Published in final edited form as:

Except Child. 2017 July ; 83(4): 359–377. doi:10.1177/0014402917690728.

Child-Level Predictors of Responsiveness to Evidence-Based Mathematics Intervention

Sarah R. Powell¹, Paul T. Cirino², and Amelia S. Malone³

¹University of Texas at Austin

²University of Houston

³Vanderbilt University

Abstract

We identified child-level predictors of responsiveness to 2 types of mathematics (calculation and word-problem) intervention among 2nd-grade children with mathematics difficulty. Participants were 250 children in 107 classrooms in 23 schools pretested on mathematics and general cognitive measures and posttested on mathematics measures. We assigned classrooms randomly assigned to calculation intervention, word-problem intervention, or business-as-usual control. Intervention lasted 17 weeks. Path analyses indicated that scores on working memory and language comprehension assessments moderated responsiveness to calculation intervention. No moderators were identified for responsiveness to word-problem intervention. Across both intervention groups and the control group, attentive behavior predicted both outcomes. Initial calculation skill predicted the calculation outcome, and initial language comprehension predicted word-problem outcomes. These results indicate that screening for calculation intervention should include a focus on working memory, language comprehension, attentive behavior, and calculations. Screening for word-problem intervention should focus on attentive behavior and word problems.

Keywords

responsiveness; intervention; mathematics; calculations; word problems

When a well-designed and competently executed randomized controlled trial produces statistically significant effects favoring the learning outcomes of children who receive a specified intervention over children who do not receive that intervention, the intervention is deemed *evidence based* (Every Student Succeeds Act, 2015). This status suggests that most children responded well to the experimental intervention. Evidence-based status does not, however, mean that all children respond adequately (Cook & Cook, 2013). In fact, few if any interventions achieve universal response, and little is known about child characteristics associated with inadequate response. This is especially the case in mathematics, where few

Correspondence concerning this article should be addressed to Sarah R. Powell, 1 University Station D5300, Austin, TX 78712. srpowell@austin.utexas.edu.

Author Note

Sarah R. Powell, Department of Special Education, University of Texas at Austin; Paul T. Cirino, University of Houston; Amelia S. Malone, Vanderbilt University.

studies have examined child-level variables associated with responsiveness in an at-risk sample. To identify child-level response variables, the relevant moderation effect involves a prerequisite cognitive ability or academic skill. With a prerequisite response moderator, the effect of intervention depends on the child's pretest cognitive ability or academic skill. That is, children with low pretest performance respond less adequately than children with greater performance in the cognitive or academic domain.

One research team has investigated whether domain-general and domain-specific variables moderated the effect of intervention. For example, in examining whether responsiveness to a fraction intervention was associated with start-of-fourth-grade whole-number calculation skill, Fuchs, Sterba, et al. (2016) did not find a significant moderator effect on posttest fraction understanding or calculation performance. Children with a range of whole-number calculation skill at pretest responded similarly to the intervention. By contrast, Fuchs, Malone, et al. (2016) determined that responsiveness to fraction word-problem intervention was associated with start-of-fourth-grade reasoning ability. That is, although most children responded adequately, those with very low reasoning did not.

Understanding how child-level variables influence children's responsiveness to intervention may help school personnel customize interventions to address at-risk children's needs. One implication of Fuchs, Malone, et al. (2016) is that children with very low reasoning ability require a different form of or a more intensive version of fraction word-problem intervention (e.g., more varied practice, hands-on materials to establish fraction concepts, a self-regulation component for attention, more intervention time). Such studies can also help researchers identify directions for expanding the overall efficacy of the intervention.

Extensions to the Literature

With the present study, we sought to extend prior work in four important ways. First, we contrasted child-level predictors of responsiveness to intervention on two forms of mathematics intervention: whole-number calculations and word problems. Each area represents a critical form of mathematics competence for success in school (National Council of Teachers of Mathematics, 2014). A calculation problem is set up for solution, whereas a word problem requires children to process text to build a problem model and then construct a number sentence for calculating the unknown. Along with this transparent difference in the nature of these tasks, correlational studies find that a different set of cognitive abilities underlies calculation and word-problem skill in representative samples (e.g., Swanson, 2011). This suggests that the child-level variables determining responsiveness to intervention may differ for these types of mathematics outcomes.

Our second extension to the literature concerned our target population: children with concurrent low performance on calculations and word problems. This focus is important because children with concurrent difficulty have more severe deficits in each domain than do children with difficulty in one area (e.g., Cirino, Fuchs, Elias, Powell, & Schumacher, 2015; Fuchs, Fuchs, Stuebing, et al., 2008; Willcutt et al., 2013). Concurrent difficulty therefore serves as an indicator of serious risk for poor mathematics learning. Moreover, approximately twice as many children have concurrent difficulty than difficulty in one area

(Fuchs, Fuchs, Stuebing, et al., 2008). In this study, we refer to this population as children with *mathematics difficulty* (MD).

Our third extension was that our intervention relied on a multi-tier support system, in which children with MD participated in instruction that was aligned across a whole-class intervention and a supplementary small-group tutoring program. In providing aligned instructional support, the hope was that small-group tutoring worked synergistically with whole-class mathematics instruction (Fuchs, Fuchs, Craddock, et al., 2008; Lembke, Garman, Deno, & Stecker, 2010). This context is important because a multi-tier framework assumes that whole-class intervention, as well as small-group tutoring, are strong. In the present study, this was the case for children with MD (Powell et al., 2015) and for the general population (Fuchs, Powell, et al., 2014). In the present study, we focused on a subset of children with MD who received both instructional tiers.

The final extension was methodological, involving the use of path analysis. *Moderation* refers to a statistical interaction between (a) a child-level variable (i.e., cognitive ability or mathematics skill) measured at the start of a study and (b) the effect between two study conditions on the study's academic outcome. With this interaction, the effect between two study conditions on the outcome is larger or smaller depending on pretest performance on an academic or cognitive variable. In most studies examining child-level variables associated with response, researchers apply cut-points to children's post-intervention performance (or improvement over the course of intervention) to classify children as adequate versus inadequate responders. Then, researchers identify which child-level variables differ significantly for the adequate and inadequate responder groups. The drawbacks of this approach are that the selected cut-point is arbitrary, and results vary with the stringency of the arbitrary cut-point (e.g., Fuchs, Compton, et al., 2008; Wise et al., 2008). Branum-Martin, Fletcher, and Stuebing (2012) therefore argued for an approach that treats performance continuously rather than categorically. Path analytic moderation analysis addresses Branum-Martin et al.'s concerns in two ways. By treating outcome performance continuously, it avoids discarding potentially informative data on the full-distribution of at-risk performance. Second and relatedly, it empirically derives the cut-point for responsiveness, as the score on the moderator variable at which the effect of intervention transitions from significance to nonsignificance.

Rationale for the Potential Moderators

As potential candidates for moderators, we relied on the individual-differences literature to identify child-level variables that predict development in calculations or word problems. Some child-level variables predict *development* in an academic area *across* environmental, instructional, or intervention conditions. In this article, we use the term *predictors* to refer to variables that exert a "main" effect of a child-level variable on the outcome. By contrast, when a child-level variable *interacts* with the effect between two conditions, the correlation between the child-level variable and the outcome is significantly stronger in one condition than in the other. For this reason, the child-level variable is associated with the magnitude of the effect *between* the two conditions. This is what we mean when we use the term *moderation*.

In the literature, predictors and moderators have been conceptualized in terms of domain-general abilities (e.g., working memory), which support performance across many academic areas, and domain-specific skills (i.e., understanding about discrete numbers), which more specifically support performance within an academic area. Geary's (2004) model of mathematics learning indicated that mathematics performance depends on a constellation of domain-general abilities and foundational mathematical skills. Our study reflects both assumptions in order to understand how response variables related to cognition and academics may moderate responsiveness to intervention.

