Contemporary Educational Psychology 60 (2020) 101811

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Do the processes engaged during mathematical word-problem solving differ along the distribution of word-problem competence?

Lynn S. Fuchs[a,*], Sarah R. Powell[b], Anna-Mária Fall[b], Greg Roberts[b], Paul Cirino[c], Douglas Fuchs[a], Jennifer K. Gilbert[a]

[a] Vanderbilt University, United States
[b] University of Texas at Austin, United States
[c] University of Houston, United States

ARTICLE INFO

Keywords:
Mathematics
Word problems
Quantile analysis
Language comprehension
Working memory
Calculations

ABSTRACT

Conventional research methods for understanding sources of individual differences in word-problem solving (WPS) only permit estimation of average relations between component processes and outcomes. The purpose of the present study was instead to examine whether and if so how the component processes engaged in WPS differ along the spectrum of WPS performance. Second graders ($N = 1130$) from 126 classrooms in 17 schools were assessed on component processes (reasoning, in-class attentive behavior, working memory, language comprehension, calculation fluency, word reading) and WPS. Multilevel, unconditional quantile multiple regression indicated that 3 component processes, calculation fluency, language comprehension, and working memory, are engaged in WPS differentially depending on students' overall word-problem skill. The role of calculation fluency and language comprehension was stronger with more competent word-problem solving ability. By contrast, the role of working memory was stronger with intermediate-level than for strong problem solving. Results deepen insight into the role of these processes in WPS and provide the basis for hypothesizing how instructional strategies may be differentiated depending on students' overall level of WPS competence.

1. Introduction

Word-problem solving (WPS), which represents a major emphasis in almost every strand of the mathematics curriculum, is also the best school-age predictor of employment and wages in adulthood (Bynner, 1997; Every Child a Chance Trust, 2009). Yet, WPs are a stumbling block for many students, and WPS presenting challenges even when calculation skill is adequate (Fuchs et al., 2008; Swanson, Jerman, & Zheng, 2008). Specific WP difficulty may occur because the cognitive processes involved in WPs differ from and are more numerous than those underlying calculation skill (e.g., Fuchs, Zumeta, et al., 2010; Fuchs et al., 2006, 2008). Thus, WP difficulty may be determined by multiple factors and may be difficult to prevent or improve once students have fallen behind peers.

For these reasons, it is important to understand the foundational academic and cognitive processes underlying individual differences in performance. In this paper, we refer to these processes as *component processes*. Such understanding may help guide effective methods to screen students for early intervention and shed light on potentially productive methods for teaching WPs and reducing performance gaps that have been established.

Toward this end, a growing body of research has identified component processes that foster WPS (e.g., Anderson, 2007; Bernardo, 1999; Fuchs et al., 2006, 2010a, 2010b, 2016; Lee, Ng, Ng, & Lim, 2004; Peng, Namkung, Barnes, & Sun, in press; Raghubar, Barnes, & Hecht, 2010; Swanson, 2016; Swanson & Beebe-Frankenberger, 2004; Van der Schoot, Bakker Arkema, Horsley, & Van Lieshout, 2009). Yet, conventional analyses used to understand sources of individual differences in WPS outcomes (e.g., standard multiple regression, structural equation modeling) only permit estimation of average relations between component processes and outcomes (Petscher & Logan, 2014). And as demonstrated by Purpura and Logan (2015) when they focused on approximate number system outcomes, conventional regression methods may fail to capture nuances in the component processes that depend on the sophistication of students' WPS performance. Although some studies investigate nonlinear relations using statistical interactions within conventional regression analysis, as Bonny and Lourenco (2013) did when investigating approximate number system outcomes, we identified none focused on WPS outcomes. Moreover, conventional regression methods for exploring the nature of interactions involve

* Corresponding author at: 228 Peabody, Vanderbilt University, Nashville, TN 37203, United States.
E-mail address: lynn.fuchs@vanderbilt.edu (L.S. Fuchs).

https://doi.org/10.1016/j.cedpsych.2019.101811

Available online 25 October 2019
0361-476X/

forming subgroups of students, whose data are fit to a slope that best conforms to individuals in that subgroup, to derive mean relations for that subgroup.

In the present study, we extended this literature on the foundational academic and cognitive processes underlying individual differences in WP performance with quantile regression analysis, which allows testing this question without forming subgroups and without constraining the functional form of relations between variables. Even without dividing the sample into decile subgroups, quantile analysis results may provide the basis for formulating hypotheses on instructional differentiation strategies that map more authentically to the small-group or center-stationed instruction schools provide on the basis of low-, intermediate-, and high-level performance on the relevant domain.

Our research question was, How do the processes engaged in WPS differ along the spectrum of the WP outcome? Our focus included six component processes: four foundational cognitive processes (in-class attentive behavior, working memory, reasoning, language comprehension) and two foundational academic skills (calculation fluency, word reading). We focused on second grade, after the first-grade burst of WPS skill has produced a broad range of individual differences (Fuchs et al., 2013).

We begin this introduction by the WPS model described by Kintsch and colleagues' (Cummins, Kintsch, Reusser, & Weimer, 1988; Kintsch & Greeno, 1985; Nathan, Kintsch, & Young, 1992). This provides the theoretical framework and frames the component processes targeted for the present investigation. After describing the text processing that occurs during WPS in ways that make engagement of foundational cognitive processes and foundational academic skills transparent, we document that each of these processes has been established via conventional analyses as a source of individual WPS differences. We then explain how the present study extends the literature.

