Hammill Institute on Disabilities

Journal of Learning Disabilities
2019, Vol. 52(3) 271–283

DOI: 10.1177/0022219418823708
journaloflearningdisabilities.sagepub.com
SAGE

Student Practice Opportunities in Core Mathematics Instruction: Exploring for a Goldilocks Effect for Kindergartners With Mathematics Difficulties

Christian T. Doabler, PhD,[1] Brian Gearin, MEd,[2] Scott K. Baker, PhD,[3] Mike Stoolmiller, PhD,[4] Patrick C. Kennedy, PhD,[2] Ben Clarke, PhD,[2] Nancy J. Nelson, PhD,[2] Hank Fien, PhD,[2] and Keith Smolkowski, PhD[5]

Abstract

Opportunities for practice play a critical role in learning complex behaviors. In the context of explicit mathematics instruction, practice facilitates systematic opportunities for students with mathematics difficulties (MD) to learn new mathematics content and apply such knowledge and skills to novel mathematics problems. This study explored whether there is an optimal amount of student practice that teachers should provide in core mathematics instruction to maximize the mathematics achievement of kindergarten students with MD, a so called "Goldilocks effect," as opposed to simply "more is better." Results from observation data collected in a large-scale efficacy trial supported the latter rather than the former. Specifically, we found that three individual practice opportunities for every explicit teacher demonstration of mathematical content was associated with increased mathematics achievement for students with MD relative to fewer practice opportunities. Implications for facilitating frequent student practice opportunities during core mathematics instruction and designing professional development for teachers who work with students with MD are discussed.

Keywords

explicit instruction, mathematics difficulties, practice opportunities, core mathematics instruction, Goldilocks effect, direct observation

Practice often plays a critical role in learning complex behaviors (Cepeda, Pashler, vul, Wixted, & Rohrer, 2006; Donovan & Bransford, 2005; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Fitts & Posner, 1967). Research from the field of neuroscience suggests that practice can lead to changes in the structure of the brain (Fields, 2005). Studies also indicate that practice improves targeted outcomes in a variety of performance-based disciplines. In music and sports, for example, findings strongly indicate that deliberate practice improves performance (Ericsson, Roring, & Nandagopal, 2007). The notion of deliberate practice in performance-based disciplines is similar to the idea of breaking academic tasks down into actionable steps. The specific steps are identified and practiced repeatedly, usually under the watchful eye of an expert coach. Steps that are more difficult for the individual, or critical in successful execution of the overall skill, are practiced more (Paumgarten, 2017).

In academics, practice often promotes understanding. Practice, when designed well, allows all students, including struggling learners, the opportunity to grasp new information, apply previously acquired knowledge and skills, and connect existing background knowledge with new content (Dunlosky et al., 2013). In the current study, practice is operationalized by teacher-initiated opportunities for individual students to independently and publicly demonstrate

[1]University of Texas at Austin, Austin, TX, USA
[2]Center on Teaching and Learning, University of Oregon, Eugene, OR, USA
[3]Southern Methodist University, Dallas, TX, USA
[4]Michigan State University, Marquette, MI, USA
[5]Oregon Research Institute, Eugene, OR, USA

Corresponding Author:
Christian T. Doabler, PhD, University of Texas at Austin, 1 University Station, D5300 SZB 408B, Austin, TX 78712, USA.
Email: cdoabler@austin.utexas.edu

their mathematical thinking and understanding through three different mediums: mathematics verbalizations, written responses, and manipulation of visual representations of mathematical ideas. Because explicit mathematics instruction has one of the strongest evidentiary bases for improving the mathematics outcomes of students with mathematics difficulties (MD; Gersten et al., 2009), our investigation of individual practice opportunities focused on those that occurred within a sequence of "explicit instructional interactions" between teachers and kindergarten students around critical mathematics content during core mathematics instruction. These interactions were deemed "explicit" because they entailed an initiating overt demonstration or explanation of a mathematical concept or skill by the teacher followed by opportunities for individual students to practice with the same mathematics content or slight variations thereof.

Notwithstanding the importance of practice, its effects are thought to be subject to the law of diminishing returns (Fitts & Posner, 1967; Ritter & Schooler, 2001). The concept of a "learning curve," for instance, describes a common dynamic whereby practice initially results in fast learning gains until a point is reached at which gains slow or level off completely. Further, in the classroom context, since the teacher has to distribute practice opportunities across many students and time is limited, the learning curve for some students could actually come back down again. This could happen for a variety of reasons such as the teacher presents fewer concepts to the group in order to fit in all the individual practice opportunities, the teacher does not distribute the practice opportunities across students optimally, or students not currently practicing get bored while other students practice and stop paying attention. As such, in the present study, we explored whether there is an optimal amount of student practice relative to teachers' overt demonstrations or explanations of mathematics content that maximizes the mathematics achievement of kindergarten students with MD (i.e., Goldilocks effect).

Studies of Explicit Mathematics Instruction in Tier 1 Kindergarten Classrooms

In kindergarten, Tier 1 mathematics instruction represents a critical tipping point in students' mathematical learning. When Tier 1 kindergarten mathematics instruction is poorly designed, the probability that students who enter school with little exposure to formal mathematics will experience persistent difficulties in later mathematics remains high (Barnes et al., 2016; Morgan, Farkas, & Wu, 2009). For many kindergarten students, Tier 1 mathematics instruction represents their first formal introduction to early mathematics. Consequently, the process of learning mathematics can be a novel and challenging experience for them. These students, who are at risk for persistent MD, may require more systematically designed opportunities to learn during Tier 1 mathematics instruction to acquire a deep understanding of mathematics. To promote outcomes among students who demonstrate academic risk, experts on effective instruction strongly suggest that Tier 1 instruction incorporate more frequent, explicit instruction (Deshler, 2015; Simmons, 2015; Vaughn, 2015). Others have similarly argued if students do not receive explicit instruction of adequate quality at Tier 1, Tier 2 instruction will not be sufficient to accelerate their learning such that the learning gap is narrowed (Baker, Fien, & Baker, 2010; Fuchs & Vaughn, 2012).

Relative to other instructional approaches, explicit mathematics instruction has garnered significant empirical support for promoting mathematics outcomes among students with or at risk for MD (Morgan, Farkas, & Maczuga, 2015). This is one reason that explicit instruction often features prominently in Multi-Tiered Systems of Support (MTSS) models, which are encouraged by state and federal law. And while the vast majority of studies involving explicit mathematics instruction conducted to date have taken place in Tier 2 settings (Gersten et al., 2009), it is encouraging that recent randomized controlled trials have begun to demonstrate the efficacy of this instructional approach in the context of Tier 1 mathematics instruction in kindergarten classrooms. Sood and Jitendra (2013), for instance, compared the impact of an explicitly designed, Tier 1 kindergarten mathematics program aimed at early number sense concepts relative to standard mathematics instructional practices. Significant differences in student mathematics achievement were reported in favor of the number sense intervention program, with effect sizes (Hedges' g) ranging from 0.55 to 1.44.

