The Effects of a Core Kindergarten Mathematics Program on the Mathematics Achievement of English Learners

Author

Christian T. Doabler, Ben Clarke, Derek B. Kosty, Scott K. Baker, Keith Smolkowski, Hank Fien

Publication History

Received 02 Dec 2014, Accepted 29 Jun 2015, Published online: 27 Dec 2019

Full Reference

Doabler, C. T., Clarke, B., Kosty, D. B., Baker, S. K., Smolkowski, K., & Fien, H. (2016). The effects of a core kindergarten mathematics program on the mathematics achievement of Spanish-speaking English learners. *School Psychology Review*, *45*, 343–361. doi: 10.17105/SPR45-3.343-361

Funding Source

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** awarded to the Center on Teaching and Learning at the University of Oregon.

Abstract

There is a dearth of research on Tier 1 instruction designed to improve the mathematics achievement of English learners. This study examined the impact of a core kindergarten mathematics curriculum on the mathematics achievement of Spanish-speaking English learners (SS-ELs). Secondary aims tested for differential response to the curriculum among SS-ELs as a function of (a) mathematics skills at the beginning of kindergarten, (b) the number of SS-ELs in classrooms, and (c) the frequency of mathematical discourse during core mathematics instruction. Data analyzed in the study were generated from a recent large-scale efficacy trial. Participants were 556 SS-ELs from 66 kindergarten classrooms. Results suggest SS-ELs in treatment classrooms made greater gains than SS-ELs in comparison classrooms on mathematics measures across the school year. Evidence of differential response to the curriculum among SS-ELs was not found. The importance of core mathematics instruction and implications for school psychologists are discussed.

The Effects of a Core Kindergarten Mathematics Program on the Mathematics Achievement of English Learners

While the societal importance of teaching for early mathematical proficiency has gained national attention (State of the Union Address, 2014), mounting evidence suggests that students from a variety of subgroups struggle to meet grade-level expectations in mathematics. Among these at-risk subgroups are English learners (ELs) or children of linguistic minority groups who lack full proficiency in English and receive language assistance. ELs represent a major presence in U.S. schools and for the last 20 years they have been the fastest growing subgroup (Francis et al., 2006; Klingner & Eppolito, 2014). Recent estimates suggest that ELs comprise 10% of the U.S. student population and that 70% of this subgroup is Spanish-speaking (Aud et al, 2013; Fry & Passel, 2009). Considering the rising presence of ELs in U.S. public schools (Aud et al., 2013) and the alarming number who have been disproportionally identified for special education (Sullivan, 2011), schools and teachers face the daunting challenge of meeting the instructional needs of ELs. Recent research shows, however, that schools are struggling to support ELs in developing mathematical proficiency.

Math achievement data from the 2013 National Assessment for Educational Progress (NAEP) indicate that 86% and 95% of fourth grade and eighth grade ELs, respectively, scored below Proficient (National Center for Education Statistics [NCES], 2013). There are also strong indications that ELs do no achieve commensurate with their English proficient peers. According to recent research, the math achievement gap between ELs and English proficient students appears early and remains relatively stable over the years (Reardon & Galindo, 2009). Since 1996, NAEP results have shown that an educationally meaningful achievement gap exists between ELs in 4th grade and their English proficient peers, and that this gap is nearly twice as large in 8th grade (NCES, 2013).

The convincing evidence that suggests ELs experience early and persistent math difficulties (MD) comes at a time when the Common Core State Standards for Mathematics (CCSS-M, 2010) have significantly raised the mathematical proficiency bar for U.S. students (Porter et al., 2011). The CCSS-M, relative to previous state standards, place greater demand on the development and use of academic language in mathematics (Dingman, Teuscher, Newton, & Kasmer, 2013). Students must now use precise mathematical language and vocabulary, verbalize and justify solution methods, and critique the reasoning of others (CCSS-M, 2010). While all students face the linguistic challenges associated with learning to use the language of mathematics in the context of the CCSS-M, these demands are compounded for ELs. They, unlike their native English-speaking peers, face the unfortunate "double demands" (Baker et al., 2014) of having to simultaneously acquire proficiency in two languages: English and mathematics (Cirillo, Bruno, & Eisenmann, 2010; Francis et al., 2006; Moschkovich, 1999).

Given the likelihood that many ELs will struggle to acquire math proficiency, a major focus of educational research and practice should be on improving the quality of core math instruction delivered in general education settings. For many students, core math instruction serves as the primary source of mathematics instruction. This is particularly true in the early elementary grades, when logistical constraints (e.g. half day programs in Kindergarten) and a primary focus on reading instruction may limit the availability of time and resources to support mathematics achievement beyond core instruction. Core math instruction, therefore, must be effectively designed and delivered to meet the instructional needs of all students, including ELs and other students at risk for MD. Evidence from recent randomized controlled trials has begun to document the utility of effective, core math instruction in promoting student math achievement, preventing MD, and reducing student need for highly intensive math interventions (Agodini & Harris, 2010; Author et al., 2008; Author et al., 2011; Fuchs, Fuchs, & Prentice, 2004).

Plausible Sources of Math Difficulties for ELs

While many factors (e.g., socio-cultural, linguistic, cognitive) may contribute to the difficulties that ELs experience in acquiring math proficiency, it is important to consider how instructional factors influence and, in some cases, initiate difficulties in mathematics. A lack of language-intensive instruction is one instructional factor that may explain why so many ELs are struggling with math and other academic areas (Arreaga-Meyer & Perdomo-Rivera, 1996; Cirillo et al., 2010; Khisty, 1995; Lee, Quinn, & Valdes, 2013; Moschkovich, 1999). Research in the areas of reading (Gersten et al., 2007; Gunn et al., 2005), social studies (Vaughn et al., 2009), and science (August et al., 2014; Lee et al., 2013) has found that the academic achievement of ELs is dependent upon meaningful opportunities to engage in the use of disciplinary language. In this study, we hypothesized that this principle would also maintain for ELs in the area of mathematics. We base this hypothesis on findings from a growing line of classroom observation research, which has shown that increased student math achievement is associated with student math verbalizations (Clements, Agodini, & Harris, 2013; Author et al., 2014). It can be argued then that classrooms, particularly those with higher percentages of ELs, should provide frequent opportunities for students to verbalize their mathematical understanding and thought processes.