In the mathematics literature for children with MD, the following domain-general cognitive processes predicted individual differences in the development of procedural calculation and word-problem skill among representative samples across environmental, instructional, or intervention conditions. Reasoning (Fuchs, Geary, et al., 2010), working memory (e.g., Geary & Widaman, 1992), and attentive behavior (e.g., Fuchs et al., 2006) predicted competency with calculation and word-problem learning. Processing speed predicted calculation learning (e.g., Fuchs, Geary, et al., 2010), and language comprehension predicted word-problem learning (e.g., Zheng, Swanson, & Marcoulides, 2011). Based on prior work, we hypothesized main effects for reasoning, working memory, and attentive behavior on both outcomes; for processing speed on the calculation outcome; and for language comprehension on the word-problem outcome. In terms of domain-specific skills, we expected main effect predictors of end-of-second-grade calculation learning to include pretest fluency with facts and understanding about whole-number magnitude (e.g., Fuchs et al., 2012). In the prediction of end-of-second-grade word-problem skill, we anticipated a main effect predictive value for pretest word-problem skill as well as understanding about number (e.g., Fuchs et al., 2012).

Extant Database for Data Analysis

We conducted our analyses on an extant database (Powell et al., 2015) documenting the evidence-based status of a multi-level calculation intervention and a multi-level word-problem intervention for children with MD. In that study, teachers were randomly assigned to calculation multi-tier intervention, word-problem multi-tier intervention, or a business-as-usual condition. Multi-tier calculation intervention improved performance on the calculation outcomes, with an effect size (ES) of 1.67 on the proximal outcome and 1.19 on the distal outcome. Multi-tier word-problem intervention enhanced performance on the proximal word-problem outcome, with an ES of 1.31, but the effect on the distal outcome was not significant (ES = 0.15). In the current set of analyses, we indexed responsiveness at the end of second grade by creating a factor score across proximal and distal measures of calculation performance and a factor score across proximal and distal measures of word-problem performance. Note that an additional study (Fuchs, Powell, et al., 2014) reported effects of the interventions with a sample that included at-risk and not-at-risk children.

Summary of Purpose

The purpose of the present analyses was to identify child-level predictors of responsiveness to two types of mathematics intervention among second-grade children with MD.

Intervention involved an evidence-based, supplementary, small-group instructional program that was aligned with an evidence-based classroom instructional program. To investigate responsiveness, we applied moderation analyses to the extant database of at-risk children, for which prior analyses indicated overall efficacy (Fuchs, Powell, et al., 2014; Powell et al., 2015). We examined eight child-level domain-general cognitive abilities and domain-specific mathematics skills as potential moderators. In isolating the effect of each moderator (an interaction between the child-level variable and the effect of intervention), we controlled for the effects of the other seven variables on the outcome.

Method

Participants

We received Institutional Review Board approval and obtained signed consent forms from teachers, consent forms from parents or guardians, and assent from children. Stratifying by school, classroom teachers were randomly assigned to three conditions: ~37.5% to calculation intervention; ~37.5% to word-problem intervention; and ~25% to control. Across four cohorts (Powell et al., 2015), we screened 1,917 children in 127 second-grade classrooms in 25 schools in a metropolitan school district using the following screening protocol to identify for MD (i.e., current difficulties with calculation and word problems): Children had to score <7 on the calculation screening measure (i.e., performance on addition sums), <6 on the word-problem screening measure (i.e., performance on single-digit word problems), and greater than the 9th percentile on at least one subtest of the two-subtest *Wechsler Abbreviated Scale of Intelligence* (WASI; Psychological Corporation, 1999). The benchmarks for the screeners were calculated using data from a pilot study where we identified pretest cut-off scores predictive of low performance on calculation and word-problem outcomes. For this study, we adjusted these benchmarks downward so that we only included students below the 40th percentile in both areas. This yielded 265 children from 110 second-grade classrooms in 25 schools.

Note that children were identified as MD in the same way regardless of condition. These children with MD comprised the lowest 14% of the entire screened sample. Of these children, 96 were in calculation intervention (36%), 91 in word-problem intervention (34%), and 78 in the control group (29%). For the present analyses, we eliminated 15 children with incomplete data on the eight child-level variables investigated as potential moderators. The 250 children with complete data were in 107 classrooms in 23 schools, and 91 were in calculation intervention (36%), 90 in word-problem intervention (36%), and 69 in the control group (28%). Table 1 features demographic information and mathematics performance by condition.

Screening Measures

We used three measures to screen for MD status. At pretest, we administered four subtests of the *Second-Grade Calculations Battery* (SGCB; Fuchs, Hamlett, & Powell, 2003): *Sums to 12*, *Sums to 18*, *Minuends to 12*, and *Minuends to 18*. All four subtests required single-digit addition or subtraction. For each subtest, children had 1 min to complete up to 25 problems.

Score was the number of correct answers. Cronbach's α ranged from .85 to .93. We used *Sums to 12* as the calculation screening measure for determining MD status.

Story Problems (Jordan & Hanich, 2000) comprised 14 problems representing combine, compare, and change word-problem types that involve single-digit addition or subtraction for solution. The test was untimed. The examiner read each word problem aloud as the children followed along on paper, and then the examiner provided time for children to write answers. A child's score was the number of correct answers. Cronbach's α was .87.

The *WASI* (Psychological Corporation, 1999) is an individually administered measure of general cognitive ability with two subtests. *Vocabulary* measured word knowledge and expressive language with 4 picture items and 37 vocabulary words. Children defined words until the end of the test or until reaching a ceiling of five consecutive errors. Split-half reliability is .86. *Matrix Reasoning* measured nonverbal reasoning with pattern completion, classification, analogy, and serial reasoning tasks. Children chose the best of five choices to complete a visual pattern for 35 items or until reaching a ceiling of four consecutive errors or four errors over five consecutive items. Reliability, as presented in the manual, is .94. The two subtest scores combine to yield an Estimated Full Scale IQ score with reliability of .92.

Moderator Measures

Addition and subtraction—To index initial calculation skill, we used two subtests of the *SGCB* (Fuchs et al., 2003). *Sums to 12* comprised 25 addition fact problems with sums from 6 to 12. Children had 1 min to write answers. *Minuends to 12* comprised 25 subtraction fact problems with minuends from 5 to 12. Children write differences for 1 min. Cronbach's α was .95.

Story problems—To index initial word-problem skill, we used *Story Problems* (Jordan & Hanich, 2000), as already described previously.

Attentive behavior—*SWAN* is an 18-item teacher rating scale (Swanson et al., n.d.) that sampled items from the Diagnostic and Statistical Manual of Mental Disorders-IV (American Psychiatric Association, 1994) criteria for Attention-Deficit/Hyperactivity Disorder for inattention (items 1 through 9) and hyperactivity or impulsivity (items 10 through 18). Items were rated from 1 (far below) to 7 (far above). We report data for the attentive behavior subscale as the average rating across the nine relevant items. We selected this subscale to index attentive behavior or the ability to maintain focus of attention. The *SWAN* correlates well with other dimensional assessments of behavior related to attention. Cronbach's α in this study was .98.

Processing speed—With *Woodcock-Johnson (WJ)-III Visual Matching* (Woodcock, McGrew, & Mather, 2001), children located and circled two identical numbers in each row of six numbers. They had 3 min to complete 60 rows. As per the test developer, reliability is .91.

Reasoning—To index initial reasoning, we used the Matrix Reasoning subtest of the *WASI* (Psychological Corporation, 1999), as already described previously.