More recently, Clarke et al. (2015) conducted a large-scale, randomized controlled trial to investigate the efficacy of the Early Learning in Mathematics (ELM) program in 129 kindergarten classrooms. Whereas classrooms randomly assigned to the treatment condition used the ELM program, a 120-lesson core mathematics program that centers on an explicit instructional framework, control classrooms continued to provide "business-as-usual" (BAU) mathematics instruction. Of the 2,600 kindergarten students who participated in the study, approximately 50% were considered at risk for MD at the start of the kindergarten year. A major finding in this efficacy work was that the ELM program (i.e., treatment classrooms) produced a pattern of change relative to BAU such that students at the low end of the pretest distribution on a standardized measure of mathematics achievement scored substantially higher at posttest compared to their BAU peers. That is, initially low-performing students in ELM classrooms benefited the most, suggesting a pattern referred to as "differential effectiveness." It is worth noting that the observed pattern implied a flatter (i.e., closer to zero) slope of the regression of posttest on pretest in ELM classrooms.

Student Practice Opportunities Situated in Explicit Instructional Interactions

While a host of variables likely mediated the positive effects of these Tier 1 explicit mathematics programs in kindergarten (Clarke et al., 2015; Sood & Jitendra, 2013), one plausible specific mediator could be the increase over BAU of frequent, meaningful student practice opportunities. Mathematics research suggests that practice with foundational concepts and skills is critical for supporting students' development of early mathematical proficiency (Clements, Agodini, & Harris, 2013). In the MD research literature, a consistent finding is the beneficial effect of explicitly designed and delivered practice opportunities on the mathematics achievement of students with MD (Doabler et al., 2015; Gersten et al., 2009; Morgan et al., 2015).

In many ways, the types of student practice opportunities investigated in the current study are not dissimilar to "opportunities to respond" (OTRs; Greenwood, Delquadri, & Hall, 1984; MacSuga-Gage & Gage, 2015; Sutherland, Alder, & Gunter, 2003) in that both include verbal and physical student responses. Moreover, like some previously investigated OTRs, the types of student practice opportunities examined here are embedded within a sequence of explicit instructional interactions, where student practice opportunities are proceeded by an overt teacher demonstration or explanations of a mathematical concept or skills. Such demonstrations are intended to scaffold students' learning so that they are better prepared to independently engage in and work with the targeted mathematical content. For example, a teacher might directly model for students how to identify which of two groups of cubes has more than the other and then facilitate a series of practice opportunities for individual students to independently practice the same skill with groups of objects that have different quantities than the original practice opportunity.

The ELM program (Clarke et al., 2015) emphasizes three types of student practice opportunities situated within sequences of explicit instructional interactions. These practice opportunities, which are operationalized in accordance to the burgeoning empirical literature on explicit mathematics instruction (Agodini & Harris, 2010; Clarke et al., 2015; Gersten et al., 2009; Sood & Jitendra, 2013), center at the individual student level and include student mathematics verbalizations, opportunities to work with concrete manipulatives, and written response opportunities. Each type of individual practice opportunity serves as not only an effective mechanism for improving student mathematics outcomes and supporting active processing of mathematical content and learning (Clements et al., 2013; Doabler et al., 2015; Gersten et al., 2009) but also as efficient ways for teachers to monitor student progress, identify potential misconceptions, foster student engagement, and differentiate instruction for students with MD. They are also meaningful constructs to teachers and curriculum developers.

The first type of individual practice is student mathematics verbalizations. In the early grades (i.e., kindergarten and first grade), before students are proficient in reading and writing, much of the student practice facilitated during mathematics instruction is mediated through language. That is, teachers verbally demonstrate or explain an idea, and students demonstrate their understanding of the idea and verbally practice applying it through a range of examples. For teachers, this verbally mediated instructional pattern has the added advantage of being overt and public—it is more time efficient to gauge how students are doing than when they are silently reading mathematics problems and then solving them through the medium of writing. For younger students, verbal practice allows them to convey their mathematical understanding and thought processes through mathematical discourse or "math talk." For example, a teacher might have a kindergarten student verbalize how she solved a "put together" word problem, asking the student to explain how she combined two groups of objects to form a larger group. Research highlights the important relationship between mathematics verbalizations and student mathematics achievement (Clements et al., 2013; Doabler et al., 2015; Gersten et al., 2009).

Another valuable type of individual practice consists of opportunities for students to work with concrete representations of mathematical ideas, such as place value blocks and 3-dimensional geometric shapes. When explicitly structured, practice with concrete materials allows students to make connections between mathematics concepts and the abstract symbols depicting those concepts (Gersten et al., 2009). The third type of practice entails solving mathematics problems with written symbols, such as solving basic number combinations. When purposefully designed (i.e., distributed across time and interleaved with different kinds of problems), such practice can help students build automaticity with mathematical procedures and operations (Fuchs et al., 2010).

A common instructional technique used during the explicit instructional interactions targeted for the current investigation is for the teacher to have one student respond to the same problem, or a very similar problem, to the one the teacher just demonstrated and then to ask other students to individually respond to more complex applications of the concept to demonstrate their understanding. In this way, the teacher can make sure that students with MD receive targeted, individual opportunities to practice applying the taught concept. For example, a teacher might have a student with MD verbalize how to use base-ten blocks to compose a teen number (e.g., 15). Then, several students, including the one involved in the initial practice opportunity, would individually be asked to demonstrate their understanding of the concept by composing other two-digit numbers from a different decade range (e.g., 20–29). These additional practice opportunities may promote a more robust and lasting understanding of mathematics among all students in general and students with MD in particular.

Research on Opportunities to Practice During Mathematics Instruction

How much practice students require to gain mathematical proficiency remains the subject of ongoing inquiry. Three single-case design studies showed that mathematics interventions with purposeful increases of practice opportunities improved multiplication performance among students with MD (Skinner, Belfiore, Mace, Williams-Wilson, & Johns, 1997; Skinner, Ford, & Yunker, 1991; Sutherland et al., 2003). However, while of importance, these studies do not shed any light on the amount of practice opportunities students with MD should receive in Tier 1 instruction.