An underdeveloped empirical research base in the area of effective math instruction for ELs can also be implicated as a contributing factor to the alarming number of ELs who struggle with math (Orosco, 2014). While significant efforts have been made in the practice of preventing reading difficulties for ELs (August et al., 2014; Baker et al., 2014; Gersten et al., 2007; Slavin

& Cheung, 2005; Vaughn, et al. 2006), few rigorously conducted studies have investigated the impact of interventions on the math achievement of ELs. For example, Janzen (2009) conducted a synthesis of the literature from 1990 to 2007 on teaching ELs in the content areas of English, math, science, and history. Within each content area, Janzen classified the findings into linguistic, cognitive, sociocultural and pedagogical sub-categories. Surprisingly, none of the 12 articles codified under the pedagogical category in math used a research methodology, such as a randomized controlled trial, quasi-experimental design, or single-case design, rigorous enough to identify and establish the causal agents of relevance to improved math achievement for ELs (Cook, 2002; Feuer, Towne, & Shavelson, 2002; Flay et al., 2005).

Since Janzen's (2009) review there has been a continued lack of rigorous experimental research on math instruction for ELs. In fact, Author and Author (in press) conducted a review of the literature base from 2000 to 2012 and found no experimental studies on math interventions with ELs. Our own review of the research from 2013 to the time of this study revealed just three math intervention studies involving ELs (i.e., Orosco, 2014; Orosco et al., 2011; Shumate, Campbell-Whatley, & Lo, 2012). *This paucity of research sheds light on the urgency to build the knowledge base on effective teaching practices and interventions designed to improve the mathematics achievement of ELs.*

Explicit Math Instruction and Its Role in Math Proficiency for ELs

One instructional approach that has strong potential for supporting ELs in developing mathematical proficiency is explicit math instruction. Over the last decade, research has begun to establish a solid evidentiary basis for using explicit math instruction to teach at-risk learners (Author et al., 2002; Bryant et al., 2011; Author et al., 2011; Dyson, Jordan, & Gluting, 2011; Gersten et al., 2009; Kroesbergen & Van Luit, 2003; National Mathematics Advisory Panel

[NMAP], 2008; Orosco, 2014; Orosco et al., 2011). For example, in a meta-analysis of 41 studies targeting students with MD, Gersten et al. (2009) found that explicit instruction had the largest impact, g = 1.22, 95% CI [0.78, 1.67], among seven dimensions of math instruction.

In this study, we hypothesized that a core math program characterized by an *explicit* and *systematic* approach to instruction would have positive effects on the math outcomes of ELs. We based this hypothesis on two factors. First, explicit math instruction shares similar characteristics to that of explicit reading instruction, an instructional approach repeatedly shown to improve reading outcomes among ELs (Baker et al., 2014; Gersten et al., 2007). Both instructional approaches, for example, expect teachers to overtly model and demonstrate what they want students to learn, and provide supportive, timely and specific academic feedback as students engage in guided and independent learning activities. Explicit reading and math instruction also incorporate visual models to teach key concepts and skills, and integrate systematic opportunities to support vocabulary development (Coyne et a., 2010; Author et al., 2012). Furthermore, both require teachers to engage students in meaningful discourse, such as productive math verbalizations, around key academic content (Gersten et al., 2007, 2009). Such academic discussions can help ELs build critical language skills in both English and mathematics (Baker et al., 2014; Cirillo et al., 2010).

A second reason why explicit math instruction may benefit ELs is because of findings from a growing line of research on core math programs (Agodini & Harris, 2010; Author et al., 2008; Author et al., 2011). Agodini and Harris (2010) investigated the effectiveness of four, commercially available first-grade elementary math programs in 39 schools. Their study found that the math achievement of students in schools that were randomly assigned to the two explicit instruction programs was significantly greater than that of students in schools that used the student-centered programs. Relatedly, findings from our own intervention research demonstrate that we can significantly increase the math achievement of struggling learners by designing and implementing math programs that incorporate validated explicit instructional design and delivery principles (Author et al., 2008; Author et al., 2011). One such program is the *Early Learning in Mathematics* (ELM) core math program.

Empirical Support for the Implementation of the ELM Math Program

While explicit math instruction is most well known for its role in small-group interventions, we have found encouraging results when using this instructional approach in core educational settings (i.e., general education classrooms). A primary focus of this line of efficacy research has been the ELM core math program. Our research team developed the ELM program to increase the mathematics achievement of kindergarten students. ELM is a yearlong, 120-lesson core program that targets kindergarten math topics identified in the CCSS-M (2010). At its core is an explicit and systematic architecture to early math instruction.

Recent efficacy research has documented preliminary empirical support for ELM's capacity to (a) improve student math achievement (Author et al., 2008; Author et al., 2011) and (b) support teachers in facilitating structured classroom discourse around critical math concepts and skills (Author et al., 2014). In a recent randomized controlled trial (Author et al., 2011), we tested the efficacy of ELM, randomly assigning 66 kindergarten classrooms to treatment and control conditions (business as usual). The student sample included approximately 1,300 kindergarten students, including at-risk and typically achieving children. Analyses revealed statistically significant effects for students in ELM classrooms over students in control classrooms on the Test of Early Mathematics Ability – 3rd Edition (TEMA-3; $t = 2.41$, $p = .02$, Hedges' $g = .15$) and Early Numeracy Curriculum Based Measurement (EN-CBM; $t = 1.99$ p

= .05, *g* = .13). We also found that at-risk students (i.e., students who scored below the 40th percentile on the TEMA-3 at pretest) significantly outperformed their at-risk control peers on both the TEMA-3 (t = 3.29, $p < .01$, g = .24) and EN-CBM total score (t = 2.54, p = .01, g = .22).