Working memory—We used two dual-task central executive subtests from the *Working Memory Test Battery for Children* (WMTB-C; Pickering & Gathercole, 2001). Each had six items at span levels from 1 through 6 to 1 through 9. Passing four items at a level moved the child to the next level. At each span level, the number of items to be remembered increased by one. Failing three items terminated the subtest. For *Listening Recall* (referred to in this paper as *working memory-sentences*), the child determined if each sentence in a series was true; then recalled the last word in each sentence. For *Counting Recall* (referred to in this paper as *working memory-counting*), the child counted a set of 4, 5, 6, or 7 dots, each on a separate card. After the last card, the child recalled the number of dots on each of the cards. As per the test developer, test-retest reliability is .91 to .93. We opted to include both subtests, rather than creating a composite variable, based on prior work suggesting individual differences in working memory for numbers versus words (Raghubar, Barnes, & Hecht, 2010).

Language comprehension—The *Woodcock Diagnostic Reading Battery* (WDRB) *Listening Comprehension* (Woodcock, 1997) measured the ability to understand sentences or passages. With 38 items, children supplied the word missing at the end of sentences or passages that progressed from simple verbal analogies and associations to discerning implications. The test manual provided examples of correct responses to guide scoring. Reliability, as reported by the test developer, is .80.

Number line—With *Number Line Estimation* (Siegler & Booth, 2004), children located Arabic numerals on a horizontal number line marked with 0 and 100 as endpoints. A number line was presented with a target number shown above the line. Children placed the target number on the number line, without a time limit. The examiner explained, "The number 50 is half of 100, so we put it halfway in between 0 and 100 on the number line." Use of these directions produced findings comparable to other work (e.g., Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007). Siegler and Opfer (2003) used group-level median placements fitted to linear and log models to make inferences about children's modal representation for making placements but used an accuracy measure for individual difference analyses. The accuracy measure was a good indicator of overall understanding of the number line and predicted later mathematics achievement (Geary, 2011). Lower scores indicated stronger performance, but we multiplied scores by -1 before conducting analyses. Test-retest reliability at second grade has been reported at .87 (Fuchs, Geary, et al., 2014).

Outcome Measures

Calculations—We used three measures. First, we indexed single-digit calculation skill with the three non-screening subtests of the *SGCB* (Fuchs et al., 2003): *Sums to 18*, *Minuends to 12*, and *Minuends to 18*. The outcome was the average of these three subtests. Second, to measure proximal double-digit calculation, we administered *Two-Digit Addition* and *Two-Digit Subtraction* of the *SGCB* (Fuchs et al., 2003). For each subtest, children had 3 min to complete up to 20 two-digit plus or minus two-digit problems. Cronbach's α for Two-Digit Addition and Two-Digit Subtraction was .96 and .87, respectively. Third, we measured distal effects with *KeyMath-Revised* (KM; Connolly, 1998) *Addition* and *Subtraction*. On KM Addition, children answered six problems with pictures and 12

problems in a written format. Problems ranged from one-digit addition to addition with decimals, fractions, and negative numbers. KM Subtraction functioned similarly with six problems presented with graphics and 12 written problems of one-, two-, and three-digit subtraction with whole numbers, fractions, decimals, and negative numbers. We administered both KM subtests individually. Cronbach's α for the two subtests was .84 and .81, respectively. Analysis indicated a single-factor solution across these seven tests (three SGCB subtests, Two-Digit Addition and Two-Digit Subtraction, and KM Addition and KM Subtraction) was a good fit to the posttest data.

Word problems—We used three measures to index word-problem performance. *Second-Grade Word Problems* (Fuchs, Zumeta, et al., 2010), the proximal measure, included 18 word problems representing combine, compare, and change problems types. Solutions required one-digit addition or subtraction. In groups, the examiner read each problem aloud as children followed along on paper. Children had 1 min to write a constructed response. Children earned credit for correct word-problem calculation and label to reflect child's processing of the problem statement and understanding of the problem's focus. Cronbach's α was .88. We measured distal effects with *KM Problem Solving* (Connolly, 1998) and the Iowa Test of Basic Skills - Data Interpretation and Problem Solving (ITBS; Hoover, Hieronymous, Dunbar, & Frisbie, 1993). KM Problem Solving included 18 word problems of increasing difficulty, which involved all four operations on word problems with transparent solution strategies; non-routine problems without clear solution strategies; and items requiring children to demonstrate comprehension of a word problem without solving it. Administration was individual, and items were presented verbally with accompanying pictures. Children answered with a verbal response. Testing was discontinued after three consecutive errors. The score was the number of correct multiple-choice answers. Cronbach's α was .74. ITBS was administered in groups. Children worked on eight word problems presented verbally and 22 written problems. Problems included one- and two-digit calculation with and without irrelevant information, with and without charts or graphs, and with and without multiple steps. The examiner read problems, and children responded using a multiple-choice format. Cronbach's α was .81. Analysis indicated a single-factor solution across measures represented a good fit to the posttest data.

Interventions

Children participated in multi-tiered calculation intervention, multi-tiered word-problem intervention, or business-as-usual control. The general education instructional tier comprised 34 whole-class lessons (two lessons per week over a 17-week period). Each lesson lasted 40 to 45 min. Research assistants (RAs) delivered all lessons in the children's classrooms, and classrooms included 15 to 26 students. Small-group intervention comprised 39 lessons. Tutoring started at the start of weeks 4 or 5 of the general education series of lessons and lasted 13 weeks (3 lessons per week). Another team of RAs provided small-group intervention in groups of 2 to 3 children, and each lesson lasted 25 to 30 min. Small-group intervention occurred outside of the child's classroom and did not occur at the same time in the school day as the general education instructional tier. That is, children received the combination of whole-class instruction plus small-group intervention. All RAs were

working on or had finished graduate degrees in education-related fields and received extensive training before working with children.

Calculation framework—Whole-class and small-group calculation intervention incorporated two areas of emphasis for developing conceptual and procedural competence with one- and two-digit addition and subtraction problems (e.g., Fuchs et al., 2013; Groen & Resnick, 1977). We focused on interconnected knowledge about number and number families. Children used number families to explore the inverse relationship between addition and subtraction. We also included instruction on place value concepts of tens and ones. The other area for emphasis was practice. Children learned counting strategies for solving one-digit problems and two- plus or minus one-digit problems without regrouping. Practice helped build fluency, which, in turn, helped improve double-digit addition and subtraction. RAs also explicitly taught efficient procedures for identifying when regrouping was required in addition and subtraction problems and how regrouping differed in addition and subtraction.

Word-problem framework—Word-problem intervention incorporated Kintsch and colleagues' framework in which word-problem solving is thought to engage problem solving and language comprehension (Cummins, Kintsch, Reusser, & Weimer, 1988; Kintsch & Greeno, 1985). This model of word-problem solving suggests general features of text apply across all word problems but that word problems differ by problem type. Children learn to apply a schema to assist with solving specific word problems, and at second grade, the three whole-number additive schemas include combine, compare, and change (Riley & Greeno, 1988).

Fidelity—To assess fidelity, all sessions were digitally audio-recorded, and we conducted in-school observations of each RA delivering intervention. Prior to the first session, we created checklists of essential information for each lesson within calculation intervention and word-problem intervention. At the end of each school year, RAs independently listened to a random while completing checklists to identify the percentage of points addressed. For small-group tutoring, the mean percentage of points addressed exceeded 95 in each condition.

Additional information—For additional information on teachers' (non-evidence-based) classroom instruction, the structure of calculation and word-problem whole-class intervention, the structure of calculation and word-problem small-group intervention, the RAs, RA training, and fidelity, see Fuchs, Powell, et al. (2014) and Powell et al. (2015).