2. Word-problem solving model

To target the component processes for investigation, we relied on Kintsch et al.'s model of WPS (e.g., Cummins et al., 1988; Kintsch & Greeno, 1985; Nathan et al., 1992), which suggests WPS is a complex undertaking. The model first assumes that general features of the text comprehension process apply across stories, essays, and WP statements, but that the comprehension strategies, the nature of the required knowledge structures, and the form of resulting macrostructures and situation and problem models differ by text type.

Based on theories of text comprehension and discourse processing (e.g., van Dijk & Kintsch, 1983), the model further posits that WP representations have three components. The first involves constructing a coherent microstructure and deriving a hierarchical macrostructure to capture the text's essential ideas. Second, the situation model supplements the text using inferences based on the problem solver's world knowledge, which includes relationships between quantities.

The problem solver uses this information to identify the third component, the problem model or schema, in which the structural relations among the quantities are formalized. This schema drives the problem solver's solution strategies. At second grade, three schemas are most common: *combine* WPs (two or more parts are combined to create a total), *compare* WPs (two quantities are compared), and *change* WPs (an event changes a starting amount to create a new ending amount). Missing information may occur in any slot of the number sentence representing the schema (e.g., in the example of the combine schema below, the total or one of the parts may be the unknown quantity).

The Kintsch model specified that building the propositional text structure, inferencing, schema induction, and applying solution strategies makes strong demands on short-term memory and language comprehension. In subsequent work (e.g., Fuchs, Fuchs, Seethaler, & Barnes, 2019), short-term memory has been reframed as working memory, because WPS requires not only briefly storing information but also sequentially updating that information in memory as the problem solver processes segments of the WP statement. This reframing is grounded in studies showing that working memory is engaged in WPS (Anderson, 2007; Lee et al., 2004; Peng et al., in press; Raghubar et al., 2010; Swanson & Beebe-Frankenberger, 2004; Swanson et al., 2008; Swanson & Sachse-Lee, 2001; Swanson, 2016; Swanson, Moran, Lussier, & Fung, 2014).

Additionally, in light of related correlational research demonstrating the role of in-class attentive behavior and reasoning in WPS (Fuchs et al., 2006, 2010a, 2010b; Swanson & Beebe-Frankenberger, 2004), the model has been expanded to include these cognitive processes. In-class attentive behavior reflects students' capacity to engage throughout the problem-solving process. Via careful attending and access to information stored in working memory, students engage in reasoning to integrate newly encountered information with information stored in working memory to logically induce relations between objects and actions described in the WP narrative and distinguish between relevant and irrelevant information.

We clarify the role of in-class attentive behavior, working memory, and reasoning ability by explaining the problem-solving process in which a competent problem solver might solve the following combine problem (Part 1 plus Part 2 equals Total or P1 + P2 = T): *Joe has 3 baseballs. Tom has 2 footballs and 2 baseballs. How many baseballs do the boys have in all?* The problem solver processes the first sentence's propositional text base to identify that the object is baseballs, the quantity is 3, the actor is Joe, but Joe's role is to be determined. This information is held in working memory. In the second sentence, propositions are similarly coded and placed in working memory, but the word *footballs* fails to match the object code in the prior sentence, signaling the number 2 footballs as perhaps irrelevant; this is added to working memory. In the last sentence (the problem's question), the quantitative proposition *how many baseballs* and the phrase *in all* cues the problem solver to identify the combine schema as the problem type that links the information saved within working memory; assign the role of superset (Total) to the question; assign subset roles (Parts 1 and Part 2) to the to-be-determined quantities in working memory; and reject 2 footballs.

Filling in these slots of the schema in this way triggers the specification of a number sentence with a missing quantity (i.e., 3 + 2 = __); then the problem solver calculates the solution. With typical school instruction, children gradually induce the combine schema as a "problem type" on their own (this is rarely explicitly taught), just as they devise their own strategies for managing the attentional, working memory, and reasoning demands posed by this problem-solving sequence.

The fourth cognitive resource involved in WPS is language comprehension. According to Kintsch and Greeno (l985), children enter school understanding important vocabulary and language constructions; with arithmetic and WP instruction, they learn to treat these words in a math-specific way (e.g., *more* becomes the more complicated construction *more than* involving sets) via mathematics instruction, which includes classroom discussions designed to sustain inquiry-based discussion and argumentation (Yackel & Cobb, 1996). In a computational simulation, Cummins et al. (1988) determined that WP representation depended heavily on language comprehension and that altering wording in minor ways dramatically affected solution accuracy.

To illustrate how WPS depends on and taxes language comprehension, consider this problem: *Joe has 3 cows. Tom has 2 barns and 3 goats. How many animals do the farmers have in all?* Compared to the first combine problem, which presented similar demands for inducing the schema, the vocabulary and constructions involving this scenario's objects increase demands on language comprehension for assigning roles in the propositional text structure. Increased language comprehension demands arise from more sophisticated representations of vocabulary involving taxonomic relations at superordinate levels and subtle distinctions among categories (cows and goats are animals; barns are not animals).