More recently, Author et al. (2014) investigated the efficacy of the ELM program in 129 kindergarten classrooms from 46 schools in Oregon and Texas. The study differed from previous investigations of ELM (Author et al., 2008; Author et al., 2011) in that it had a specific focus on teacher outcomes rather than student math achievement. Author and colleagues examined whether ELM increased teachers' facilitation of high-quality instructional interactions. Findings suggested that ELM stimulated more opportunities for teachers and students to engage in high-quality mathematical discussions compared to classrooms in the control condition. Specifically, higher rates of math verbalizations by groups of students (t = 5.09, $p < .001$, g = .91) and individuals (t = 3.30, p = .001, g = .57) were found in ELM classrooms.

Purpose of the Current Study and Research Questions

A clear and compelling need exists to build a scientific knowledge base of effective instructional practices and programs aimed at increasing the math achievement of ELs (McCardle, Mele-McCarthy, Cutting, Leos, & D'Emilio, 2005). To this end, the purpose of this study was to investigate the impact of the ELM core mathematics program on the mathematics achievement of ELs. To our knowledge, no similar studies have been conducted.

In addition to studying the efficacy of ELM, we also tested a set of *a priori* student and classroom-level predictors of differential response to the ELM program. Because level of mathematical knowledge at kindergarten entry is a proxy of risk status and has been found to be one of the strongest predictors of later math achievement (Duncan et al., 2007; Morgan, Farkas

& Wu, 2009; Judge & Watson, 2011), we examined whether the effects of ELM differed by ELs' initial skill performance in math. Additionally, we investigated whether the extent to which ELs are distributed across classrooms influenced the efficacy of ELM. Recent studies suggest that classrooms with higher percentages of disadvantaged students, including ELs, produce lower student math achievement (Isenberg et al., 2013). We also tested whether the use of mathematical language in classrooms was a predictor of differential response to ELM. As math standards increasingly emphasize the need for students to understand math concepts and demonstrate their mathematical understanding through verbal explanations of what math problems are asking and how they can be solved, this study offered an important opportunity to study the impact of an innovative approach in math with ELs. It may be, for example, that by facilitating multiple opportunities for students to verbally express their mathematical thinking and problem solving, ELs have more opportunities to understand math concepts and procedures, and improve their math skills. We believe this study was the first to investigate such a hypothesis, particularly in the context of testing the impact of a core math program with a strong emphasis on mathematics discourse.

In summary, this study was guided by the following four research questions:

1. What is the effect of the ELM program on the math achievement among ELs?
2. Do math skills at the beginning of kindergarten, as measured by the TEMA-3, predict differential response to the ELM program among ELs?
3. Does the number of ELs in classrooms predict differential response to the ELM program among ELs?
4. Does the frequency of math discourse used in classrooms predict differential response to the ELM program among ELs?

Method

Research Design and Database

This study conducted a secondary analysis of data collected during a large-scale efficacy trial funded by the Institute of Education Sciences (IES) and designed to investigate the efficacy of the ELM kindergarten intervention program (Author et al., 2011). The ELM efficacy trial was conducted in Oregon and Texas, respectively, during the 2008-2009 and 2009-2010 school years (Author et al., 2011; Author et al., 2014). Blocking on schools, 129 kindergarten classrooms were randomly assigned to either treatment (ELM; n = 68) or comparison (district-approved kindergarten mathematics instruction; n = 61) conditions. Thus, the ELM efficacy trial treated classrooms as the primary unit of analysis. In all, the original sample included 2,598 kindergarten students attending 129 classrooms in 46 schools. The ELM efficacy trial collected data from participating students to: (a) document demographic characteristics and (b) measure gains in student mathematics achievement from the beginning to the end of kindergarten. Classroom observations were also conducted in both conditions at fall, winter, and spring time points to measure the amount of math classroom discourse used during core math instruction.

Prior to analyzing the data, we established two inclusion criteria for what would constitute an eligible ELM efficacy trial classroom. A classroom was considered eligible if it: (a) enrolled students considered as ELs and (b) provided complete student demographic data related to students’ English language status. From the original sample of 129 kindergarten classrooms, 22 private school classrooms were dropped because they did not include ELs. We also excluded 14 public school classrooms and 3 charter public school classrooms because they provided incomplete EL status information. In total, our analytical sample included 90 kindergarten classrooms with 708 considered as ELs. Data analyzed in the current study included student mathematics achievement data collected from the 708 ELs and observational data documented in the 90 kindergarten classrooms.

Teacher and Student Sample

The 90 classrooms (48 treatment, 42 control) were from 31 schools located in 3 school districts in Oregon and 2 school districts in Dallas, Texas. Teachers in treatment classrooms delivered the ELM curriculum. In comparison classrooms, teachers provided district-approved kindergarten mathematics instruction. Of the 90 classrooms, 83 were located in public schools and 7 were in charter public schools. All charter school classrooms were located in Texas. Public school classrooms were located in schools eligible for Title 1 funding. Table 1 provides descriptive information about the classrooms and teachers by condition and region. Of the 90 classrooms, 75 provided a full-day kindergarten program and 15 provided a half-day program. All half-day classrooms were located in Oregon. Math instruction in all classrooms was delivered in English. Average class size for treatment and comparison classrooms was M = 21.8 (SD = 5.2) and M = 21.2 (SD = 4.1), respectively. The 90 participating classrooms were taught by 91 teachers. One comparison classroom in Oregon had two teachers, each working a half-day schedule. All teachers participated for the duration of the ELM efficacy trial.

Nested within the 90 classrooms were 708 EL kindergarten students. Of the 708 ELs, 407 and 301 were in treatment and comparison classrooms, respectively. As shown in Table 1, ELs in both conditions were similar in terms of age, ethnicity, gender, Spanish spoken as the primary language, and percentage of students identified for special education. Students were determined as being ELs based upon participating school district processes and policies, which varied across the study.