Procedure

We administered the screening measures in September and the pretest measures in October. RAs delivered whole-class intervention from November to March, and a separate team of RAs delivered small-group intervention from December to March. We posttested children in late March. On individual measures, we audio-recorded each administration, and sampled 20% of tapes for rescoring. Agreement exceeded 99%. RAs independently entered responses on 100% of the test protocols for each measure on an item-by-item into two separate

databases. We rectified the discrepancies between the two databases to reflect the child's original response.

Data Analysis

Prior to conducting analyses, we transformed scores on each measure to sample-based z -scores. The moderation analyses tested whether a pretest skill or ability interacted with the main effect contrasting the intervention against the control group (and thereby determined children's responsiveness to the intervention). Following Preacher and Hayes (2008; Hayes, 2013), we used an ordinary least squares path analytical framework. Moderation analysis, based on a path analytic framework, has been used previously to investigate responsiveness to intervention (e.g., Fuchs, Malone, et al., 2016). The Preacher and Hayes analyses permitted (a) conclusions about which pretest child-level variables interact with the main effects of intervention and (b) empirical information about where along the distribution of that pretest ability or skill the effect of intervention transitions from significant to nonsignificant. (Note that Barrett, 2007, suggested a minimum sample size of 200, whereas Kline, 2016, indicated a minimum of 10 subjects per parameter. Although these analyses did not meet the Barrett criterion, they did meet the Kline guideline [i.e., 11 parameters for each model, with sample sizes of 159 and 160]. Moreover, allowances were made for populations that are restricted in size, as in this case.)

Eight pretest mathematics skills and cognitive abilities were evaluated as potential moderators (i.e., determinants of children's responsiveness to intervention): two pretest mathematics skills (single-digit addition and subtraction fluency or story problems; number line understanding) and six pretest cognitive abilities (reasoning, attentive behavior, processing speed, working memory-sentences, working memory-counting, and language comprehension). In each analysis, one of the eight pretest variables was tested as the moderator, while controlling for the other seven pretest variables. Each of these moderation analyses was a test of a statistical interaction between (a) the intervention effect (intervention versus control group) and (b) the pretest ability or skill (treated continuously). A significant interaction or moderation indicated that the level of the child's performance on the pretest variable was associated with responsiveness to intervention.

To follow-up each significant interaction, we relied on the Johnson-Neyman technique (Bauer & Curran, 2005; Hayes & Matthes, 2009), which derived the value along the full continuum of the pretest ability or skill at which the effect of intervention transitioned from statistically significant to nonsignificant. In this way, the regions of significance for the intervention effect were empirically derived, and corresponded to values of the pretest ability or skill at which children demonstrated adequate response to the intervention: children with pretest scores exceeding the value (empirically derived cut-point) at which the effect of intervention transitioned to significance. The regions of nonsignificance corresponded to values of the pretest ability or skill at which children did not benefit from the intervention: children with pretest scores falling below the value (empirically derived cut-point) at which the effect of intervention transitioned to nonsignificance. Each significant moderation (interaction) effect demonstrated that a pretest ability or skill was linked to children's

responsiveness to intervention and provided the cut-point on the pretest variable associated with responsiveness.

This approach did not require arbitrarily set cut-scores for designating responsiveness. Instead, the Johnson-Neyman analysis empirically derived that point by testing the difference between intervention and control groups for every value of the pretest ability or skill. For each significant moderator, we provide a figure to help readers visualize the nature of the moderator (interaction or responsiveness) effect.

Results

Moderators of Responsiveness to Calculation Intervention

Table 2 shows the raw and standard score means, *SDs*, and correlations for variables included in the analyses. Table 3 features results of the responsiveness (moderation) analyses. We identified two significant moderator effects. Working memory-sentences moderated the main effect between the calculation intervention versus the control group on the calculation outcome. For this model, R^2 was .62, $F(10,149) = 23.86$, $p < .001$. The path coefficient for the interaction was 0.22 (S.E. = 0.10), $t = 2.15$, $p = .03$. To probe this interaction between (a) working memory-sentences and (b) the effect of calculation intervention over the control group, we relied on the Johnson-Neyman technique (Bauer & Curran, 2005; Hayes & Matthes, 2009). This technique derives the value, along the full continuum of the moderator, at which the effect of X on Y transitions from statistical significance to nonsignificance. Figure 1, top panel, is a visualization of this interaction. The x-axis shows illustrative points along the distribution of working memory percentile ranks for the present sample. As shown, the effect of the calculation intervention was smaller for children with lower working memory compared to children with more adequate working memory. That is, the difference between the black bars (showing the calculation intervention group's scores) and white bars (showing the control group's scores) decreased as children's working memory capacity decreased. Even so, the effect favoring the calculation intervention over the control group remained significant.

The second moderator of the main effect favoring calculation intervention over the control group involved language comprehension. For the model, R^2 was .62, $F(10,149) = 23.82$, $p < .001$. The path coefficient for the interaction was 0.22 (S.E. = 0.11), $t = 2.11$, $p = .04$. We again relied on the Johnson-Neyman technique to probe this interaction. Figure 2 is a visualization of this interaction. The x-axis shows illustrative points along the distribution of language comprehension percentile ranks for the present sample. As shown, the effect of the calculation intervention was smaller for children with lower language comprehension compared to children with more adequate language comprehension: The difference between the black bars (i.e., intervention group) versus the white bars (i.e., control group) decreased as children' language comprehension decreased. The effect favoring the calculation intervention over the control group transitioned from significant to nonsignificant at 2.79 *SDs* below the *sample* mean on language comprehension.

Moderators of Responsiveness to Word-Problem Intervention

We ran eight analogous models contrasting the word-problem intervention group against the control group on the word-problem factor outcome score, but substituting pretest word-problem skill (i.e., *Story Problems*) for single-digit fluency (see Tables 2 and 3). We identified no moderators of the main effect favoring word-problem intervention over the control group.

Predictors of Calculation and Word-Problem Learning Across Groups

Across intervention and control groups, start-of-second-grade teacher rating of attentive behavior was a significant predictor of both mathematics outcomes (see Table 3). Also, initial calculation skill was a significant predictor of the calculation outcome, and pretest language comprehension was a significant predictor of the word-problem outcome.

Discussion

With the present set of analyses, our goal was to deepen insight into the child-level variables that contribute to inadequate response to evidence-based mathematics intervention. Few studies have addressed this issue in mathematics, and the present set of analyses extends prior work in four ways. First, we investigated potential domain-general and domain-specific moderators of responsiveness to intervention for two contrasting conditions – calculations and word problems. Proficiency with calculations and word problems are foundational for future mathematics success (Fuchs et al., 2012; Vukovic et al., 2014), yet little is known about potential moderators of success in these areas among second-grade children with MD. We investigated eight potential moderators of intervention.

Second, our target population included children with concurrently low calculation and word-problem performance. These children are especially at risk for more severe mathematics deficits (e.g., Fuchs, Fuchs, Stuebing, et al., 2008) and tend to experience poorer outcomes on higher-order mathematics than children with deficits in just one area of mathematics. Identifying potential moderators of intervention for these groups of children has important implications for designing future effective interventions.

Third, both active treatment conditions involved a multi-tier support system, in which children with MD participated in instruction that was aligned across the general education classroom and the supplementary small-group intervention program. That is, children in the calculation or word-problem condition received whole-class instruction combined with small-group intervention. This is a novel design, as most responsiveness-to-intervention randomized control trials do not include two evidence-based tiers of support for children with MD, even though multi-tier systems of support specify two strong instructional tiers.