This model of WPS thus identifies in-class attentive behavior, working memory, reasoning, and language comprehension as key cognitive processes engaged in WPS. Additionally, computational skill is transparently required to succeed with the problem-solving task, and word-reading skill should facilitate access to the text (even when problems are read aloud while presented to students, as done in the present study). We targeted these six component processes in the present investigation.

3. Prior research and how the present study extends that literature

Prior research, which has focused exclusively on the estimation of average relations between component processes and WP outcomes, provides evidence for a unique role for in-class attentive behavior (e.g., Fuchs et al., 2006, 2010), working memory (e.g., Geary & Widaman, 1991; Swanson & Beebe-Frankenberger, 2004; Swanson et al., 2008; Swanson & Sachse-Lee, 2001), reasoning (e.g., Fuchs et al., 2012), and language comprehension (e.g., Fuchs et al., 2006, 2010; Swanson, 2006). The contribution of calculation and word-reading skill has also been established (e.g., Fuchs et al., 2006, 2007; Swanson, 2006).

What extends the present study beyond earlier work is that instead of using conventional multiple regression analysis, we employed *quantile* multiple regression to examine the engagement of each component process as a function of WPS ability, while controlling for the effects of the other five component processes. Quantile regression is well suited for answering questions about differential effects along the distribution of an outcome skill, when the interest is subgroups of learners. Purpura and Logan (2015), who assessed the effects of the approximate number system and mathematical language on preschool mathematics knowledge, contrasted the use of conventional regression against quantile regression. They illustrated that the mean effects revealed via conventional regression analysis failed to detect the more nuanced pattern of effects detected with quantile regression.

In the present study, the interest was deepening insight on the component processes engaged at varying levels of WPS sophistication. The hope is that findings provide direction for hypothesizing and then testing potential strategies for differentiating instruction for low-level problem solvers versus intermediate-level problem solvers versus high-level problem solvers. This parallels schools' reliance on special education or small-group Tier 2 intervention within responsiveness-to-intervention (RTI) or a multi-tier system of supports (MTSS) to meet the needs of students with identified or at risk for mathematics learning disabilities. In many schools, RTI or MTSS intervention occurs during a school-wide block, where students are grouped by low-, intermediate-, and high-level levels of academic performance to differentiate among the needs of these groups (Fuchs & Fuchs, 2016).

Quantile regression answers questions about differential effects along the distribution of an outcome skill in subgroups of learners in a way that maximizes information in the analytic sample (Koenker & Bassett, 1978). Rather than categorizing students into ability subgroups and then comparing those groups, it compares students at different points along the distribution using all available data to make the estimation. Quantiles are the cut-points dividing the range of observations into equal, continuous intervals; in this study 10 quantiles (or deciles). Data points closer to the quantile of interest are weighted more than data points farther away. The quantile regression parameter estimates the change in a specified quantile on the outcome variables (in this study, WPS) by a one unit change in each predictor variable (in this study, each component process). This allows comparing quantiles (percentiles) on WPS may be more or less affected by each predictor variable. (For additional information on quantile analysis, see Koenker, 2015.)

It is important to note that in the present study, to avoid results that fail to generalize across quantiles, as is the case with conditional quantile regression, we relied on *un*conditional quantile regression models. In conditional quantile regression, effects are difficult to interpret when they vary between conditional quantiles because the weights or coefficients do not "aggregate" to the population quantile values. Unconditional quantile regression instead permits regression logic in terms of quantiles. Our analyses also accounted for the nested structure of the database to control for the effects of schools, teachers, and classrooms.

4. Method

4.1. Participants

Data collection for the original intervention study (Fuchs et al., 2014) was approved by the university's Institutional Review Board. Teachers and parents provided written consent; students provided verbal assent. Consent rates were 100% for teachers, 78% for parents, and 98% for students. In the original study, we selected participants from 1917 children with parent or guardian consent and child assent in 127 second-grade classrooms taught by 96 teachers in 18 schools in a metropolitan school district across four cohorts (one per year for 4 years; some teachers participated in more than one cohort, with some teachers having more than one classroom in the study, hence the discrepancy between number of classrooms and teachers).

Selection of the sample occurred in three steps. First, we conducted whole-class screening with these 1917 children on WP and calculation measures. The WP measure was the outcome measure used in the present study's quantile analysis (see Measures). The calculation measure was one subtest *(Sums to 12)* of the calculation fluency component process used in the present study (see Measures); in this sample, Cronbach's alpha (α) for this one subtest was 0.85–0.93. Addressing both domains in the context of the present analysis to identify the sample is helpful because it promotes generalizability to students who are identified for mathematics intervention in schools, which have a heavy focus on calculations in their screening for math intervention.

In the second screening step, we randomly sampled from these strata to achieve a sample that oversampled at-risk students but represented the entire distribution, as called for in the original study design. Third, we excluded children scoring below the 9th percentile on both of two intelligence subtests of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999); this was to ensure students had the requisite abilities to participate in a meaningful way. Note that some children completed none or only a small portion of pretesting because they moved early in the school year or became unavailable due to school scheduling changes.

