi:10.1177/0022219412442150

Gersten, R., 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.

Gersten, R. M., & Dimino, J. (2001). The realities of translating research into classroom practice. *Learning Disabilities Research & Practice*, *16*, 113–122. doi:10.1111/0938-8982.00013

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2013). Longitudinal pathways from math intrinsic motivation and achievement to math course accomplishments and educational attainment. *Journal of Research on Educational Effectiveness*, *6*, 68–92. doi:10.1080/19345747.2012.698376

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

Hudson, P., & Miller, S. P. (2006). *Designing and implementing mathematics instruction for students with diverse learning needs*. Boston, MA: Pearson Education.

Kameenui, E. J., & Simmons, D. (1990). *Designing instructional strategies: The prevention of academic learning problems*. Columbus, OH: Merrill.

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.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. doi:10.2307/2529310

Landrum, T. J., & Tankersley, M. T. (2004). Science in the schoolhouse: The uninvited guest. *Journal of Learning Disabilities*, *37*, 207–212.

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

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

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, & B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education Division of Behavioral, and Social Sciences and Education. Washington, DC: National Academy Press.

Nelson-Walker, N. J., Fien, H., Kosty, D. B., Smolkowski, K., Smith, J. L. M., & Baker, S. K. (2013). Evaluating the effects of a systemic intervention on first-grade teachers' explicit reading instruction. *Learning Disability Quarterly*. Advance online publication. doi:10.1177/0731948712472186

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*, 109–119.

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

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

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the classroom observations of student–teacher interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly*, *27*, 316–328. doi:10.1016/j.ecresq.2011.09.004

Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education*, *22*, 113–121.

Vaughn, S., & Dammann, J. E. (2001). Science and sanity in special education. *Behavioral Disorders*, *27*, 21–29.

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. (2013). *What Works Clearinghouse: Procedures and standards handbook (Version 3.0)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.