Other studies provide initial evidence for the importance of practice in core mathematics instruction and suggest that amount of practice matters. Clements et al. (2013) used observation data to examine the relationships between "teacher-directed" instructional practices and the mathematics outcomes of first and second grade students. The observation data were collected in over 600 first and second grade classrooms during a large-scale, randomized controlled trial focused on the efficacy of four different mathematics curricula. A low-inference observation measure was employed to document the frequency of nearly 100 items associated with the instructional practices of the four curricula. Analyses offered mixed results. While not statistically significant in first grade classrooms, results suggested that the frequency of individual student mathematics verbalizations were related to increased mathematics achievement in second grade.

Similarly, Doabler et al. (2015) explored the associations between the rate (per minute) of explicit instructional practices, including student practice opportunities, and gains in student mathematics outcomes using data collected during a randomized controlled trial focused on testing the efficacy of the ELM mathematics program (Clarke et al., 2015). Doabler et al. (2015) used a low-inference observation measure to document the frequency of individual and group practice opportunities. Similar to the current study, the student practice examined was teacher-initiated and included opportunities for students to independently demonstrate their mathematical thinking and understanding through mathematics verbalizations, written responses, and use of visual representations of mathematical ideas. However, unlike the current study, such practice was not investigated within sequences of explicit instructional interactions. Analyses of approximately 400 observations conducted in 129 kindergarten classrooms indicated that students in classrooms with more frequent individual practice opportunities made substantively important gains in mathematics outcomes from the beginning to the end of kindergarten. Specifically, students increased overall performance on a standardized mathematics outcome measure focused on foundational, whole number concepts and problem-solving skills. Group practice (i.e., choral response opportunities), however, was not found to be a statistically significant predictor of students' mathematics achievement (Doabler et al., 2015).

While findings from this prior observational research suggest that frequent practice is important in students' mathematical learning (Clements et al., 2013; Doabler et al., 2015; Skinner, Belfiore et al., 1997; Skinner et al., 1991; Sutherland et al., 2003), these studies do not identify an optimal amount of student practice relative to teachers' explicit demonstrations and explanations of mathematical content needed to maximize the mathematics achievement of students with MD (i.e., Goldilocks effect). Because time is a precious resource in schools, particularly when working with struggling learners (Kame'enui, 1993), identifying just the "right amount" of practice could help teachers use their mathematics instructional time more efficiently and effectively. For instance, smaller quantities of practice, such as for every teacher demonstration there would be on average of one practice opportunity (i.e., a 1:1 ratio), may enable teachers to maintain a faster pace and more quickly teach new concepts. However, additional practice, such as a 1:3 ratio of teacher demonstrations to student practice, may allow teachers to promote a deeper understanding of a targeted concept or skill among struggling learners. Moreover, more frequent practice may increase student mastery of material taught and give students a better opportunity to apply their newly acquired knowledge and skills in novel problem-solving contexts. On the other hand, more time spent on individual practice reduces the time for teachers to introduce new concepts and may slow the pace to the point where some students get bored and stop paying attention.

In sum, the question of whether individual, student practice opportunities relative to an explicit teacher demonstration of mathematical content are subject to the law of diminishing returns and have an optimal level is of great practical importance (Ritter & Schooler, 2001). For instance, identifying specific amounts of student practice may help further refine how teachers can optimally deliver during Tier 1 instruction to accelerate the mathematical learning of students with MD. Additionally, information gained from such investigations could help curriculum developers design more effective mathematics programs for teaching mathematics to the full range of students.

Purpose of the Study

The purpose of the present study was to explore whether there was a Goldilocks effect for the ratio of individual student practice opportunities to teacher demonstrations during Tier 1 mathematics instruction. Little research has been conducted on optimal ratios of student practice for students with MD. As such, we extend previous research on the frequency of observed student practice opportunities during mathematics instruction (Clements et al., 2013; Doabler et al., 2015; Skinner, Belfiore et al., 1997; Skinner et al., 1991; Sutherland et al., 2003) by addressing a critical question: Is there an optimal ratio of student practice relative to every overt teacher demonstration or explanation of mathematics content that maximizes the mathematics achievement of kindergarten students with MD (i.e., a Goldilocks effect).

Method

Data Source

This study is a secondary analysis of data collected during a randomized controlled trial funded by the Institute of Education Sciences and designed to test the efficacy of the ELM kindergarten mathematics program (Clarke et al., 2015). The ELM Efficacy Trial took place in Oregon and Texas during the 2008–2009 and 2009–2010 school years, respectively. A total of 129 kindergarten classrooms participated in the efficacy trial, of which 68 were randomly assigned to use the ELM core mathematics program (treatment) and 61 were randomly assigned to continue implementing BAU mathematics instruction (control). Data analyzed in the current study include student mathematics outcomes collected at pretest and posttest and observed rates (per minute) of individual student practice opportunities and teacher demonstrations and explanations captured in sequences of explicit instructional interactions during Tier 1 kindergarten mathematics instruction.

Participants

Kindergarten classrooms. Participants were recruited from 129 classrooms across 46 schools (32 public, 11 private, and 3 charter) from Oregon and Texas. All private and charter schools were in three school districts in Texas. Of the 129 classrooms (64 Oregon, 65 Texas), 112 provided full-day kindergarten, and 17 provided half-day kindergarten. All half-day kindergarten classrooms were in Oregon. One full-day classroom in Oregon operated 4 days per week. The 129 classrooms included 16 bilingual education classes, but all mathematics instruction was conducted in English. Average class size was 21 students (SD = 3.8).

Teachers. The 129 classrooms were taught by 130 teachers (98% female; 69% White, 20% Hispanic, and 11% another ethnic group). Two half-time teachers taught one classroom. In terms of background and experience, 129 teachers held certification, 39% held a graduate degree, and 51% had completed college-level coursework in algebra. Nearly all teachers had 7 or more years of total teaching experience.

Students. A total of 2,708 kindergarten students (47.3% female) participated in the study. Approximately 50% began kindergarten below the 25th percentile on the Test of Mathematics Ability–3rd Edition (TEMA-3) and thus were considered at risk for MD. The treatment condition included 1,475 students; the control condition included 1,233 students. Student demographic data were only available for those students who attended one of the 32 participating public schools. In those public schools, students were 56.7% White, 16.3% African American, 15.3% American Indian, 8.3% Asian, and <1% Pacific Islander. An average of 76% of the student population was eligible for free or reduced-price lunch programs. Approximately 29% qualified for Limited English Proficiency (LEP) services and 5% received special education.