The present analysis hence incorporated children who had complete data on the variables of interest, which were collected at the start of second grade, with 1130 children from 126 classrooms (95 teachers) in 17 schools. (Because data were collected before intervention began, intervention condition is not relevant in this report.) The sample was 51% female; 14% non-native English speaking; 84% receiving subsidized school lunch; and 42% African-American, 24% white Hispanic, 27% white non-Hispanic, 3% Kurdish, and 4% other. Students were 8.72 years on average in the second month of second grade, with a mean IQ of 93.60 (*SD* = 13.22) as measured on the WASI (Wechsler, 1999); a mean reading standard score of 100.14 (*SD* = 15.79) on the Wide Range Achievement Test (WRAT; Wilkinson, 1993); and a mean WRAT-Arithmetic standard score of 93.88 (*SD* = 12.79). Average standard scores on other norm-referenced tests employed in this analysis were each within 1 *SD* of the norming sample's mean.

4.2. Measures

All reliability estimates refer to α unless otherwise noted.

4.2.1. Descriptive measures

WASI (Wechsler, 1999) was used to index general cognitive ability (split-half reliability = 0.92). It includes the Vocabulary and Matrix Reasoning subtests. The *Wide Range Achievement Test* (WRAT-3; Wilkinson, 1993) was used to index mathematics performance. It requires children to write answers to calculation problems of increasing difficulty. In this sample, α was 0.93.

4.2.2. Quantile analysis outcome

Following Jordan and Hanich (2000), *Word Problems* comprises 14 brief word problems involving change, combine, compare, and equalize relationships and requiring single-digit addition or subtraction for solution (i.e., sums of 7, 8, or 9 or subtrahends of 6, 7, 8, or 9; there are no addends or minuends of zero or one; answers to the subtraction problems are from 2 to 6). The tester reads each item aloud; students have 30 sec to respond and can ask for re-reading(s) as needed. The score is the number of correct answers. In this sample, α was 0.89.

4.2.3. Quantile analysis component processes.

In-class attentive behavior was indexed using the *SWAN*, an 18-item teacher rating scale (Swanson et al., 2012) sampling items from the *Diagnostic and Statistical Manual of Mental Disorders-IV* (APA, l994) criteria for Attention-Deficit/Hyperactivity Disorder for inattention (items 1–9) and hyperactivity/impulsivity (items 10–18). Validity is supported in the literature (Arnett et al., 2013; Lakes, Swanson, & Riggs, 2012; Swanson et al., 2012). Teachers rated each item as 1 = Far Below, 2 = Below, 3 = Slightly Below; 4 = Average, 5 = Slightly Above, 6 = Above, 7 = Far Above. We report data for the in-class attentive behavior subscale as the average rating across the nine items that index the ability to maintain focus of attention. This score correlates well with other dimensional assessments of behavior related to attention (Swanson et al.). In this sample, α was 0.99.

To index **calculation fluency**, we administered two subtests of single-digit addition and two subtests of single-digit subtraction from the *Calculations Battery* (Fuchs, Hamlett, & Powell, 2003): *Sums to 12, Sums to 18, Minuends to 12, and Minuends to 18*. For each, students have 1 min to complete 25 problems. We combined the subtests to create one score. On a sub-sample of 79 students, test-retest reliability was 0.91.

To measure **language comprehension**, we used the *Listening Comprehension* subtest of the *Woodcock Diagnostic Reading Battery* (Woodcock, 1997), which measures the ability to understand sentences or passages. With 38 items, students supply the word missing at the end of sentences or passages that progress from simple verbal analogies and associations to discerning implications. Split-half reliability is 0.80 at ages 5–18.

For **reasoning**, we relied on the *Wechsler Abbreviated Scale of Intelligence – Matrix Reasoning* (Wechsler, 1999), which includes items of pattern completion, classification, analogy, and serial reasoning. We opted for a nonverbal measure of reasoning to avoid confounding verbal reasoning with language ability. For each of the 32 items, children select one of five options that best completes a visual pattern. As reported in the test manual, test-retest reliability for children ages 6–11 years is 0.76. The standard error of measurement reported for this instrument ranges from 2–4 for 6 to 11-year-old children.

To index **working memory**, we used the dual-task central executive *Listening Recall* subtest from the *Working Memory Test Battery for Children* (WMTB-C; Pickering & Gathercole, 2001), with which the child determines if each sentence in a series is true and then recalls the last word in each sentence. It has six items at span levels from 1–6 to 1–9. Passing four items at a level moves the child to the next level. At each span level, the number of items to be remembered increases by one. Failing three items at a given span terminates the subtest. We used the trials correct score. Test-retest reliability on 82 students was 0.84.

We indexed **word-reading skill** with the Wide Range Achievement Test (WRAT-3; Wilkinson, 1993), with which students read a list of words in increasing difficulty. In this sample, α was 0.98.

4.3. Procedure

Testers were trained to criterion and used standard directions for administration. All testers were graduate students pursuing a degree in an education-related field or full-time research coordinators. Testers participated in a 1-day training in which they learned how to administer all tests. Before administration of the individual assessments, each tester administered the test battery to one of the full-time research coordinators. The testers conducted assessments in the second and third months of second grade: the calculations and word-problem test in large groups; the other measures individually. For 15% of group-administered tests, paper protocols were double-scored to index agreement, which exceeded 99%. All individual sessions were audiotaped; 15% of tapes, stratified by tester, were re-scored by an independent scorer. Scoring accuracy exceeded 99%.