The fourth extension was relying on path analytic moderation analysis, which treats child-level variables continuously and empirically derives cut-points for responsiveness. When a child-level variable interacts with an effect between an intervention and control condition, the correlation between the child-level variable and the outcome is stronger in one than in the other condition, and the child-level variable is associated with the magnitude of the effect between the two conditions. We considered eight cognitive abilities and mathematics

skills as potential moderators, and in each analysis, we controlled for the effect of the other seven.

For analyses contrasting the calculation intervention against the control condition on the calculation outcome, we identified two significant moderator effects. The first was working memory. As visualized in the top panel of Figure 1, the control group did similarly poorly, regardless of children's working memory capacity; that is, the correlation between working memory and outcome score was not significant in the control group. The condition in which children received the multi-level calculation intervention performed substantially better than the control group. This was true throughout the range of initial working memory scores.

Yet, in the calculation intervention, the correlation between working memory and the posttest calculation score was significant: Intervention outcomes were stronger for children with greater working memory capacity than for intervention children with less working memory capacity. With this correlation in the intervention condition, the ES favoring calculation intervention over the control group decreased as working memory decreased. At the 90th percentile of the sample's distribution of working memory, the ES was 1.45; at the 10th percentile, 0.89. Although the effect favoring the calculation intervention over the control group remained statistically significant within the observable range of this moderator, this pattern along with the significance of the interaction effect, indicates that working memory is associated with responsiveness to intervention. This finding illustrates the importance of moderator analyses even for generally efficacious interventions. Although effects were strong throughout the range of performance, findings suggest that schools consider additional strategies to address the needs of children with extremely limited working memory capacity, just as results indicate that researchers might further strengthen the intervention to address the needs of all learners.

The nature of the second moderator effect, involving language comprehension, was similar. As visualized in Figure 2, the control group did similarly poorly, regardless of children's language comprehension ability, with a nonsignificant correlation between language comprehension and the outcome score. In the calculation condition, outcome scores were substantially better than for the control group throughout the range of initial language comprehension performance, but the correlation between language comprehension and posttest calculation score was significant: Intervention outcomes were stronger for children with more intact language comprehension than children with weaker language comprehension. The ES favoring the calculation condition over the control group decreased as the sample's language comprehension scores decreased. At the 90th percentile, the ES was 1.51; at the 10th percentile, 0.94, and the effect of intervention over control transitioned from significance to nonsignificance at 2.79 *SDs* below the sample's mean on language comprehension. Here, too, although effects were strong throughout the range of performance, findings suggest that schools incorporate additional strategies to address the needs of specific children, this time those with extremely low language comprehension ability. Furthermore, researchers need to strengthen the intervention to accommodate the needs of all children with MD.

To identify directions for strengthening the calculation intervention, it is important to consider why children's working memory and language comprehension might be associated with responsiveness in this domain. Intervention focused on interconnected knowledge about number, involving ideas and principles about whole numbers, and the place value system. It also provided children with counting strategies for solving one-digit problems and two- plus or minus one-digit problems without regrouping. Conveying these ideas and strategies depended heavily on the RA teacher's and RA tutor's verbal language. Although the intervention design strived to encourage simple, direct language even as it incorporated the use of manipulatives and visuals, children with stronger language comprehension demonstrated a clear advantage. This suggests the need to incorporate additional non-language instructional supports in further strengthening this evidence-based intervention. It also suggests that educators must incorporate adjustments for children with very low language comprehension.

At the same time, executing the strategies we taught for deriving problem solutions is demanding of working memory capacity. For example, in using the counting up strategy to solve 9 minus 3, the child "puts the 3 in his or her head" (i.e., to remember to count up from 3) and then counts up to 9 holding up one finger for each count (e.g., "4, 5, 6, 7, 8, 9"). The answer is the number of fingers used during counting up. With double-digit addition and subtraction, children must utilize working memory for calculation within each place-value column as well as for regrouping concepts and notation. We provided practice designed to build children's accuracy and fluency with these and the other strategies we taught. Even so, children with stronger working memory capacity demonstrated a clear advantage. This suggests the need to incorporate additional instructional supports to compensate for working memory deficits.

Moreover, the calculation intervention exerted a strong main effect over the control condition (Powell et al., 2015). The present set of analyses extends this work by indicating that the main effect favoring calculation over control remains strong, even with a moderator effect and the main effects of seven other child-level abilities controlled in the model and with the factor score incorporating proximal and distal calculation outcomes.

The multi-level word-problem intervention proved even more robust. For the word-problem intervention, no moderator effects were identified. The main effect favoring word-problem intervention over the control group was strong throughout the distribution on all eight of the tested moderators, while the complex model also controlled the main effects of seven other child-level abilities or skills and when the factor score incorporated word-problem outcomes.

These analyses provide the basis for contrasting child-level variables that predict development in calculations versus word problems – *across* environmental, instructional, or intervention conditions. In investigating these main effect predictors, our analyses are more stringent than most prior work, because we controlled for the effects of a highly efficacious intervention in our models, even as we restricted the range of performance on the outcome by studying children with MD. Even so, attentive behavior emerged as a significant predictor of both outcomes. Additionally, initial calculation skill uniquely explained posttest

calculation performance, while language comprehension uniquely explained posttest word-problem skill.

This further substantiates the importance of these abilities in mathematics development, as has been shown in prior work for attentive behavior (e.g., Fuchs, Geary, et al., 2010), foundational calculation skill for later calculation performance (Fuchs et al., 2012), and language comprehension for word problems (e.g., Fuchs, Geary, et al., 2010). In this study, every unit increase in attentive behavior was associated with a 0.16 *SD* increase in later calculation skill and with a 0.33 *SD* increase in later word-problem solving. Every unit in initial calculation skill was associated with a 0.26 *SD* increase in the calculation outcome, and every unit in language comprehension led to a 0.20 *SD* increase in the word-problem outcome.

Implications for Practice

As just noted, results from this study indicated that teachers must be mindful of the possibility that children with extremely low working memory or language comprehension ability may require additional non-language instructional supports to optimize responsiveness to calculation intervention. At the same time, it is important to note that attentive behavior emerged as a significant overall (main effect) predictor even in the context of highly effective calculation or word-problem interventions that included evidence-based self-regulation strategies. This suggests the need to extend the typical kinds of strategies incorporated into evidence-based interventions, perhaps with regular feedback to children on their learning progress and by motivating at-risk learners to use that feedback to take additional responsibility for directing components of their school program, perhaps in terms of homework.

Study Limitations and Conclusions

The impact study (Fuchs, Powell, et al., 2014; Powell et al., 2015) on which the present analyses were conducted was limited in three important ways. First, research staff, who were not certified teachers, delivered the intervention. Research staff did receive careful training, support, and supervision to implement the interventions with fidelity. We would recommend that schools provide adequate training and support to building staff as staff deliver these or any evidence-based interventions. To the extent that this cannot be ensured, overall impact and child responsiveness are likely to decrease. At the same time, the interventions used in this study are fully specified in manuals, with complete sets of materials available for implementation. This provides the basis for wide distribution for accurate implementation and for conducting effectiveness studies, which are needed to quantify intervention effects when school personnel deliver intervention with the typical levels of support available in the schools.

Second, although the estimates of ES from the impact study were based on high quality randomized control trials, the moderator analyses reported in the present analyses were correlational. Therefore, conclusions about which cognitive processes moderate the effects of the calculation intervention should be interpreted with caution. Third, the targeted population was children with concurrent difficulty across calculations and word problems.

This population is important, given that these learners experience inferior outcomes than children with difficulty in just one domain and that the majority of children with MD experience difficulty across these domains. Even so, it would be potentially instructive to study the child-level variables associated with responsiveness for children with initial difficulty in calculations versus initial difficulty with word problems. Such a design may expand insights in important ways, and future research taking this approach is needed.