ELM

ELM is a core (Tier 1) kindergarten mathematics program that promotes the development of mathematical proficiency in five domains of kindergarten mathematics: (a) counting and cardinality, (b) operations and algebraic thinking, (c) number and operations in base 10, (d) measurement and data, and (e) geometry. Precise mathematics vocabulary is also a cornerstone of the ELM program and thus is prioritized throughout its lessons. Classroom teachers deliver the program's 120 lessons in whole class settings. Each lesson lasts for approximately 45 minutes and provides pedagogical support for teachers to provide explicit models and explanations of new mathematical content and facilitate frequent guided and independent practice opportunities for individual students and the group at large. For example, an ELM teacher might have an individual student verbalize the steps involved for sorting geometric shapes by their different attributes. Across the school year, teachers in the treatment condition received four 6-hour professional development sessions on (a) evidence-based principles of mathematics instruction, (b) the instructional design principles of ELM, and (c) ELM's mathematical content.

Standard District Mathematics Instruction

Mathematics instruction provided in the 61 control classrooms consisted of BAU mathematics instruction, as represented by various published curricula and teacher-developed materials (e.g., Texas Mathematics curriculum, Everyday Mathematics). Observations revealed that instruction in the 61 control classrooms primarily focused on whole number concepts, followed by concepts of geometry and measurement. This instruction was delivered through a variety of instructional formats, including small groups and whole-class activities.

Measures

All participating students were administered two mathematics outcome measures at the start (fall) and end (spring) of their kindergarten school year.

TEMA-3. Mathematics achievement was measured with the TEMA-3 (Ginsburg & Baroody, 2003), a 72-item norm-referenced measure of early number sense. For student-level reliability, the publisher-reported estimates of internal consistency exceed .92, and alternate-form and test-retest reliabilities exceeded .80. Concurrent validity coefficients with four commonly used tests of mathematics ranged from .55 to .91. For classroom-level reliability, the intraclass correlation coefficient (ICC) for classrooms for the pretest TEMA-3 was .26, and the average classroom reliability of pretest TEMA-3 across all 129 classrooms was .85. Given the high level of reliability at the classroom level, we used the observed classroom average of the pretest TEMA-3 to represent classroom-level effects of the pretest TEMA-3 on outcomes in all the multilevel SEMs.

Early numeracy curriculum-based measurement (EN-CBM). EN-CBM (Clarke & Shinn, 2004) is a set of four fluency-based measures of early number sense: oral counting, number identification, quantity discrimination, and strategic counting with strings of numbers. Prior research reported a predictive validity coefficient of $r = .81$ between an EN-CBM total score and the TEMA-3 (Clarke et al., 2015). In this study, the total score on the EN-CBM in the fall of kindergarten, as computed as the sum across the four measures, served as a predictor of TEMA-3 to reduce potential bias from missing TEMA-3 data at both pre- and posttest.

Classroom Observations of Student-Teacher Interactions–Mathematics (COSTI-M). Trained research staff used the COSTI-M (Doabler et al., 2015; Smolkowski & Gunn, 2012) to document the frequency of explicit mathematics instructional practices. The COSTI-M is a low-inference observation measure that has been empirically validated across four federally funded efficacy trials (Clarke et al., 2015; Clarke et al., 2016; Fien et al., 2015; Smolkowski & Gunn, 2012). Doabler et al. (2015) reported predictive validity of the COSTI-M with the TEMA-3 ($p = .004$, Pseudo-$R^2 = .08$) and the EN-CBM ($p = .017$, Pseudo-$R^2 = .05$). The COSTI-M measures the number and rate of teacher demonstrations, individual student practice opportunities, group practice opportunities, and teacher-provided academic feedback. The latter two were not included in the current study because they have not been found to be statistically significant predictors of students' mathematics achievement (Doabler et al., 2015). Instead, the current study focused on (a) individual practice opportunities, given their predictive utility of student mathematics outcomes reported in prior research (Clements et al., 2013; Doabler et al., 2015; Gersten et al., 2009), and (b) teacher demonstrations, given that some nontrivial amount of teacher-led instruction, such as a teacher demonstrating how to decompose a teen number into a 10 and some 1s, is clearly necessary for initiating and scaffolding individual response opportunities among students with MD, especially those in kindergarten.

As operationalized in the COSTI-M, teacher demonstrations represent explicit explanations and demonstrations of mathematics content. For example, observers coded a teacher demonstration when a teacher used a think-aloud technique to overtly describe the attributes of three-dimensional shapes. Individual student practice opportunities consist of a single student verbalizing or physically demonstrating his or her mathematical understanding with and without support from the teacher. For example, observers would code two separate individual response opportunities if a teacher had the same student identify a 2-dimensional shape and then verbally state the attributes of the shape. In this study, we examined individual practice opportunities and teacher demonstrations that occurred within sequences of explicit instructional interactions. As previously noted, these sequences consisted of an initial teacher's demonstration or explanation of mathematical content followed by one opportunity for an individual student to practice or a series of separate individual practice opportunities.

Trained observers administered the COSTI-M in all 129 participating kindergarten classrooms. Observations were scheduled in advance and occurred in the fall, winter, and spring, with approximately 6 weeks separating each observation round. One observation was planned per classroom for each observation round. Observers remained in each classroom for the duration of mathematics instruction, with an average observation lasting 46 minutes ($SD = 19$ minutes).

Observers received approximately 14 hours of training across three sessions. Training focused on direct observation procedures, kindergarten mathematics instruction, and procedures associated with the use of the COSTI-M. Interobserver agreement, which was represented by ICCs, indicated that observers reliably used the COSTI-M. The ICCs ranged from .61 to .99, which based on guidelines proposed by Landis and Koch (1977) represented substantial to nearly perfect interobserver reliability. In the analyses reported here, stability ICCs were .40 and .26 for individual practice opportunities and teacher demonstrations, respectively, which given two to three occasions of observation implies quite modest reliabilities (.62 and .45, respectively) for a construct score based on the observed rates. As such, we used a latent variable approach to eliminate bias due to low reliability and obtain a more accurate estimate of the effects of the rates of teacher demonstrations and individual student practice opportunities on student outcomes.

Missing data covariates. To minimize potential bias from missing outcome data, two demographic variables were included as missing data covariates (i.e., auxiliary variables): student LEP status and school State-Type status (two dummy indicators, Texas-public and Texas-private or -charter vs. Oregon-public as the omitted reference category). Details about the role of the missing data covariates in the model are given below.

Statistical Analysis

These analyses extend the previously reported differential effectiveness of ELM (Clarke et al., 2015) by testing the extent to which specific ratios of individual student practice opportunities to teacher demonstrations predicted classroom-level differences in the slope of the posttest on pretest TEMA-3 regression, regardless of condition (i.e., treatment or control) in the larger ELM Efficacy Trial (Clarke et al., 2015). Specifically, we model quadratic and linear latent representations of specific ratios of individual student practice opportunities to explore whether there is an optimal level of individual practice opportunities relative to a teacher's explicit demonstration or explanation of mathematical content and the mathematics achievement of students with MD. These analytic models do not include the ELM (treatment) vs. standard practice (control) distinction because our goal was not to test for mediation but to test whether specific practice-to-demonstration ratios predict a flatter slope, and thus greater differential effectiveness, regardless of the condition to which a teacher happened to be assigned as part of the original study. Our modeling process consisted of two complimentary steps. First, we constructed and tested a latent measurement model to characterize differences among teachers in ratios of individual student practice opportunities to teacher demonstrations. Second, we evaluated the extent to which the latent ratio predicted the post on pre slope of classroom-level TEMA-3.