4.4. Data analysis

We used unconditional quantile regression to evaluate the extent to which the component processes involved in WPS for average students are engaged similarly for students at different levels of the WPS measure. As noted in the introduction, ordinary least squares regression consistently estimates the impact of an independent variable, X, on the *unconditional mean* of an outcome variable, Y. Quantile regression (QR) "goes beyond the mean," allowing model estimates at points along Y other than (or in addition to) its mean, by computing a distribution's quantiles.

However, whereas in regular OLS, conditional means average to the relevant population mean (i.e., the unconditional mean), quantiles in QR do not average to their unconditional population analogue. Therefore, QR estimates generally do not represent the impact of X on the conditional quantiles that comprise the distribution of Y. Although the QR framework provides a pragmatic approach to understanding the differential impacts of covariates along the distribution of an outcome, the framework that pervades the applied literature relies on a *conditional* QR method. By assessing the impact of a covariate on a quantile of the outcome conditional on specific values of other covariates, conditional QR may generate results that do not generalize across all quantiles, to the extent that conditional quantiles vary.

Unconditional QR (UQR), by contrast, yields interpretable results because it marginalizes (i.e., estimates values in the subset of variables without reference to values of the other variables) the effect over the distributions of other covariates in the model. This is accomplished by defining quantiles prior to fitting regressions (Killewald & Bearak, 2014; Porter, 2015). In the present analysis, we fit UCR models (Firpo, Fortin, & Lemieux, 2009) to understand the relation between component processes (reasoning, in-class attentive behavior, working memory, language comprehension, calculation fluency, and word reading) and WPS across the unconditional quantiles representing the WPS distribution.

We relied on Firpo el al.'s (2009) two-step approach, (a) calculating a re-centered influence function (RIF) for each quantile of interest (here, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90) and (b) conducting separate regressions for each quantile with the RIFs as dependent variables. This two-step process allows for estimates of the independent variable's effect and, more importantly, how that effect differs across the distribution of the dependent variable (Porter, 2015).

We followed the UQR models with three comparisons for each predictor (i.e., component process): the conditional coefficients at the 0.20 versus 0.50 quantiles, at the 0.20 versus 0.80 quantiles, and at the 0.50 versus 0.80 quantiles. These percentile values were selected to represent low, average, and high performance, which corresponds to the basis on which students are often grouped for instructional differentiation in schools. To ease interpretation, all variables were z-transformed prior to analysis.

In all models, we accounted for variance associated with the nested

Table 1
Means, Standard Deviations (SDs), and correlations.

			Correlations				
	Mean	SD	AB	CF	LC	R	WM
In-class attentive behavior (AB)	38.25	(12.16)					
Calculation fluency (CF)	21.93	(11.45)	0.38				
Language Comprehension (LC)	16.79	(4.76)	0.25	0.21			
Reasoning (R)	11.87	(5.89)	0.29	0.29	0.24		
Working memory (WM)	7.13	(3.59)	0.32	0.22	0.42	0.33	
Word-reading skill (WR)	26.26	(5.20)	0.45	0.34	0.31	0.30	0.35

Note. In-class attentive behavior is Strengths and Weaknesses of ADHD Symptoms and Normal Behavior rating scale (Swanson et al., 2012). Calculation fluency is Second-Grade Calculations Battery: Sums to 12, Sums to 18, Minuends to 12, Minuends to 18 (Fuchs et al., 2003). Language comprehension is Woodcock Diagnostic Reading Battery - Listening Comprehension (Woodcock, 1997). Reasoning is Wechsler Abbreviated Scale of Intelligence - Matrix Reasoning (Wechsler, 1999). Working memory is Working Memory Test Battery for Children - Listening Recall (Pickering & Gathercole, 2001). Word reading is Wide Range Achievement Test – Reading (Wilkinson, 1993).

data structure. Students were nested in classes/teachers (i.e., subsets of students shared the same class/teacher); classes/teachers were nested in schools (i.e., subsets of classes/teachers shared the same school); and classes were crossed with teachers (i.e., the 126 classes were led by 95 teachers, some of whom participated in multiple years of the cohort study, meaning that multiple classes shared – or were crossed with – the same teacher). We fit an unconditional multilevel model with WPS as the outcome, random effects at level 3 (school) and level 2 (teachers and classes). Because variance at the teacher level was 0, the model was rerun with the level 2 random effect removed. For schools, the ICC was 0.10; for classes, 0.14. We modeled variance at the school and class levels as multilevel mixed-effects regression in STATA V. 15, using the *rifreg* command to create RIFs and the *mixed* command to fit multilevel regressions.

5. Results

We report raw score means, SDs, and correlations in Table 1 and report *UQR* parameters in Table 2. Descriptively, the smallest to largest coefficient, averaged across the nine quantiles, was word reading (0.07), reasoning (0.16), working memory (0.18), in-class attentive behavior (0.19), language comprehension (0.23), and calculation fluency (0.30). We interpret one coefficient here as a guide for interpreting other coefficients: In the 0.10 quantile model, a 1 *SD* increase in calculation fluency was associated with a 0.32 *SD* increase in WPS, controlling for the effect of other component processes.

Controlling for the effect of other predictors, calculation fluency and reasoning were significantly related to WPS across the distribution, but the strength of these associations differed across some quantiles. Controlling for the effects of other predictors, in-class attentive behavior and language comprehension were significantly related to WPS at every quantile except 0.10; working memory at every quantile except the 0.90; and word-reading skill only at 0.10, 0.20, and 0.80.