These study limitations notwithstanding, we draw three major conclusions. First, in both interventions (i.e., calculation intervention and word-problem intervention), effects were strong throughout the range of initial pretest scores among children with MD who experienced concurrent difficulty with calculations and word problems. Initial domain-specific performance did not moderate intervention effects for either intervention. That is, neither initial calculation skill nor initial understanding about number was associated with children's responsiveness to calculation intervention. Neither initial word-problem skill nor initial understanding about number was associated with responsiveness to word-problem intervention. This suggests strong efficacy for these interventions and indicates their added value within a multi-tier system.

Our second conclusion is that whether domain-general cognitive abilities are associated with responsiveness to mathematics intervention depends on the intervention's focus: calculations versus word problems. For children receiving a multi-tier calculation intervention, working memory and language comprehension were associated with responsiveness to intervention. By contrast, we identified no domain-general abilities that were active in influencing responsiveness to the word-problem intervention. Future research may therefore productively focus on specific sub-domain outcomes in the search for moderators of responsiveness to intervention in order to provide insights on strategies for expanding the range of interventions available in the schools.

Finally, we reiterate that attentive behavior was the only domain-specific or domain-general variable that predicted performance across the intervention and control groups on calculations as well as word problems. As already noted, this suggests the continued importance of focusing on attentive behavior among children with MD and identifying methods to enhance self-regulation in this population.

Acknowledgments

This research was supported by Awards Number R01HD053714, R24HD075443, and Core Grant #HD15052 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to Vanderbilt University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th. Author; Washington, DC: 1994.
- Barrett P. Structural equation modeling: Adjudging model fit. *Personality and Individual Differences*. 2007; 42:815–824. doi:10.1016/j.paid.2006.09.018.

- Bauer DJ, Curran PJ. Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*. 2005; 40:373–400. doi:10.1207/s15327906mbr4003_5. [PubMed: 26794689]
- Branum-Martin L, Fletcher JM, Stuebing KK. Classification and identification of reading and math disabilities: The special case of comorbidity. *Journal of Learning Disabilities*. 2012; 46:490–499. doi:10.1177/0022219412468767. [PubMed: 23232442]
- Cirino PT, Fuchs LS, Elias JT, Powell SR, Schumacher RF. Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities*. 2015; 48:156–175. doi:10.1177/0022219413494239. [PubMed: 23851137]
- Connolly, AJ. KeyMath-Revised. American Guidance; Circle Pines, MN: 1998.
- Cook BG, Cook SC. Unraveling evidence-based practices in special education. *The Journal of Special Education*. 2013; 47:71–82. doi:10.1177/0022466911420877.
- Cummins DD, Kintsch W, Reusser K, Weimer R. The role of understanding in solving word problems. *Cognitive Psychology*. 1988; 20:405–438. doi:10.1016/0010-0285(88)90011-4.
- Every Student Succeeds Act (ESSA) of 2015. :114–95. Pub. L.
- Fuchs D, Compton DL, Fuchs LS, Bryant J, Davis GN. Making “secondary intervention” work in a three-tier responsiveness-to-intervention mode: Findings from the first-grade longitudinal reading study of the National Research Center of Learning Disabilities. *Reading and Writing*. 2008; 21:413–436. doi:10.1007/s11145-007-9083-9.
- Fuchs LS, Compton DL, Fuchs D, Powell SR, Schumacher RF, Hamlett CL, Vukovic RK. Contributions of domain-general cognitive resources and difference forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology*. 2012; 48:1315–1326. doi:10.1037/a0027475. [PubMed: 22409764]
- Fuchs LS, Fuchs D, Compton DL, Powell SR, Seethaler PM, Capizzi AM, Fletcher JM. The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*. 2006; 98:29–43. doi:10.1037/0022-0663.98.1.29.
- Fuchs LS, Fuchs D, Craddock C, Hollenbeck KN, Hamlett CL, Schatschneider C. Effects of small-group tutoring with and without validated classroom instruction on at-risk students’ math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*. 2008; 100:491–509. doi:10.1037/0022-0663.100.3.491. [PubMed: 19122881]
- Fuchs LS, Fuchs D, Stuebing K, Fletcher JM, Hamlett CL, Lambert WE. Problem-solving and computation skills: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology*. 2008; 100:30–47. doi:10.1037/0022-0663.100.1.30. [PubMed: 20057912]
- Fuchs LS, Geary DC, Compton DL, Fuchs D, Hamlett CL, Bryant JV. The contributions of numerosity and domain-general abilities to school readiness. *Child Development*. 2010; 81:1520–1533. doi:10.1111/j.1467-8624.2010.01489.x. [PubMed: 20840238]
- Fuchs LS, Geary DC, Compton DL, Fuchs D, Schatschneider C, Hamlett CL, Changas P. Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology*. 2013; 105:58–77. doi:10.1037/a0030127. [PubMed: 24065865]
- Fuchs LS, Geary DC, Fuchs D, Compton DL, Hamlett CL. Sources of individual differences in emerging competence with numeration understanding versus multidigit calculation skill. *Journal of Educational Psychology*. 2014; 106:482–498. [PubMed: 25284885]
- Fuchs LS, Hamlett CL, Powell SR. Second-grade calculations battery. 2003 Available from L. S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.
- Fuchs LS, Malone A, Schumacher RF, Namkung JM, Hamlett CL, Jordan NC, Siegler RS, Gersten R, Changas P. Supported self-explaining during fraction intervention. *Journal of Educational Psychology*. 2016; 108:493–508. doi:10.1037/edu0000073.
- Fuchs LS, Powell SR, Cirino PT, Schumacher RF, Marrin S, Hamlett CL, Changas PC. Does calculation or word-problem instruction provide a stronger route to pre-algebraic knowledge? *Journal of Educational Psychology*. 2014; 106:990–1006. [PubMed: 25541565]
- Fuchs LS, Sterba SK, Fuchs D, Malone A. Does evidence-based fractions intervention address the needs of very low-performing students? *Journal of Research on Educational Effectiveness*. 2016; 9:662–677. doi:10.1080/19345747.2015.1123336.