< Table 1 here >

ELM Intervention

ELM is a core kindergarten mathematics program that consists of 4 quarterly teacher manuals, each containing 30 daily lessons. Math content is systematically introduced, reviewed and extended through ELM's explicit instructional design framework. Each manual offers scripted guidelines to support teachers in demonstrating key math content, delivering timely academic feedback, and facilitating frequent practice opportunities for students, including structured verbal interactions between teachers and students, and among students, around key math content. Such practice opportunities are systematically designed to help students build mathematical proficiency, and develop mathematical language and vocabulary. To promote conceptual understanding, lessons incorporate frequent opportunities for students to work with visual representations of math ideas, such as 3-D shapes, counting blocks and numbers lines.

Mathematics domains targeted in ELM include: (a) counting and cardinality, (b) operations and algebraic thinking, (c) number and operations in base ten, (d) measurement and data, (e) geometry, and (f) precise mathematics vocabulary. Daily lessons last approximately 45 minutes in duration and include (a) whole-class and small-group activities focused on new mathematical content, (b) judicious review of previously learned material, and (c) worksheet activities that provide students extended practice with previously taught concepts and skills. Problem solving activities are introduced every five lessons to help students practice newly acquired problem solving skills and engage in "real world" mathematical problems, such as collecting categorical data and representing the data on a graph. Treatment teachers implemented the ELM program five days per week in whole-class settings.

Professional development. Treatment teachers received four professional development workshops related to program implementation. Each workshop lasted six hours and corresponded with the ELM quarterly teacher manuals. For example, the first workshop was conducted prior to the start of the school year and focused on lessons 1-30. Each workshop centered on research-based principles of math instruction and the instructional design and delivery features of the ELM program. Workshops also offered treatment teachers opportunities to practice with sample lessons and receive feedback from the ELM curriculum team.

Treatment Fidelity. Implementation fidelity of the ELM program was assessed three times in each treatment classroom. Teachers' adherence to the program was documented using a rating scale ranging from 0 (did not implement), 0.5 (partial implementation) to 1.0 (full implementation). Author et al. (2014) reported moderate levels of fidelity in the fall (M = .86, SD =.13), winter (M = .87, SD =.15), and spring (M = .87, SD =.14) and found no evidence of contamination between ELM and comparison classrooms.

Comparison Classrooms

Classrooms randomly assigned to the comparison condition provided standard district practices (business-as-usual). All comparison classroom teachers were asked to provide 45 minutes of daily math instruction. Instruction in the comparison classrooms entailed teacher-developed activities and a variety of commercially available math programs, including *Everyday Mathematics*, *Houghton Mifflin*, *Scott Foresman*, *Texas Mathematics,* and *Bridges in Mathematics.* Teachers in the comparison condition used a variety of instructional formats to deliver instruction, including whole-class instruction, center-based activities and peer-to-peer learning.

Measures

Student Measures

Students were assessed at pretest and posttest on measures of foundational aspects of number sense and whole number understanding. The assessment battery included a general outcome measure of students' procedural and conceptual knowledge of whole numbers, and a set of early mathematics curriculum-based measures that focused on discrete skills of number sense. Trained staff administered all student measures, with data collection meeting acceptable reliability criteria (i.e., implementation fidelity of .95 or higher).

Test of early mathematics ability-third edition. The Test of Early Mathematics Ability-Third Edition (TEMA-3; Pro-Ed, 2007) is a standardized, norm-referenced, individually administered measure of beginning mathematical ability. The TEMA-3 assesses mathematical understanding at the formal and informal levels for children ranging in age from 3 to 8 years 11 months. The TEMA-3 addresses children's conceptual and procedural understanding of mathematics, including counting and basic calculations. The TEMA-3 reports alternate-form and test-retest reliabilities of .97 and .82 to .93, respectively. For concurrent validity with other math outcome measures, the TEMA-3 manual reports coefficients ranging from .54 to .91. Standard scores were used in the analyses.

Early numeracy-curriculum-based measurement measures (EN-CBM). EN-CBM (Clarke & Shinn, 2004) consists of four, 1-minute fluency-based measures. The Oral Counting measure requires students to orally rote count as high as possible and the discontinue rule applies after the first counting error. The Number Identification measure requires students to orally identify numbers between 0 and 10. Quantity Discrimination requires students to name which of two visually presented numbers between 0 and 10 is greater. The Missing Number measure requires students to name the missing number from a string of three numbers (0-10), with the unknown number in the first, middle, or last position. Author et al. (2014) report concurrent validity coefficients between EN-CBM total scores and the TEMA-3 scores at pretest (r = .87) and posttest (r = .81). Average test-retest reliability of EN-CBM was reported as .89 (Author et al., 2014). A total EN-CBM score was computed as the sum across all subtests and used in subsequent analyses.

Observations of Core Mathematics Instruction.

To measure the frequency of math discourse used during core math instruction, project staff observed all 129 intervention and comparison classrooms. In Oregon, intervention and comparison classrooms were observed three times (fall, winter, and spring). Classrooms in Texas were observed two times (winter and spring). In the aggregate, 314 classroom observations were conducted in the 90 classrooms. Of the 314 observations, 74 served as interobserver reliability checks. Author et al. (2014) reported intraclass correlation coefficients (ICCs) that ranged from .67 to .95, suggesting substantial to nearly perfect interobserver reliability.

Trained observers documented the frequency of mathematical discourse in both conditions using the Classroom Observations of Student-Teacher Interactions–Mathematics (COSTI-M), a modified version of a classroom-level observation instrument designed by Smolkowksi and Gunn (2012). Data were collected on four student-teacher interaction behaviors associated with productive mathematical discourse: (a) teacher demonstrations, (b) teacher-provided academic feedback, (c) group responses, and (d) individual 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.