Latent variable models of teacher demonstrations and individual practice opportunities. As is typical for rate variables, the distributions were positively skewed. Thus, we log-transformed the rates (adding a small positive constant as a continuity correction) to better approximate the multinormality assumptions of latent variable models. Log transformation also offers a second important advantage: It transforms a ratio of rates (i.e., the number of individual practice opportunities per teacher demonstration) into a difference score of log rates, which are easy to create and work with in standard SEM software. Because specification of the measurement models for teacher demonstrations and individual practice opportunities were identical (see Doabler et al., 2018 for details), the logged rates from the repeated classroom observations were specified as indicators of a single latent variable for each instructional component. All factor loadings were constrained to 1 and the indicator intercepts were constrained to 0, making the measurement model a random intercept model in which the latent factor captures the differences between teachers in teacher demonstrations and student practice opportunities that were stable across the school year.

The separate measurement models for teacher demonstrations and individual practice opportunities were then combined into a single model, creating (a) a latent difference score representing latent logged individual practice opportunities minus latent logged teacher demonstrations (Raykov, 1992) and (b) a latent intercept score representing latent logged teacher demonstrations. For simplicity of exposition, we shall refer to the difference score as the ratio.

Ratios of individual practice opportunities. The last step in the modeling process was to combine the two-level student achievement model with the latent classroom-level model. In this combined model, latent variables representing rates of teacher demonstrations and the latent ratio for student practice to teacher demonstrations, along with classroom-level pretest TEMA-3, were used to predict classroom-level posttest TEMA-3 random slope. To test for the possibility of an optimal level of the ratio, we used Mplus 7.31 (Muthén & Muthén, 2015) to create latent quadratic versions of the latent ratio (Klein & Stoolmiller, 2003). The linear and quadratic effects were both used as predictors of the random slope. For higher ratios of student practice to teacher demonstrations to be predictive of student differential effectiveness, the latent ratio would need to have a negative effect on the random slope (i.e., make it flatter). If there was also an optimal level of the ratio, the quadratic trend of the latent ratio would be positive and, in combination with the linear effect, the overall fitted relation would resemble a U, J, or backwards J, with the very bottom of the U, J, or backwards J representing the optimal ratio of student practice opportunities.

Missing data. Because rates of missing student pretest data on the TEMA-3 were higher for initially low-skilled students (as measured by pretest EN-CBM), we attempted to account for the missing pretest data to make the missing-at-random assumption more plausible (Graham, 2009), minimize potential bias, and maximize power. We included two auxiliary classroom-level variables with no missingness as predictors of outcomes because they were related to rates of missingness, outcomes, or both: study site (Oregon vs. Texas) and type of school (traditional public vs. private or charter public). We also included two auxiliary student-level variables, LEP status and pretest EN-CBM, as predictors of outcomes, both of which had lower rates of missingness than the TEMA-3 and were correlated with missingness, the TEMA-3, or both. We also included all student-level pretest data (i.e., TEMA-3 and EN-CBM) in the missingness portion of the model, which necessitated using a more complicated approach to estimation (i.e., Monte Carlo numerical integration; MCNI) but allowed us to retain 98% of the student sample.

All SEMs were modeled in Mplus, using robust full-information maximum likelihood estimation. As in prior work, we first centered pretest TEMA-3 standard scores at the 98th percentile value of the pretest TEMA-3 distribution (i.e., 127) and rescaled both pretest and posttest TEMA-3 standard scores to prevent convergence issues by dividing the publisher-derived scaled scores (M = 100, SD = 15) by 20 (Doabler et al., 2018). All p values are two-tailed. Model fit was evaluated using a combination of the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the chi-square p value. Prior to developing our statistical models, we carefully examined univariate distributions of the instructional rate (i.e., teacher demonstrations and individual practice) and student outcome variables, checking for outliers and nonnormal distributions. We also inspected bivariate scatter plots at both the teacher and student levels to check for substantial departures from linearity and outliers.

Results

Missing Data and Descriptive Statistics

Our overall student sample size was 2,708 students nested within 129 classrooms. For the two-level models with both student and teacher data, we excluded 30 students who were missing fall LEP status and 28 students who were missing all pretest (i.e., TEMA-3 and EN-CBM) data, resulting in a total analytic sample of 2,650 students (98% of students) nested in 129 classrooms. For the one-level models with only teacher data, we dropped one additional classroom that was missing data on both teacher demonstrations and individual student practice opportunities, resulting in an analytic sample of 128 classrooms (99%). Descriptive statistics for student- and teacher-level variables are reported in Table 1. For example, across the year, participating classrooms averaged between .51 and .60 individual practice opportunities per minute and between .55 and .67 teacher demonstrations per minute, with average ratios of individual practice to teacher demonstration between 1.12 and 1.24. Note that COSTI-M data were not collected in Texas schools during Observation 1, resulting in a smaller sample at that point.

Latent Variable Models of Teacher Demonstrations and Individual Practice Opportunities

Results of analyses evaluating the latent rate models of explicit mathematics instruction in isolation are presented in detail in Doabler et al. (2018). The latent rate model that included a ratio of teacher demonstrations to individual student practice opportunities fit the data well (chi-square = 32.98, df = 32, p = .4190, RMSEA = 0.003, CFI = 0.987, TLI = 0.986). Because the model had a fairly large number of parameters given the moderate teacher sample size, we checked the robustness of the results by estimating a Bayesian version of the same model using noninformative priors. Posterior medians and 95% credibility intervals were very similar to their multiple linear regression counterparts. The latent ratio was correlated –0.42 with latent logged teacher demonstrations, indicating that classrooms with lower ratios tended to have more teacher demonstrations. Both types of schools in Texas had significantly fewer teacher demonstrations than Oregon public schools. Texas charter and private schools had lower latent ratios than Oregon public schools, but Oregon and Texas public schools were not significantly different. Classroom average pretest TEMA-3 was not predictive of either the latent ratio or latent logged teacher demonstrations. In sum, the hypothesized teacher-level model fit the data well, so we proceeded to modeling the quadratic and linear effects of the latent ratio.