Fig. 1 presents a visualization of the UQR results. The *x*-axes represent deciles, and the *y*-axes are the UQR coefficients, with the graphs illustrating the association between each component process and WPS as a function of decile. The figure shows, for example, that the unique predictive utility of language comprehension for WPS tends to increase from the 0.10 (coefficient = 0.08) to the 0.80 (coefficient = 0.37) quantile, whereas the importance of word reading for word-problem solving trends lower as the quantile increases.

To evaluate whether the utility of each predictor component process reliably differs at key points in the spectrum of WPS ability, we contrasted conditional coefficients for the same component process at low (0.20), intermediate (0.50), and high (0.80) quantiles. Table 3 shows results. The relation between language comprehension and WPS differed significantly for each contrast: It was significantly stronger at the intermediate versus the low quantiles, at the high versus the low quantiles, and at the higher versus the intermediate quantiles. The relation between calculation fluency with WPS was stronger at the high than at the intermediate quantile. In reverse direction, however, the relation between working memory and WPS was stronger at the intermediate than at the high quantile.

6. Discussion

The Kintsch model (Cummins et al., 1988; Kintsch & Greeno, 1985; Nathan et al., 1992) suggests WPS is a complex undertaking, which requires building the propositional text structure, formulating inferences based on world knowledge, deriving the problem model or schema, and applying solution strategies. As Kintsch et al. originally hypothesized and with later extensions to that model (e.g., Fuchs, Gilbert, Fuchs, Seethaler, & Martin, 2018), this undertaking makes strong demands on in-class attentive behavior, working memory, reasoning ability, and language comprehension.

Table 2
Unconditional quantile regression results.

		Quantiles of word-problem solving								
		0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
In-class attentive behavior	Estimate	0.08	**0.18**	**0.15**	**0.15**	**0.20**	**0.26**	**0.31**	**0.19**	**0.18**
	SE	0.05	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05
Calculation fluency	Estimate	**0.32**	**0.28**	**0.27**	**0.26**	**0.27**	**0.33**	**0.38**	**0.37**	**0.18**
	SE	0.05	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05
Language comprehension	Estimate	0.09	**0.15**	**0.17**	**0.16**	**0.24**	**0.29**	**0.33**	**0.37**	**0.28**
	SE	0.05	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05
Reasoning	Estimate	**0.11**	**0.14**	**0.21**	**0.20**	**0.13**	**0.15**	**0.18**	**0.13**	**0.17**
	SE	0.05	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05
Working memory	Estimate	**0.26**	**0.19**	**0.18**	**0.17**	**0.23**	**0.23**	**0.27**	**0.11**	−0.03
	SE	0.05	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05
Word-reading skill	Estimate	**0.15**	**0.10**	0.06	0.05	0.06	0.06	0.06	**0.11**	0.08
	SE	0.05	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05

Note. Coefficients in bold are statistically significant at $p < .05$. In-class attentive behavior is Strengths and Weaknesses of ADHD Symptoms and Normal Behavior rating scale (Swanson et al., 2012). Calculation fluency is Second-Grade Calculations Battery: Sums to 12, Sums to 18, Minuends to 12, Minuends to 18 (Fuchs et al., 2003). Language comprehension is Woodcock Diagnostic Reading Battery - Listening Comprehension (Woodcock, 1997). Reasoning is Wechsler Abbreviated Scale of Intelligence - Matrix Reasoning (Wechsler, 1999). Working memory is Working Memory Test Battery for Children - Listening Recall (Pickering & Gathercole, 2001). Word reading is Wide Range Achievement Test – Reading (Wilkinson, 1993).