In the COSTI-M, teacher demonstrations reflect a teacher providing mathematical information in an overt and clear manner. Teacher demonstrations are considered a hallmark of explicit math instruction and include a teacher's explanations, verbalizations of thought processes, or physical demonstrations of mathematics content. Research from the learning sciences suggests that explicit demonstrations are a more efficient and effective way of presenting critical academic content to students compared to less-explicit teaching methods, such as discovery and problem-based learning, and inquiry-based teaching, (Kirschner, Sweller & Clark, 2006; Mayer, 2006). Academic feedback reflects a teacher's explanation of an incorrect student response or a verification of a correct student response. Evidence suggests that providing specific informational feedback about a student response or action improves learning and helps students understand how they performed during the process of learning (Halpern, Aronson, Reimer, Simpkins, Star, & Wentzel, 2007; Hattie & Timperley, 2007; Shute, 2008)

Group responses entail a concurrent mathematical verbalization from two or more students. When prompted and facilitated well, they present an opportunity to engage all students in a mathematical task, such as an entire class stating how the additive identity property applies when adding zero to another whole number. Individual responses reflect one student verbalizing or physically demonstrating the answer to a mathematical problem. When interspersed with group responses, individual responses allow teachers the ability to monitor the mathematical understanding of individual students. Empirical studies point to that fact that for young students to learn mathematics, they must be given frequent opportunities to engage in productive math discourse (Author et al., in press; Gersten et al., 2009), such as the group and individual responses captured by the COSTI-M. To avoid coding extraneous conversation, such as student "call-outs", group and individual responses were only coded if requested by the teacher.

Statistical Analysis

We assessed intervention effects on TEMA-3 standard scores and EN-CBM raw scores with a mixed-model (multilevel) time × condition analysis (Murray, 1998) to account for the intraclass correlation associated with students nested within classrooms, the level of random assignment. The analysis tested differences between conditions on change in outcomes from the fall (T1) to spring (T2) of kindergarten, with gains for individual students clustered within classrooms. The statistical model included time, condition, and the time × condition interaction, with time coded 0 at T1 and 1 at T2 and condition coded 0 for control and 1 for ELM. Analyses were based on 90 classrooms that included at least one EL and had complete student demographic data about EL status.

We also explored differential response to the ELM intervention as a function of various student- and classroom-level variables. We expanded the statistical model for this secondary aim to include a predictor and its interaction with condition, time, and the time × condition term; resulting in a three-way interaction, all corresponding two-way interactions, and individual (conditional) effects. The three-way interaction of the predictor, time, and condition provided an estimate of whether condition effects varied by the predictor.

Model estimation. We fit models to our data with SAS PROC MIXED version 9.2 (SAS Institute, 2009) using restricted maximum likelihood and included all available data, whether or not students' scores were present at both time points. Maximum likelihood estimation with all available data produces 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). Most missing data involved students who were absent on the day of assessment or transferred to a new school.

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. Regression methods have been found quite robust to violations of normality and outliers have a limited influence on the results in a variety of multilevel modeling scenarios (Bloom, Bos, & Lee, 1999; Donner & Klar, 1996; Fitzmaurice et al., 2004; Hannan & Murray, 1996; Murray et al., 2006). 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 of results, we computed an effect size, Hedges' g (Hedges, 1981), for each fixed effect. Hedges' g, recommended by the What Works Clearinghouse (WWC, 2011), represents an individual-level effect size comparable to Cohen's d (Cohen, 1988).

Results

Table 2 provides descriptive statistics for the primary outcome measures used to evaluate the impact of the ELM intervention among ELs. ELM and comparison classrooms did not significantly differ on any demographic characteristics or outcome measures collected at pretest.

< Table 2 here >

Efficacy

We tested the hypothesis that ELs in ELM classrooms experienced greater gains on TEMA-3 and EN-CBM during kindergarten than ELs in comparison classrooms. Complete results are summarized in Table 3, including the ICC for gains as described by Murray (1998, see p. 301). ELs in ELM classrooms statistically significantly outperformed ELs in comparison classrooms on the TEMA-3 (g = 0.24, p = .0395) and a trend-level effect was obtained for the EN-CBM (g =

0.20, p = .0553).

Differential Response

The tests of differential response included a predictor and its interactions with time, condition, and time × condition. The three-way predictor × time × condition interaction term indicates differential response to treatment. For each outcome measure, we tested for differential response to ELM as a function of (a) pretest student performance as measured by TEMA-3 and (b) the following classroom characteristics: number of ELs in the classroom; rate of group responses; rate of individual responses; rate of group and individual responses combined; and rate of teacher models, group responses, individual responses, and teacher-provided feedback combined. We found no statistically significant three-way interactions (p’s > .2610). Thus, our analyses were unable to offer clear evidence of differential response to the ELM intervention among ELs.

Discussion

This study examined the impact of a 120-lesson core kindergarten mathematics program on the mathematics achievement of ELs. For EL students in ELM classrooms there was a significant effect on one of the two outcome measures, TEMA 3 (g = 0.24, p = .0395). Impact on the second outcome measure was positive and trended towards statistical significance, EN-CBM (g = 0.20, p = .0553). Overall, results would be classified as substantively important positive effects (WWC, 2011). Analyses examining differential response found no difference in response to ELM for ELs by initial skill status, the number of ELs in the classroom, and a set of student-teacher interaction behaviors theorized to facilitate mathematical discourse used in core mathematics instruction.

We believe the finding that initial skill status, as measured by the TEMA-3 at the beginning of kindergarten, did not influence the impact of ELM for ELs is encouraging because it suggests that ELM essentially had the same positive impact for all ELs regardless of the amount of informal math knowledge they had prior to school entry. In other words, ELM seemed to work equally well across a range of skill levels. It may be that the instructional design features of the ELM program are configured in a manner that can help support the majority of EL kindergarten students in developing early mathematics proficiency. For example, ELs may gain a deep understanding of how numbers work through ELM's sequence of instruction, which strategically intersperses concrete instructional examples with abstract representations of numbers. ELs may also benefit from the way in which ELM uses simpler instructional examples rather than complex ones to introduce and teach new math concepts and vocabulary. It is plausible that these introductory instructional examples and problem contexts help engage the existing understandings and experiences of ELs and, in turn, allow them to achieve early success with new math content.