- Fuchs LS, Zumeta RO, Schumacher RF, Powell SR, Seethaler PM, Hamlett CL, Fuchs D. The effects of schema-broadening instruction on second graders' word-problem performance and their ability to represent word problems with algebraic equations: A randomized control study. *The Elementary School Journal*. 2010; 110:440–463.
- Geary DC. Mathematics and learning disabilities. *Journal of Learning Disabilities*. 2004; 37:4–15. doi: 10.1177/00222194040370010201. [PubMed: 15493463]
- Geary DC. Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*. 2011; 47:1539–1552. doi:10.1037/a0025510. [PubMed: 21942667]
- Geary DC, Hoard MK, Byrd-Craven J, Nugent L, Numtee C. Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*. 2007; 78:1343–1359. doi:10.1111/j.1467-8624.2007.01069.x. [PubMed: 17650142]
- Geary DC, Widaman KF. Numerical cognition: On the convergence of componential and psychometric models. *Intelligence*. 1992; 16:47–80. doi:10.1016/0160-2896(92)90025-M.
- Groen G, Resnick L. Can pre-school children invent addition algorithms? *Journal of Educational Psychology*. 1977; 69:645–652. doi:10.1037//0022-0663.69.6.645.
- Hayes, AF. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Guilford; New York, NY: 2013.
- Hayes AF, Matthes J. Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*. 2009; 41:924–936. doi: 10.3758/BRM.41.3.924. [PubMed: 19587209]
- Hoover, HD., Hieronymous, AN., Dunbar, SB., Frisbie, DA. Iowa Test of Basic Skills, Form K. Riverside; Itasca, IL: 1993.
- Jordan NC, Hanich LB. Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities*. 2000; 33:567–578. [PubMed: 15495398]
- Kintsch W, Greeno JG. Understanding and solving word arithmetic problems. *Psychological Review*. 1985; 92:109–129. doi:10.1037/0033-295X.92.1.109. [PubMed: 3983303]
- Kline, RB. Principles and practice of structural equation modeling. 4th. Guilford Press; New York, NY: 2016.
- Lembke ES, Garman C, Deno SL, Stecker PM. One elementary school's implementation of response to intervention (RTI). *Reading and Writing Quarterly*. 2010; 26:361–373. doi: 10.1080/10573569.2010.500266.
- National Council of Teachers of Mathematics. Principles to action. Author; Reston, VA: 2014.
- Pickering, S., Gathercole, S. Working Memory Test Battery for Children. Psychological Corporation; London, England: 2001.
- Powell SR, Fuchs LS, Cirino PT, Fuchs D, Compton DL, Changas PE. Effects of a multi-level support system on calculation, word-problem, and pre-algebraic learning among at-risk learners. *Exceptional Children*. 2015; 81:443–470. doi:10.1177/0014402914563702. [PubMed: 26097244]
- Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*. 2008; 40:879–891. doi:10.3758/brm.40.3.879. [PubMed: 18697684]
- Psychological Corporation. Wechsler Abbreviated Scale of Intelligence (WASI) manual. Author; San Antonio, TX: 1999.
- Raghubar KP, Barnes MA, Hecht SA. Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*. 2010; 20:110–122. doi:10.1016/j.lindif.2009.10.005.
- Riley MS, Greeno JG. Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction*. 1988; 5:49–101. doi:10.1207/s1532690xci0501_2.
- Siegler RS, Booth JL. Development of numerical estimation in young children. *Child Development*. 2004; 75:428–444. doi:10.1111/j.1467-8624.2004.00684.x. [PubMed: 15056197]
- Siegler RS, Opfer JE. The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*. 2003; 14:237–243. [PubMed: 12741747]

- Swanson HL. Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*. 2011; 103:821–837. doi:10.1037/a0025114.
- Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, M., Sergeant, J., McCleary, R. Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN Ratings Scales. n.d.Retrieved August 19, 2007, from www.adhd.net
- Vukovic RK, Fuchs LS, Geary DC, Jordan NC, Gersten R, Siegler RS. Sources of individual differences in children's understanding of fractions. *Child Development*. 2014; 85:1461–1476. doi: 10.1111/cdev.12218. [PubMed: 24433246]
- Willcutt EG, Petrill SA, Wu S, Boada R, Defries JC, Olson RK, Pennington BF. Comorbidity between reading disability and math disability: concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities*. 2013; 46:500–516. doi: 10.1177/0022219413477476. [PubMed: 23449727]
- Wise JC, Pae HK, Wolfe CB, Sevcik R, Morris RD, Lovett M, Wolf M. Phonological awareness and rapid naming skills of children with reading disabilities and children with disabilities who are at risk for mathematics difficulties. *Learning Disabilities Research and Practice*. 2008; 23:125–136. doi:10.1111/j.1540-5826.2008.00270.x.
- Woodcock, RW. Woodcock Diagnostic Reading Battery. Riverside; Itasca, IL: 1997.
- Woodcock, RW., McGrew, KS., Mather, N. Woodcock-Johnson III tests of cognitive ability. Riverside; Itasca, IL: 2001.
- Zheng X, Swanson HL, Marcoulides GA. Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*. 2011; 110:481–498. doi:10.1016/j.jecp.2011.06.001. [PubMed: 21782198]

Figure 1.
Interaction between working memory-sentences and the effect of calculation intervention over the control group.

Figure 2.

Interaction between language comprehension and the effect of calculation intervention over the control group.

Table 1**Demographics and Mathematics Performance by Condition**

Variable	Condition					
	Calculation intervention		Word-problem intervention		Control	
	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)	%	<i>M</i> (<i>SD</i>)
Demographics						
Male	49		60		41	
Subsidized lunch	93		87		90	
Race: African American	55		55		58	
White non-Hispanic	11		16		19	
White Hispanic	30		22		17	
Other	4		7		6	
Special education	9		9		14	
English learner	18		17		16	
Pretest mathematics						
Number Line		17.97 (6.04)		19.25 (6.37)		18.62 (6.35)
Single-Digit Fluency		11.53 (5.08)		11.58 (5.10)		11.99 (5.17)
Story Problems		2.97 (1.55)		3.21 (1.58)		3.12 (1.59)
Posttest mathematics						
Single-Digit Fluency		44.54 (16.38)		29.84 (11.66)		21.97 (9.83)
Double-Digit Accuracy		21.92 (8.98)		13.04 (7.01)		10.03 (6.71)
KeyMath-Calculations		16.42 (4.03)		13.53 (3.04)		12.33 (3.53)
Second-Grade Word Problems		2.87 (1.83)		5.70 (3.21)		3.35 (1.91)
KeyMath-Problem Solving		2.26 (1.44)		2.40 (1.24)		2.28 (1.40)
ITBS-Data Interpretation & Problem Solving		13.48 (3.79)		13.17 (3.40)		12.90 (3.23)

Note. ITBS = Iowa Test of Basic Skills.

Table 2

Means, Standard Deviations (SDs), and Correlations

Calculation Intervention and Control Groups (<i>n</i> = 160)												
Variables	Raw Score		Standard Score		Correlations							
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	AS	A	P	R	W-S	W-C	L	N
Pretest Add and Subtract (AS)	11.73	(5.53)	NA									
Attentive Behavior (A)	31.45	(11.35)	NA		.21							
Processing Speed (P)	10.33	(2.28)	90.28	(11.99)	.14	.12						
Reasoning (R)	12.82	(5.49)	42.69	(7.05)	.25	.29	.20					
Working Memory-Sentences (W-S)	5.49	(3.58)	76.98	(16.73)	.12	.36	.18	.42				
Working Memory-Counting (W-C)	12.71	(3.91)	79.38	(13.66)	.11	.16	.07	.24	.26			
Language Comprehension (L)	14.76	(4.30)	82.02	(11.29)	.14	.16	.08	.07	.30	.11		
Number Line (N)	18.27	(6.33)	NA		.32	.18	.07	.10	.21	.12	.05	
Posttest Calculation Skill (C)	0.00	(1.00)	NA		.31	.40	.16	.25	.24	.21	.05	.22

Word-Problem Intervention and Control Groups (<i>n</i> = 159)												
Variables	Raw Score		Standard Score		Correlations							
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	SP	A	P	R	W-S	W-C	L	N
Pretest Story Problems (SP)	3.17	(1.56)	NA									
Attentive Behavior (A)	29.88	(10.33)	NA		.09							
Processing Speed (P)	10.25	(2.26)	89.49	(11.91)	-.03	.20						
Reasoning (R)	12.50	(4.73)	42.79	(6.98)	.23	.25	.19					
Working Memory-Sentences (W-S)	5.24	(3.66)	77.26	(19.29)	.22	.29	.13	.31				
Working Memory-Counting (W-C)	12.53	(3.69)	78.27	(12.41)	.15	.20	.04	.23	.18			
Language Comprehension (L)	15.33	(4.67)	83.92	(12.99)	.10	.04	.07	.11	.33	.05		
Number Line (N)	19.00	(6.51)	NA		.18	.08	.08	.04	.19	.10	.13	
Posttest Word Problems (WP)	0.00	(1.00)	NA		.18	.41	.16	.22	.27	.13	.27	.17

Note. AS = Pretest Add and Subtract on *Second-Grade Calculations Battery* (Fuchs et al., 2003); A = Attentive Behavior from Strengths and Weaknesses of ADHD-Symptoms and Normal Behavior (Swanson et al., 2004); P = Processing Speed on Visual Matching from Woodcock-Johnson III (Woodcock et al., 2001); R = Reasoning on Matrix Reasoning from Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999); W-S = Working memory-sentences is Listening Recall from the Working Memory Test Battery-Children (WMTB-C; Pickering & Gathercole, 2001); W-C = Working memory-counting is Counting Recall from the WMTB; L = Listening Comprehension from Woodcock Diagnostic Reading Battery (WDRB; Woodcock, 1999); N = Number Line Estimation (Siegler &

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Booth, 2004); C = Posttest Calculation is a factor score across single-digit addition and subtraction fluency (Fuchs et al., 2003), double-digit calculation skill (Fuchs et al., 2003), and KeyMath Addition and Subtraction (Connolly, 1998); NA = Not Applicable.