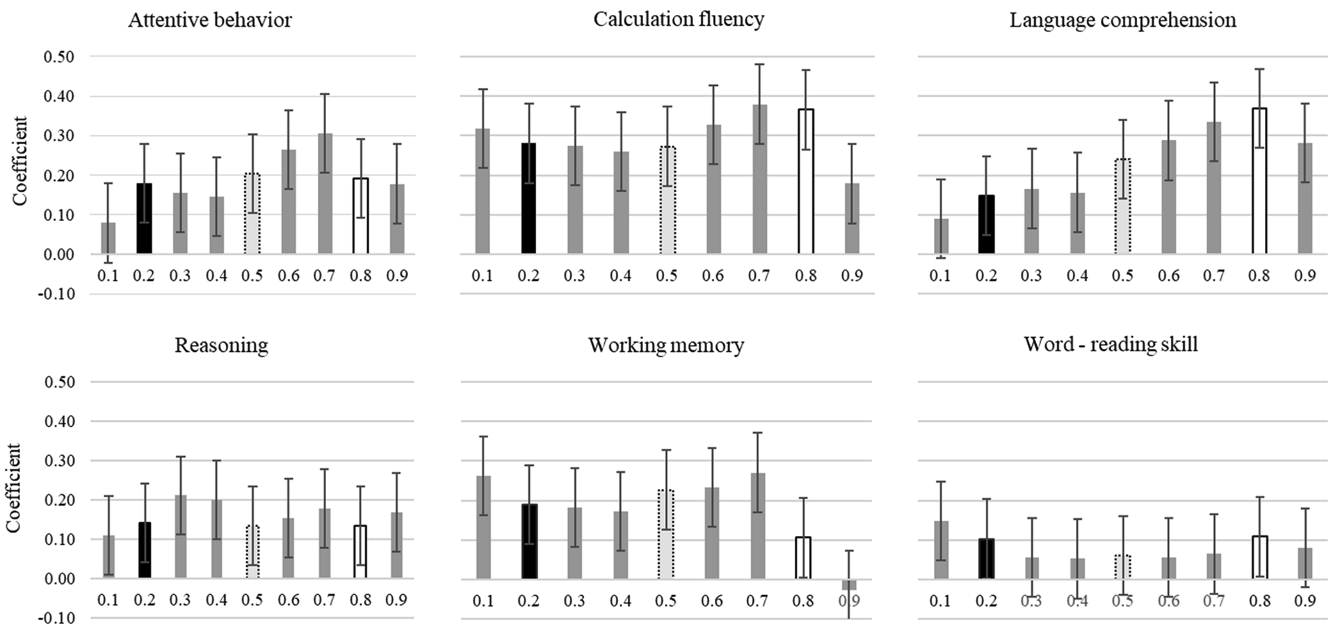

Fig. 1. UQR results showing the unique relation of each predictor to the WPS outcome. Error bars represent 95% confidence intervals. The focal quantiles are coded as follows: solid bars = low-level WPS quantile; light gray bars = intermediate-level WPS quantile; clear bars = high-level WPS quantile. Note that although the bars may appear fixed, they are not the same; they are dynamic.

Table 3
Comparison of coefficients across quantiles.

		0.50 vs. 0.20	0.80 vs. 0.50	0.80 vs. 0.20
In-class attentive behavior	Estimate	0.01	−0.01	0.00
	SE	0.05	0.05	0.06
Calculation fluency	Estimate	0.00	**0.10**	0.10
	SE	0.05	0.05	0.06
Language comprehension	Estimate	**0.10**	**0.13**	**0.24**
	SE	0.05	0.05	0.06
Reasoning	Estimate	0.00	−0.01	−0.01
	SE	0.04	0.05	0.06
Working memory	Estimate	0.03	**−0.12**	−0.09
	SE	0.05	0.05	0.06
Word-reading skill	Estimate	−0.03	0.04	0.01
	SE	0.05	0.05	0.06

Note. Coefficients in bold are statistically significant at $p < .05$. In-class attentive behavior is Strengths and Weaknesses of ADHD Symptoms and Normal Behavior rating scale (Swanson et al., 2012). Calculation fluency is Second-Grade Calculations Battery: Sums to 12, Sums to 18, Minuends to 12, Minuends to 18 (Fuchs et al., 2003). Language comprehension is Woodcock Diagnostic Reading Battery - Listening Comprehension (Woodcock, 1997). Reasoning is Wechsler Abbreviated Scale of Intelligence - Matrix Reasoning (Wechsler, 1999). Working memory is Working Memory Test Battery for Children - Listening Recall Pickering & Gathercole, 2001). Word reading is Wide Range Achievement Test – Reading (Wilkinson, 1993).

Prior studies document how each of these cognitive processes, as well as calculation fluency (Swanson & Beebe-Frankenberger, 2004) and word-reading skill (Fuchs et al., 2006), contribute to individual differences in WPS performance. Yet, prior work is limited to the estimation of average relations across the spectrum of WPS skill. In the present study, we extended this literature by instead asking how the engagement of each process differs as a function of second-grade WPS, while controlling for the effects of each of the other five component processes as well as variance attributable to schools, classrooms, and teachers.

We first corroborated prior work by demonstrating a role for five of the six component processes. Controlling for the effect of other predictors, all but the one component process was significantly related to the WPS outcome at each quantile, with only a few exceptions. Calculation fluency and reasoning were significantly related to WPS at each quantile; in-class attentive behavior and language comprehension at every quantile except the 0.10; working memory at every quantile except 0.90. By contrast, the hypothesized sixth component process, word-reading skill, was significantly related to WPS only at the 0.10, 20, and 0.80 quantiles. This was surprising given prior evidence (Fuchs et al., 2006) of a relation between word-reading skill and WPS, even when problems are read aloud to students as in the present study. This suggests the possibility that the positive unique contribution detected via more conventional analysis is attributable to the significant relation at the low- and high-end quantiles, without being robustly applicable. This was not, however, born out by the tests of differences between coefficients, and explanation awaits findings of future work designed to probe the role of word reading in WPS, perhaps via experimental methods.

Meanwhile, the tests of differences between coefficients, which provide greater stringency for considering significance versus non-significance in relations as a function of WPS ability, provided interesting insights. For working memory, results indicated a significantly higher regression weight for the role of working memory at the intermediate than at the high quantile. This negative association suggests that students with intermediate WPS skill, relative to those with high WPS skill, rely more on working memory when solving WPs. This demonstration of inconsistency for this relation depending on overall WPs competence, revealed in the present study's sample, echoes the individual differences literature: Some studies document such a role (e.g., Swanson & Beebe-Frankenberger, 2004; Swanson et al., 2008; Swanson & Sachse-Lee, 2001) while others do not (e.g., Fuchs et al., 2008, 2010b). The UQR finding suggests that such inconsistency in the literature reflects different levels of WPS competence among samples. This possibility is supported in a meta-analytic finding (Peng et al., in press), in which a stronger relation between working memory and mathematics performance was demonstrated for students with complicated mathematics learning disabilities than for students with math-only learning disabilities or for typically developing students.