The finding of no differential response based on the initial skill levels of ELs is also somewhat surprising given that a previous study of ELM (Author et al., 2011) revealed that the program was more effective for students considered at-risk for math difficulties at the start of the kindergarten year (i.e., TEMA-3 pretest scores at or below the 40th percentile on the TEMA-3) than students considered on-track for developing mathematical proficiency (i.e., TEMA-3 pretest scores above the 40th percentile on the TEMA-3). While only 11% of ELs in the current sample tested above the 40th percentile on the TEMA-3 at pretest, we believe a strong case can be made from this study that typically-achieving ELs, like their EL peers who are at-risk for math difficulties at the start of kindergarten, may require explicit and systematic core math instruction given their limited proficiency in the English language. Converging evidence from investigations of the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) Cohort longitudinal dataset indicate that early math skills acquired in kindergarten are critical for acquiring proficiency in later mathematics and building knowledge in other content areas, including reading and science (Classens & Engel, 2013; Duncan et al., 2007; Morgan et al., 2009; Judge & Watson, 2011). Therefore, it is seems reasonable that all ELs may need an explicit and systematic core math program to make a successful start in kindergarten mathematics and begin to tackle the “double demands” of simultaneously learning the languages of English and mathematics.

Limitations and Future Research

A number of critical limitations should be considered when examining the findings from the study. First, the designation of a student as an EL was based on district methodology that varied widely across the different districts in the study. In part, this reflects the reality of actual practice (McCardle et al., 2005; Rueda & Windmueller, 2006). Thus caution should be exercised in extrapolating results of the current study to students and districts that may employ different classification methods. As with any study conducted within a unique geographic and demographic sample, a focus should be on replicating results across an array of diverse sites and participants (Cook, 2014; Flay et al., 2005).

Future research should explore the potential of differential impact based on initial skill level among ELs with a greater range of math understanding at the start of kindergarten. This is of particular importance as the field moves to implementing multi-tiered approaches to math instruction (Fuchs, Fuchs, & Compton, 2012). Greater insights into which students are likely to respond to Tier 1 core math instruction will be critical in building effective service delivery models. Establishing the impact of a core math program for ELs will also help determine if an EL student who fails to make sufficient growth is the result of lower levels of English language skills or due to difficulties specific to mathematics. That is, if the core program is effective for

ELs in general, an EL's non-response could more readily be attributed to a true deficit in mathematics. In a response to intervention (RtI) or multi-tier model, these EL students would be considered in need of a Tier 2 intervention. Because work investigating the impact of mathematics intervention programs is severely limited for ELs (Janzen, 2008), future research should not only focus on the effectiveness of core programs, but also be linked to ongoing efforts to develop and evaluate effective Tier 2 and 3 programs (Author et al., 2009; NMAP, 2008). As a result, this will better allow schools the opportunity to provide a full continuum of support for ELs as they learn mathematics.

We hypothesized that ELM's architecture of instruction would be a key ingredient in impacting student outcomes because of its incorporation of explicit instructional design principles, such as overt teacher demonstrations and deliberate student practice opportunities. To some extent, this general hypothesis was supported as ELM had a positive impact on the math achievement of ELs. However, additional analyses revealed the effect of the ELM program did not vary as a function of rates of math discourse during core math instruction (e.g., rate of group responses). That is, the impact of ELM was essentially the same between classrooms with high and low rates of math discourse opportunities. While ELM's scripted teacher manuals ensure frequent math verbalizations that target key math concepts and skills, this result suggests that the rate of these discourse opportunities may be less meaningful for ELs than the quality of the math verbalization opportunities. Research suggests that language-intensive math instruction is critical to students' development of math knowledge (Author et al., in press; Gersten et al., 2009; Kilpatrick, Swafford, & Findell, 2001). However, less is known about when math discourse opportunities should take place during the learning process and to what extent should they be cognitively demanding for students with limited proficiency in English. Toward this end, future research should focus on developing standardized observation protocols that document not only the extent to which math discourse opportunities occur during core math instruction but also the quality of these instructional interactions. This dual quantity-quality approach would likely provide a more comprehensive picture of whether all students, including ELs, are receiving effective, evidence-based math instruction and meaningful access to grade-level math content.

In this study, there is some concern about how directly we were able to evaluate math performance on the content taught as part of the ELM program. Given that ELM addresses multiple math domains (i.e. counting and cardinality, operations and algebraic thinking, number and operations in base ten, measurement and data, geometry, and precise mathematics vocabulary), our use of the TEMA-3 and EN-CBM as outcome measures may not have aligned fully with the content coverage of ELM. The TEMA-3 focuses primarily on whole number understanding and the EN-CBM measures focus on discrete aspects of number sense (e.g. magnitude comparison). These foundational skills are covered early in the scope and sequence of ELM's first quarterly teacher manual. Future research should utilize a more proximal assessment that directly links to the concepts taught across the math domains of ELM.

Another important issue to consider is the use of English-based math assessments to screen and identify ELs who may be in need of math intervention services. While the primary first language of this study's sample was Spanish, all math assessments were administered in English. Consequently, potential language barriers may have impacted ELs' pretest and posttest performances. Future research involving ELs should administer standardized, math assessments in both English and Spanish. This would allow for comparisons between students' math skills in English and Spanish.

Lastly, it should be noted that while we theorize that the instructional architecture of ELM was the primary agent impacting outcomes, treatment teachers were provided professional development on the ELM program and on effective teaching strategies and behaviors. While the professional development was primarily centered on ELM components and implementation, a stronger research design would have controlled for the impact of professional development by providing control teachers with general experiences on teaching strategies and behaviors. Doing so would have eliminated professional development as an alternate possible cause and confound when interpreting the study results.

Implications for School Psychologists

We believe core math programs can act as a valuable first line of defense in impacting the mathematics achievement of ELs. Core math programs serve as an instructional mechanism from which teachers can deliver important math content and support students' development of math proficiency (Author et al., 2012). When judiciously developed, core math programs align with theories of children's mathematical thinking and learning, and logically reflect the hierarchical structure of math (Clements, 2007). Despite this, a growing line of curriculum evaluation research indicates that many market leading core math programs lack the instructional design and delivery principles that are empirically validated to increase the math achievement of students with or at-risk for math difficulties (Bryant et al., 2008; Author et al., 2012; Sood & Jitendra, 2007).