Ratios of Individual Practice Opportunities

For our hypothesized model with a fixed intercept and random slope, the quadratic trend for the latent ratio was very small, close to 0, and not significant (results not presented). That is, within the observed data, we were unable to identify a point at which a higher ratio no longer increased or even decreased mathematics outcomes (i.e., the hypothesized optimal ratio). Consequently, we dropped the quadratic effects and focused on the linear effects of the latent ratio. The linear effect was positive and significant, indicating that higher latent ratios of practice to demonstrations flattened the random slope, providing evidence of differential achievement benefitting students with MD. As a sensitivity analysis, we checked for an interaction between the treatment condition and the latent ratio to ensure that the effect was the same in both ELM and control conditions. The interaction term was not significant, indicating that the effect of the latent ratio was not dependent on the ELM program. Full results of this model are shown in Table 2.

To contextualize these findings, we used the model parameters provided in Table 2 to compute raw and standardized effect sizes. We considered students with MD with a range of very low to low pretest standard scores of 57, 64, 68, and 75 on the TEMA-3, which correspond to the 2nd, 6th, 10th, and 20th percentile values in our data. To facilitate interpretability, we considered, as a reference, typical teachers in Oregon who had an observed rate of teacher demonstrations per minute of .45 (i.e., about 1 demonstration every 2 minutes). Given that there are practical limits on the number of student

Table 1. Descriptive Statistics for Student- and Teacher-Level Variables in the Two-Level Models.

Variable	Occasion	*N*	*M*	*SD*	Min	Max	Skew	Kurt
Student-level variables								
LEP status		2,650	0.29	0.45	0	1	0.93	−1.13
EN-CBM pretest		2,337	75.58	51.79	0	261	0.62	−0.28
TEMA-3 pretest		2,212	90.29	17.06	55	145	0.17	−0.28
TEMA-3 posttest		2,383	101.41	14.73	55	145	−0.11	0.16
Teacher-level variables								
Private		129	0.25	0.43	0	1	1.18	−0.62
Texas		129	0.26	0.44	0	1	1.13	−0.73
Individual practice opportunities (per minute)	Observation 1	59	0.58	0.46	0.00	1.85	1.08	0.69
	Observation 2	127	0.60	0.45	0.03	3.06	1.86	7.08
	Observation 3	128	0.51	0.40	0.00	2.00	1.42	2.42
Teacher demonstrations (per minute)	Observation 1	59	0.67	0.41	0.09	2.08	1.10	1.47
	Observation 2	127	0.57	0.44	0.00	2.38	1.84	4.19
	Observation 3	128	0.55	0.40	0.00	2.17	1.35	2.04
Individual practice–teacher demonstration ratio	Observation 1	59	1.12	0.91	0.26	4.64	1.69	2.93
	Observation 2	127	1.24	0.75	0.12	4.32	1.55	3.36
	Observation 3	128	1.12	0.62	0.16	3.63	0.93	1.28
Individual practice–teacher demonstration log-transformed difference	Observation 1	59	−0.16	0.71	−1.33	1.53	0.50	−0.73
	Observation 2	127	0.04	0.60	−2.12	1.46	−0.39	0.89
	Observation 3	128	−0.06	0.61	−1.82	1.29	−0.47	−0.27
Log-transformed rate of teacher demonstrations	Observation 1	59	−0.18	0.44	−1.09	0.85	−0.07	−0.24
	Observation 2	127	−0.32	0.48	−1.39	0.97	0.33	0.13
	Observation 3	128	−0.34	0.47	−1.39	0.88	0.20	−0.38
Log-transformed rate of individual practice opportunities	Observation 1	59	−0.33	0.54	−1.39	0.74	0.02	−0.66
	Observation 2	127	−0.28	0.49	−1.28	1.20	0.00	−0.36
	Observation 3	128	−0.39	0.50	−1.39	0.81	0.07	−0.24

Note. EN-CBM = early numeracy curriculum-based measure; LEP = dichotomous Limited English Proficiency indicator variable; *N* = sample size at the relevant level of the model; Private = dichotomous charter and private school indicator variable; TEMA-3 = Test of Early Mathematics Ability–3rd Edition; Texas = dichotomous state indicator variable. COSTI-M data were not collected in Texas schools during Observation 1, resulting in a lower sample size.

practice opportunities that can be completed in a minute, we computed the difference in student gains on the TEMA-3 for a classroom that facilitated three individual practice opportunities for every teacher demonstration (i.e., 3:1 ratio) compared to a classroom that provided one practice opportunity for every teacher demonstration (i.e., 1:1 ratio). The differences in gains for pretest scores of 57, 64, 68, and 75 are, respectively, 9.4, 8.4, 7.9, and 7.0 raw points on the TEMA-3. Given the normative *SD* of 15, these result in Hedges' *g* effect sizes of .63, .56, .53, and .47, respectively, which represent medium to medium-large effects by most standards. In other words, if teachers were to facilitate three individual student practice opportunities for every teacher demonstration, the end-of-year mathematics achievement for the initially lowest scoring students would benefit substantially. Based on our data, facilitating three individual student practice opportunities for every teacher demonstration represents a feasible amount of practice that provides substantial benefits with respect to gains in mathematics achievement for students with MD.

Discussion

The purpose of the present study was to explore whether there was a Goldilocks effect for the ratio of individual student practice opportunities to teacher demonstrations during Tier 1 mathematics instruction. We did not find evidence for a Goldilocks effect: The quadratic effect of the latent ratio was small and not significant. In the absence of a Goldilocks effect, we explored whether there was a simpler linear relationship (i.e., more is better) in which a particular number of individual practice opportunities for every teacher demonstration predicted mathematics achievement for students with MD. In this analysis, the latent ratio had a strong effect on the random slope, suggesting more frequent individual practice opportunities per teacher demonstration increases end-of-year mathematics achievement for students with MD. Specifically, for students who are initially low scoring at the start of the kindergarten school year, providing three individual student practice opportunities for every explicit teacher demonstration or explanation provides substantial benefits on gains in mathematics achievement. For example,

Table 2. Two-Level Models of TEMA-3 Outcome Model.