Table 3

Moderation Analyses

Moderators for Calculation Intervention and Control Groups (<i>n</i> = 160)															
Effects	Add and Subtract		Attentive Behavior		Processing Speed		Reasoning		Working Memory-Sentences		Working Memory-Counting		Language Comprehension		Number Line
	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	
Constant	−0.01	−0.01(.990)	−0.01	−0.22(.826)	−0.00	−0.03(.978)	−0.00	−0.07(.947)	0.00	0.02(.986)	−0.00	−0.01(.966)	0.00	0.06(.950)	0.00 0.01(.991)
Moderator	0.26	4.44(<001)	0.15	2.59(.010)	0.04	0.76(.450)	0.03	0.43(.665)	0.11	1.81(.073)	0.10	1.81(.072)	0.01	1.18(.239)	0.04 0.77(.441)
Intervention	1.24	11.54(<001)	1.24	11.63(<001)	1.24	11.55(<001)	1.24	11.58(<001)	1.24	11.70(<001)	1.24	11.52(<001)	1.24	11.72(<001)	1.24 11.53(<001)
Moderator × Intervention	−0.03	−0.31(.761)	0.12	1.12(.266)	0.06	0.55(.586)	0.12	1.12(.245)	0.22	2.15(.033)	0.03	0.24(.813)	0.22	2.11(.036)	0.02 0.21(.836)
Control variables															
Add and Subtract	NA		0.26	4.52(<.001)	0.26	4.42(<.001)	0.26	4.49(<.001)	0.26	4.52(<.001)	0.26	4.43(<.001)	0.26	4.56(<.001)	0.26 4.44(<.001)
Attentive Behavior	0.15	2.61(.010)	NA		0.16	2.70(.008)	0.16	2.75(.007)	0.16	2.85(.005)	0.16	2.67(.008)	0.16	2.72(.007)	0.15 2.64(.009)
Processing Speed	0.05	0.88(.381)	0.05	0.91(.364)	NA		0.05	0.90(.369)	0.05	0.95(.346)	0.05	0.85(.396)	0.04	0.79(.431)	0.05 0.86(.389)
Reasoning	0.03	0.42(.672)	0.02	0.39(.694)	0.03	0.46(.644)	NA		0.01	0.15(.882)	0.03	0.46(.647)	0.01	0.11(.911)	0.03 0.46(.643)
Working Memory-Sentences	0.09	1.42(.157)	0.10	1.52(.131)	0.09	1.45(.148)	0.09	1.39(.167)	NA		0.09	1.41(.161)	0.11	1.70(.091)	0.09 1.41(.161)
Working Memory-Counting	0.10	1.80(.073)	0.11	1.97(.051)	0.10	1.78(.077)	0.11	1.98(.049)	0.10	1.84(.064)	NA		0.95	1.78(.078)	0.10 1.81(.072)
Language Comprehension	0.04	0.67(.504)	0.04	0.70(.484)	0.03	0.63(.532)	0.03	0.55(.582)	0.04	0.87(.386)	0.04	0.65(.515)	NA		0.04 0.64(.525)
Number Line	0.04	0.81(.420)	0.05	0.83(.500)	0.04	0.78(.436)	0.03	0.61(.544)	0.04	0.77(.442)	0.04	0.80(.428)	0.05	0.95(.342)	NA
Moderators for Word-Problem Intervention and Control Groups (<i>n</i> = 159)															
Effects	Story Problems		Attentive Behavior		Processing Speed		Reasoning		Working Memory-Sentences		Working Memory-Counting		Language Comprehension		Number Line
	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	
Constant	0.00	0.03(.978)	−0.00	−0.06(.952)	−0.00	−0.01(.989)	0.00	0.00(.999)	0.00	0.00(.999)	0.00	0.02(.988)	−0.00	−0.01(.994)	−0.00 −0.04(.965)
Moderator	0.07	1.03(.303)	0.34	4.53(<.001)	0.06	0.80(.422)	0.07	0.97(.331)	0.02	0.25(.800)	0.02	0.27(.787)	0.20	2.78(.006)	0.11 1.60(.111)
Intervention	0.41	3.00(.003)	0.40	3.03(.003)	0.42	3.03(.003)	0.42	3.03(.003)	0.42	3.02(.003)	0.42	3.01(.003)	0.42	3.02(.002)	0.42 3.06(.003)
Moderator × Intervention	−0.13	−0.90(.371)	0.10	0.72(.470)	0.11	0.78(.436)	0.08	0.54(.588)	−0.00	−0.02(.982)	0.10	0.72(.475)	0.01	0.10(.924)	0.14 0.98(.328)
Control variables															
Story Problems	NA		0.07	0.91(.364)	0.08	1.04(.302)	0.07	0.96(.338)	0.08	1.04(.302)	0.07	1.01(.316)	0.07	1.01(.315)	0.07 0.98(.327)

Moderators for Calculation Intervention and Control Groups (<i>n</i> = 160)																
Effects	Add and Subtract		Attentive Behavior		Processing Speed		Reasoning		Working Memory-Sentences		Working Memory-Counting		Language Comprehension		Number Line	
	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)	Coeff	<i>t</i> (<i>p</i>)
Attentive Behavior	0.34	4.54(<.001)		NA	0.33	4.41(<.001)	0.34	4.51(<.001)	0.33	4.41(<.001)	0.34	4.54(<.001)	0.33	4.40(<.001)	0.32	4.27(<.001)
Processing Speed	0.06	0.88(.382)	0.06	0.90(.370)		NA	0.07	0.96(.337)	0.05	0.95(.346)	0.06	0.89(.374)	0.07	0.94(.349)	0.07	0.96(.338)
Reasoning	0.08	1.07(.287)	0.08	1.02(.310)	0.08	1.03(.303)		NA	0.08	0.99(.323)	0.08	1.06(.289)	0.05	0.69(.494)	0.08	1.07(.286)
Working Memory-Sentences	0.03	0.27(.784)	0.03	0.35(.726)	0.03	0.36(.722)	0.03	0.31(.760)		NA	0.03	0.29(.770)	0.07	0.96(.339)	0.02	0.26(.794)
Working Memory-Counting	0.01	0.17(.867)	0.02	0.31(.753)	0.01	0.19(.853)	0.02	0.30(.763)	0.02	0.21(.836)		NA	0.03	0.29(.791)	0.02	0.21(.834)
Language Comprehension	0.21	2.91(.004)	0.21	2.87(.005)	0.20	2.75(.007)	0.20	2.69(.008)	0.20	2.78(.006)	0.21	2.82(.005)		NA	0.29	2.78(.006)
Number Line	0.11	1.53(.128)	0.12	1.69(.093)	0.11	1.60(.112)	0.11	1.60(.112)	0.12	1.61(.109)	0.12	1.62(.108)	0.12	1.61(.109)		NA

Note. Coeff = Coefficient; NA = Not Applicable.

Note. Coeff = Coefficient; NA= Not Applicable.