School psychologists are skilled to assist schools and teachers in enhancing core math programs that fail to incorporate explicit instructional design and delivery principles. The purpose of such enhancements is to make core instruction more focused, explicit, and systematic for at-risk learners (Author et al., 2008; Author et al., 2012). For example, school psychologists will be able to ensure that core math programs organize instruction around the big ideas of mathematics and conspicuously teach those key concepts and skills through overt teacher demonstrations and visual representations of math ideas. They can also inspect whether core programs offer frequent and high-quality opportunities for students to practice with critical math content. We believe these enhancements can help ensure core math instruction is designed to meet the instructional needs of ELs.

Additionally, school psychologists will be able to assist teachers in making the math vocabulary presented in core math programs more precise and accessible for ELs. There is an increasing awareness of the need to embed academic vocabulary instruction within content area instruction and for that instruction to provide ample opportunities for ELs to engage with advanced content (Baker et al., 2014; Caldwell, Karp, Bay-Williams, Rathmell, Zbiek et al., 2011; Francis et al., 2006; Gersten et al., 2007). The ELM program employs an *extended* and *embedded* approach to vocabulary instruction (Coyne et al., 2009) to explicitly teach vocabulary central to success within mathematics (e.g., addition, inches) and academic words that are important across other content areas (e.g., explain, justify). We contend that this approach is reflective of best practices in the field because it prioritizes vocabulary and allows all students, including ELs, to actively use key vocabulary in a variety of math contexts.

Multi-tiered approaches to math instruction call for the delivery of high quality instruction to occur in Tier 1, core educational settings. Integral to any multi-tiered model is the collection of evidence for whether teachers deliver core math programs with fidelity and understanding. As suggested by O'Donnell (2008), documenting implementation fidelity is necessary to "determine whether unsuccessful outcomes are due to an ineffective program or due to failure to implement the program and its conceptual and methodological underpinnings as intended" (p. 42). School psychologists can support schools in correlating fidelity of implementation data with student math outcomes to obtain estimates for whether increased math achievement is due to higher levels of implementation. These correlations can also help determine if additional professional development is needed to improve the implementation of core math programs.

Finally, it is particularly imperative for schools to determine the most effective allocation of limited resources to meet the math needs of ELs. Early elementary classrooms often have a primary focus on early literacy, with supports including evidence-based core and intervention programs, reading coaches, and mandated reading instruction blocks (e.g. 90 minutes). Time constraints may be exacerbated for ELs who are in need of support to facilitate their development of English. School psychologists can assist schools in determining efficient approaches to ensure the math needs of EL learners are served.

Conclusion

Despite the preponderance of evidence that a successful start in mathematics is critical for all students (Classens & Engel, 2013; Morgan et al., 2009), and the fact that a concerning number of ELs are struggling to acquire proficiency in mathematics (NCES, 2013), there is an alarming shortage of empirical literature on effective mathematics instruction for ELs (Author, 2012). The current study addresses the urgent need for research in this critical area. While findings from this study are limited, they do indicate promise and should serve as a watershed for future research. Future studies will hopefully allow the field to begin building a research base on effective instructional practices for teaching mathematics to ELs.

References

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness*, *3*, 199–253. doi: 10.1080/19345741003770693

Arreaga-Mayer, C., & Perdomo-Rivera, C. (1996). Ecobehavioral analysis of instruction for at-risk language-minority students. *Elementary School Journal,* 96, 245–258.

Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., and Zhang, J. (2013). *The Condition of Education 2013* (NCES 2013-037). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch.

August, D., Branum-Martin, L., Cárdenas-Hagan, E., Francis, D. J., Powell, J., Moore, S., & Haynes, E. F. (2014). Helping ELLs Meet the Common Core State Standards for Literacy in Science: The Impact of an Instructional Intervention Focused on Academic Language. *Journal of Research on Educational Effectiveness*, *7*, 54-82.

Baker, S.K., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (NCEE 2014-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from: http://ies.ed.gov/ncee/wwc/publications_reviews.aspx.

Bloom, H. S., Bos, J. M., & Lee, S. W. (1999). Using cluster random assignment to measure program impacts. *Evaluation Review, 23*, 445-469. doi:10.1177/0193841X9902300405

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(1), 7-23.

Caldwell, J. H., Karp, K., & Bay-Williams, J. M. (2011). *Developing essential understanding of addition and subtraction for teaching mathematics in prekindergarten–grade 1* (Essential understanding series). Reston, VA: National Council of Teachers of Mathematics.

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.

Clements, D. H. (2007). Curriculum research: Toward a framework for "research-based curricula." Journal for Research in Mathematics Education, 38, 35–70.

Common Core State Standards Initiative. (2010). Common core state standards for mathematics. from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

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

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

Cook, T. D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, *24*(3), 175-199. doi: 10.3102/01623737024003175

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli Jr, R., Ruby, M., Crevecoeur, Y. C., & Kapp, S. (2010). Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness*, *3*(2), 93-120.

Dingman, S., Teuscher, D., Newton, J. A., & Kasmer, L. (2013). Common Mathematics Standards in the United States. *The Elementary School Journal*, *113*(4), 541-564.

Donner, A., & Klar, N. (1996). Statistical considerations in the design and analysis of community intervention trials. *Journal of Clinical Epidemiology, 49*, 435–439. doi:10.1016/0895-4356(95)00511-0

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(6), 1428–1446. doi: 10.1037/0012-1649.43.6.1428

Dyson, N. I., Jordan, N. C., & Glutting, J. (2011). A Number Sense Intervention for Low-Income Kindergartners at Risk for Mathematics Difficulties. *Journal of Learning Disabilities*. doi: 10.1177/0022219411410233

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher,* 31(8), 4–14. doi: 10.3102/0013189x031008004

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., ... & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention science*, *6*(3), 151-175.

Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Research-based recommendations for the use of accommodations in large-scale assessments. Book 3 in Practical Guidelines for the Education of English Language Learners* Retrieved from http://www.centeroninstruction.org/files/ELL3-Assessments.pdf

Fry, R., & Passel, J. S. (2009). *Latino children: A majority are US-born offspring of immigrants*. (No. 110). Washington, DC: Pew Hispanic Center. Retrieved from http://pewhispanic.org/files/reports/110.pdf

Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to Mathematical Problem-Solving Instruction Comparing Students at Risk of Mathematics Disability With and Without Risk of Reading Disability. *Journal of Learning Disabilities*, *37*(4), 293-306.

Gersten, R., Baker, S.K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective Literacy and English Language Instruction for English Learners in the Elementary Grades: A Practice Guide* (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the Development of Reading Skill Through Supplemental Instruction Results for Hispanic and Non-Hispanic Students. *The Journal of special education*, *39*(2), 66-85.

Halpern, D. F., Aronson, J., Reimer, N., Simpkins, S., Star, J. R., & Wentzel, K. (2007). *Encouraging girls in math and science. IES Practice Guide.* (No. NCER 2007-2003). Washington, DC: National Center for Education Research.

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.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77,* 81-112.

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

Isenberg, E., Max, J., Gleason, P., Potamides L., Santillano, R., Hock, H., & Hansen, M. (2013). "Access to Effective Teaching for Disadvantaged Students." Washington, DC: U.S. Department of Education, Institute of Education Sciences

Janzen, J. (2008). Teaching English language learners in the content areas. *review of Educational research*, *78*(4), 1010-1038.

Judge, S., & Watson, S. M. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. *The Journal of Educational Research*, *104*, 147–157.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: Mathematics Learning Study Committee, National Research Council.

Khisty, L. L. (1995). Making inequality: Issues in language and meanings in mathematics teaching with Hispanic students. In W. G. Secada, E. Fennema, & L. B. Adajian (Eds.), *New directions for equity in mathematics instruction* (pp. 279-297). Cambridge: Cambridge University Press

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

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist,* 41(2), 75–86. doi: 10.1207/s15326985ep4102_1

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33(1), 159–174.

Lee, O., Quinn, H., Valdés, G., (2013). Science and Language for English Language Learners in Relation to Next Generation Science Standards and with Implications for Common Core State Standards for English Language Arts and Mathematics *Educational Researcher, 42,* (223-233) doi:10.3102/0013189X13480524

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

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning?. *American Psychologist*, *59*(1), 14-19.

McCardle, P., Mele-McCarthy, J., & Leos, K. (2005). English language learners and learning disabilities: Research agenda and implications for practice. *Learning Disabilities Research & Practice*, *20*(1), 68-78.

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

Moschkovich, J. (1991). Association supporting the participation of English language learners in mathematical discussion. *For the Learning of Mathematics, 19*, 11–19.

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

National Center for Education Statistics (2013). The Nation's Report Card: A First Look: 2013 Mathematics and Reading (NCES 2014-451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

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

Orosco, M. J. (2014). Word Problem Strategy for Latino English Language Learners at Risk for Math Disabilities. *Learning Disability Quarterly*, *37*(1), 45-53. doi: 10.1177/0731948713504206

Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2013). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education*, *47*(2), 96-107.

Pro-Ed. (2007). *Test of early mathematics ability, third edition.* Austin, TX: Author.

Reardon, S. F., & Galindo, C. (2009). The Hispanic-white achievement gap in math and reading in the elementary grades. *American Educational Research Journal,* 46(3), 853–891. doi: 10.3102/0002831209333184

SAS Institute. (2009). SAS/STAT® 9.2 user's guide (2nd ed.). Cary, NC: SAS Institute, Inc. Retrieved March 10, 2010, from the SAS Product Documentation web site: 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

Shumate, L., Campbell-Whatley, G. D., & Lo, Y. Y. (2012). Infusing culturally responsive instruction to improve mathematics performance of Latino students with specific learning disabilities. *Exceptionality*, *20*(1), 39-57.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1),153-189.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research,* 75, 247–284. doi: 10.3102/00346543075002247

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(2), 316–328. doi: 10.1016/j.ecresq.2011.09.004

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

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., ... & Francis, D. J. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, *39*(1), 56-73.

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research on Educational Effectiveness*, *2*(4), 297-324.

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

Table 1

Descriptive Information for Students and Classrooms by Condition

	ELM	Comparison
Student characteristics		
Number of students *n*	407	301
Age *M* (*SD*)	5.6 (0.5)	5.6 (0.5)
Male *n* (%)	214 (53)	152 (51)
Hispanic *n* (%)	360 (89)	263 (87)
First language *n* (%)		
English	7 (2)	9 (3)
Spanish	328 (81)	228 (76)
Other	14 (3)	12 (4)
Eligible for special education *n* (%)	24 (6)	21 (7)
Classroom characteristics		
Number of classrooms *n*	48	42
Number of students per class *M* (*SD*)	21.8 (5.2)	21.2 (4.1)
Program structure *n* (%)		
Full-day program	41 (85.4)	34 (81.0)
Half-day program	7 (14.6)	8 (19.0)

Note. *M* = mean, *SD* = standard deviation. Age was computed as of the beginning of the study (i.e., 10/1/2008 for the Oregon cohort and 10/1/2009 for the Texas cohort)

Table 2

Descriptive Statistics for Outcome Measures by Assessment Point and Condition

Measure		Fall		Spring	
		ELM	Comparison	ELM	Comparison
TEMA-3	*M*	78.4	79.8	94.3	91.4
	(*SD*)	(13.5)	(13.4)	(13.3)	(13.0)
	N	297	202	356	270
EN-CBM	*M*	41.2	41.0	138.4	124.6
	(*SD*)	(33.3)	(35.1)	(53.4)	(57.3)
	N	344	242	355	271

Note. TEMA-3 = Test of Early Mathematics Ability-Third Edition; EN-CBM = early numeracy curriculum-based measure; *M* = Mean; *SD* = Standard deviation.