Effect	Variable 1	Variable 2	Est.	*SE*	*p*
Student level					
Regressions	TEMA-3 posttest ON	LEP status	−0.074	0.033	0.024
	EN-CBM pretest ON	LEP status	−0.192	0.033	0.000
		TEMA-3 pretest	0.954	0.016	0.000
Covariance	EN-CBM pretest WITH	TEMA-3 posttest	0.037	0.005	0.000
Means	TEMA-3 pretest		−1.919	0.043	0.000
Intercepts	EN-CBM pretest		1.859	0.033	0.000
Variances	TEMA-3 pretest		0.751	0.028	0.000
Residual variances	TEMA-3 posttest		0.162	0.007	0.000
	EN-CBM pretest		0.245	0.011	0.000
Classroom level					
Regressions	Random slope ON	Demonstration rate	−0.037	0.098	0.703
		Latent difference	−0.162	0.044	0.000
		Texas	0.016	0.044	0.709
		Private	−0.017	0.039	0.672
		TEMA-3 pretest class average	0.004	0.002	0.072
	Demonstration rate ON	Texas	−0.282	0.071	0.000
	Demonstration rate ON	Private	−0.199	0.085	0.020
	Demonstration rate ON	TEMA-3 pretest class average	0.006	0.004	0.139
	Latent difference ON	Texas	0.147	0.103	0.154
		Private	0.323	0.12	0.007
		TEMA-3 pretest class average	−0.004	0.006	0.539
	TEMA-3 posttest ON	Texas	0.191	0.071	0.007
		Private	0.035	0.049	0.472
		TEMA-3 pretest class average	0.012	0.004	0.003
Covariance	Difference WITH	Demonstration rate	−0.022	0.017	0.200
Intercepts	TEMA-3 posttest		6.186	0.036	0.000
	Individual practice rate		−0.334	0.053	0.000
	Demonstrate rate		−0.201	0.04	0.000
	Random slope		0.628	0.02	0.000
Residual variances	Individual Practice 1–3		0.151	0.018	0.000
	Demonstrations 1–3		0.165	0.019	0.000
	Difference		0.101	0.028	0.000
	Demonstration rate		0.038	0.018	0.035
	Random slope		0.005	0.001	0.000

Note. Demonstration rate = rate of teacher demonstrations; EN-CBM = early numeracy curriculum-based measure; latent difference = latent difference between logged individual student practice opportunities and logged teacher demonstrations; LEP = Limited English Proficiency indicator variable; private = charter and private school indicator variable; TEMA-3 = Test of Early Mathematics Ability–3rd Edition; Texas = school district indicator variable.

our model implies that students at the 2nd and 20th percentiles on fall mathematics skill in classrooms that provide a 3:1 student practice to teacher demonstration ratio had Hedges' *g* effect sizes on the TEMA-3 of .63 and .47, respectively, compared to similarly performing students in classrooms with fewer practice opportunities per demonstration.

While preliminary, our findings suggest there may be practical value in teachers trying to facilitate about three individual response opportunities for each demonstration or explanation they provide, particularly when their instruction includes students with MD. However, it is important to note that our results do not address the impact that even more frequent individual practice has on student mathematics achievement, such as four or five practice opportunities for each teacher demonstration. It is possible that, if teachers had provided more practice opportunities, we would have found even higher levels of our targeted student mathematics outcomes. Nevertheless, the present results are important because they suggest that providing up to three practice opportunities can be an effective use of classroom time, especially for students with MD. Though this finding may be subject to qualification along a number of critical dimensions, such as the quality of student practice, it provides a logical starting point for future research and some much-needed insight into the types of practical questions teachers have about intensifying mathematics instruction for students with MD.

Limitations

Several limitations should be considered when interpreting our results. First, there was a limited number of observations conducted per classroom. Although three observations per classroom exceeds the number of observations typically conducted in observational research (Pianta & Hamre, 2009) and large-scale efficacy trials (Clements et al., 2013), additional data may have provided a more robust estimate of student practice. In the larger ELM efficacy trial, the number of observations conducted in each classroom was based on available resources. Second, a total of 16 classrooms reported providing bilingual instruction during the school day. While the mathematics instruction in these particular classrooms was delivered in English, we found no statistically significant differences between bilingual and monolingual classrooms with respect to rates of individual practice opportunities (p = .082). In addition, LEP was included in the model as a predictor of outcomes, primarily as a missing data covariate, but this also suggests that language differences did not interfere with the modeling.

Third, our analysis did not include teachers' provision of academic feedback. While academic feedback is an effective method for providing students with information on their performances with mathematical tasks, it was not considered in the current study because prior research with the COSTI-M suggests that it is highly correlated with individual practice opportunities (Doabler et al., 2018). This correlation is likely a function of how academic feedback is operationalized by the COSTI-M, in which academic feedback is coded after a teacher-prompted student practice opportunity. Fourth, our observation system captures the frequency of student practice but does not document its quality. Although investigation of instructional quality data is highly important (Pianta & Hamre, 2009), the principal investigators of ELM Efficacy Trial prioritized the COSTI-M over a high-inference observation system because it directly maps onto the ELM program's theory of change, which specifies that frequent, appropriately designed practice mediates ELM's impact on student mathematics achievement. Nonetheless, future research should include instructional quality data as the potential utility of student practice may depend on both its quantity and quality.

Finally, the present study did not differentiate practice opportunities by mathematical content or practice type (e.g., verbalizations vs. use of concrete mathematics materials). Instead, we examined instructional content overall (i.e., across both complex and foundational mathematics concepts and skills) and practice as a single category. This decision was based primarily on how the observation data were collected. However, it is plausible that more complex content, such as solving word problems, requires more frequent practice opportunities.

Implications for Research and Practice

While preliminary, our findings have implications on several interrelated fronts. First, the finding that three individual student practice opportunities for every teacher demonstration led to higher levels of end-of-year mathematics achievement for initially low-skilled students aligns with previous research (Clements et al., 2013; Morgan et al., 2015). While practice is important for all students, it is essential for students who receive little exposure to mathematics prior to school entry (Barnes et al., 2016; Clarke et al., 2015). As such, a recommendation is that teachers engage students who enter kindergarten at risk for MD in frequent opportunities to practice with foundational mathematics content. If judiciously integrated with overt teacher demonstrations, practice can help these at-risk students learn new mathematical content and transfer acquired knowledge and skills to solve novel mathematics problems. However, we recognize that there are practical limits on the extent to which a teacher can adjust core mathematics instruction to provide a larger number of individual practice opportunities for every explanation or demonstration. For instance, some amount of teacher explanation is likely needed to initiate individual practice, and there are likely temporal constraints on how long a teacher can spend on a given activity (e.g., due to the school's schedule and fluctuations in student interest and willingness to practice).

Our results also have implications for designing professional development for teachers who work with students with MD. Providing three individual practice opportunities for every demonstration seems like a feasible goal for teachers to strive for in their core mathematics instruction. As such, it seems reasonable that teachers could learn how to increase the amount of individual practice that students with MD receive in general education classrooms. Finally, we encourage researchers to further investigate student practice opportunities. Observation data analyzed in the current study were collected in Tier 1 kindergarten mathematics settings. Future research should consider expanding this line of research into other instructional formats (e.g., small group settings) and grade levels. Exploring for a Goldilocks effect for student practice in other areas may help teachers better support students with MD in becoming mathematically proficient.

Conclusion

The current study represents one of the first efforts to examine for a Goldilocks effect for the ratio of individual student practice opportunities to teacher demonstrations during Tier 1 mathematics instruction. While evidence for such an effect did not surface, our results, while preliminary, did indicate a "more is better" finding, suggesting that students with MD benefit most when core mathematics instruction offers higher ratios of student practice to teacher demonstrations relative to lower ratios. More frequent, explicitly designed practice may allow students with MD to gain a deeper and more lasting understanding of foundational mathematics concepts and skills.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Ben Clarke, Scott K. Baker, and Hank Fien are eligible to receive a portion of royalties from the University of Oregon's distribution and licensing of certain ELM-based works. Potential conflicts of interest are managed through the University of Oregon's Research Compliance Services.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grants R305A150037 and R305A080699 and by the National Science Foundation through Grant 1503161 awarded 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.

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*, 1–20.

Barnes, M. A., Klein, A., Swank, P., Starkey, P., McCandliss, B., Flynn, K., . . .Roberts, G. (2016). Effect of tutorial interventions in mathematics and attention for low-performing preschool children. *Journal of Research on Educational Effectiveness*, *1*, 155–178.

Cepeda, N. J., Pashler, H., vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380. doi: 10.1037/0033-2909.132.3.354

Clarke, B., Baker, S., Smolkowski, K., Doabler, C.T., Strand Cary, M., & Fien, H. (2015). Investigating the efficacy of a core kindergarten mathematics curriculum to improve student mathematics learning outcomes. *Journal of Research on Educational Effectiveness*, *8*, 303–324. doi: 10.1080/19345747.2015.1116034

Clarke, B., Doabler, C. T., Smolkowski, K., Kurtz-Nelson, E., Baker, S., Fien, H., & Kosty, D. (2016). Testing the immediate and long-term efficacy of a tier 2 kindergarten mathematics intervention. *Journal of Research on Educational Effectiveness*, *9*, 607–634. doi: 10.1080/19345747.2015.1116034

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

Clements, D. H., Agodini, R., & Harris, B. (2013). *Instructional practices and student math achievement: Correlations from a study of math curricula* (NCEE Evaluation Brief No. 2013-4020). Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences.

Deshler, D. D. (2015). Moving in the right direction but at what speed, and how smoothly? *Remedial and Special Education*, *36*, 72–76.

Doabler, C. T., Baker, S. K., Kosty, D., Smolkowski, K., Clarke, B., Miller, S. J., & Fien, H. (2015). Examining the association between explicit mathematics instruction and student mathematics achievement. *Elementary School Journal*, *115*, 303–333.

Doabler, C.T., Stoolmiller, M., Kennedy, P. C., Nelson, N. J., Clarke, B., Gearin, B., Fien, H., Smolkowski, K., & Baker, S. K. (2018). Do components of explicit instruction explain the differential effectiveness of a core mathematics program for students with mathematics difficulties? A mediated moderation analysis. *Assessment for Effective Intervention*. doi: 10.1177/1534508418758364

Donovan, M. S., & Bransford, J. D. (Eds.). (2005). *How students learn: History, mathematics and science in the classroom*. Washington, DC: National Academies Press.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*, 4–58.

Ericsson, K. A., Roring, R., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies*, *18*, 3–56. doi: 10.1080/13598130701350593

Fields, R. D. (2005). Myelination: An overlooked mechanism of synaptic plasticity? *The Neuroscientist*, *11*, 528–531.

Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson, N. J., & Chapparro, E. (2015). An examination of the efficacy of a multitiered intervention on early reading outcome for first grade students at risk for reading difficulties. *Journal of Learning Disabilities*, *48*, 602–621. doi: 10.1177/0022219414521664

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks Cole.

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.

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

Ginsburg, H. P., & Baroody, A. J. (2003). *Test of early mathematics ability–third edition (TEMA-3)*. Austin, TX: ProEd.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1984). Opportunity to respond and student academic performance. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 58–88). Columbus, OH: Merrill.

Kame'enui, E. J. (1993). Diverse learners and the tyranny of time: Don't fix blame; fix the leaky roof. *The Reading Teacher*, *46*, 376–383.

Klein, A. G., & Stoolmiller, M. (2003). Detecting latent interaction effects in behavioral data. *Methods of Psychological Research*, *8*, 113–126.

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

MacSuga-Gage, A. S., & Gage, N. A. (2015). Student-level effects of increased teacher-directed opportunities to respond. *Journal of Behavioral Education*, *24*, 273–288.

Morgan, P. L., Farkas, G., & Maczuga, S. (2015). Which instructional practice most help first-grade students with and without mathematics difficulties. *Educational Evaluation and Policy Analysis*, *37*, 184–205. doi: 10.3102/0162373714536608

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.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (Version 7.31). Los Angeles, CA: Muthén & Muthén.

Paumgarten, N. (2017, November 27). Confidence game: The intense discipline of a skiing prodigy. *The New Yorker*, 36–45.

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

Raykov, T. (1992). Structural models for studying correlates and predictors of change. *Australian Journal of Psychology*, *44*, 101–112.

Ritter, F. E., & Schooler, L. J. (2001). The learning curve. In W. Kintch, N. Smelse, & P. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (Vol. 13, pp. 8602–8605). Amsterdam: Pergamon.

Simmons, D. (2015). Instructional engineering principles to frame the future of reading intervention research and practice. *Remedial and Special Education*, *36*, 45–51.

Skinner, C. H., Belfiore, P. J., Mace, H. W., Williams-Wilson, S., & Johns, G. A. (1997). Altering response topography to increase response efficiency and learning rates. *School Psychology Quarterly*, *12*, 54–64.

Skinner, C. H., Ford, J. M., & Yunker, B. D. (1991). A comparison of instructional response requirements on the multiplication performance of behaviorally disordered students. *Behavioral Disorders*, *17*, 56–65.

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*, *44*, 48–57.

Sood, S., & Jitendra, A. K. (2013). An exploratory study of a number sense program to develop kindergarten students' number proficiency. *Journal of Learning Disabilities*, *46*, 328–346. doi:10.1177/0022219411422380

Sutherland, K. S., Alder, N., & Gunter, P. L. (2003). The effect of varying rates of opportunities to respond to academic requests on the classroom behavior of students with EBD. *Journal of Emotional and Behavioral Disorders*, *11*, 239–248.

Vaughn, S. (2015). Building on past successes: Designing, evaluating, and providing effective treatments for persons for whom typical instruction is not effective. *Remedial and Special Education*, *36*, 5–8.