Working memory may demonstrate a stronger relation with WPS for students with intermediate WP skill compared to those with stronger WP skill because intermediate-level WPS is more effortful and less fluent than is the case for stronger problem solvers. Part of this more effortful WPS involves greater reliance on working memory as intermediate-level problem solvers sequentially process information chunks within WP statements. Meanwhile, stronger engagement of working memory for intermediate- over high-level problem solvers, but not for low over intermediate- or high-level problem solvers, may reflect limited working memory capacity or less effective reliance on working memory capacity for low-level problem solvers.

At the same time, a different pattern of findings emerged for language comprehension, whereby its relation with WPS was significantly *weaker* at the 0.20 than at the 0.50 quantile, at the 0.20 than at the 0.80 quantile, and at the 0.50 than at the 0.80 quantile. These positive associations mirror the general trend, reflected in Fig. 1, in which the unique predictive value of language comprehension for WPS tends to increase from the 0.10 (URQ coefficient = 0.08) to the 0.80 (UQR coefficient = 0.37) quantile. Finding higher levels of language comprehension engagement for higher WPS quantiles corroborates correlational studies indicating that as language comprehension improves, so does WPS (Bernardo, 1999; Fuchs et al., 2016, 2018, 2010; Swanson & Beebe-Frankenberger, 2004; Van der Schoot et al., 2009).

The present UQR analysis deepens insight into the power of this relation by revealing constrained language comprehension processing for low-level problem solvers, even as the moderate recruitment of language comprehension for intermediate-level problem solvers is more constrained than for high-level problem solvers. That is, for high-level WPS, more sophisticated language comprehension reflects the range of language processing necessary to provide the strongest *relation* between language comprehension and WPS of the three quantiles.

Finding greater engagement of language comprehension in higher-level WPS, where language comprehension is stronger than for intermediate range problem solvers who in turn rely more on language comprehension than lower-level problem solvers, may seem counterintuitive. One might assume that stronger language comprehension means less need to recruit that ability. Yet, in many walks of life, a seemingly effortless and sophisticated performance relies on competent execution of component processes or subskills. For example, stealing a basketball to drive across court and dunk the ball presents an effortless picture, even though it engages a sophisticated set of subskills, without which the intended action fails. Present findings suggest the same applies to the recruitment of language comprehension in competent WPS. At the same time, one would expect that the expert basketball player relies less on working memory to execute the maneuver than would be the case for the intermediate- or low-level basketball player's attempt at the same maneuver.

Similar (not identical) to the pattern that emerged for language comprehension, the final significant difference in correlational strength as a function of quantile indicates a weaker relation between calculation fluency and WPS at the 0.20 and 0.50 quantiles than at the 0.80 quantile. Restricted calculation skill among these low- and intermediate-level second graders, compared to high-level problem solvers, may help explain the weaker relation here. This would argue for the importance of instructional focus designed to promote both the number knowledge that underlies competent arithmetic reasoning via reflective discourse (Cobb, Boufi, McClain, & Whitenack, 1997) and the strategies that support efficient calculations as paths toward improving WPS (Fuchs et al., 2013).

Alternatively, however, the weaker association with calculation fluency and WPS for low- and intermediate- versus high-level problem solvers may be due to a common and error-fraught approach observed in immature WPS, whereby children add the numbers in WP statements without engaging in mathematical reasoning to first build a problem statement. This produces WP errors approximately 50% of the time (because half the problems require subtraction) even if the calculation accuracy for that incorrect operation is correct. This alternative possibility underscores the importance of promoting mathematical reasoning to support the building of problem models when solving mathematical WPs.

With respect to educational implications for the present study's findings on working memory and language comprehension, it is important to note evidence of limited transfer from isolated cognitive process training to academic performance (e.g., Melby-Lervåg & Hulme, 2013). Therefore, when addressing limitations in these cognitive or linguistic processes, compensatory strategies may be required. For working memory, this might include, for example, teaching students strategies to compensate for constrained working memory capacity, such as circling or highlighting information in the problem narrative as they are identified as relevant; crossing out irrelevant information as they are identified as irrelevant; selecting and underlining within the narrative a good candidate word label for the problem answer; and drafting a visual or number sentence representation of the problem model as children listen or read and revising that representation as additional information comes on line.

With respect to language comprehension, it is interesting to consider results in light of prior work demonstrating that simplifying the wording of WP statements dramatically affects average solution accuracy by facilitating access to the problem's meaning (e.g., Cummins, 1991; Davis-Dorsey, Ross, & Morrison, 1991; De Corte et al., 1985; Vicente, Orarntia, & Verschaffel, 2007). Present UQR findings suggest such simplification may support performance not just at the mean, but instead throughout the spectrum of WPS ability. This reveals the need for future studies to explore whether WPS skill moderates the improvement in WPS performance accomplished via re-wording of WP statements. If the phenomenon proves robust across, then graduated rewording of WP statements, designed to systematically sensitize students to salient language constructions for understanding WPs, may prove an innovative and helpful strategy for improving WP comprehension and solution accuracy across the range of learners. In the meantime, present findings indicate the need to build students' language comprehension to facilitate understanding of and increased reliance on the math-laden vocabulary and grammatical constructions found in WP narratives to support WP-model building. This includes teacher-led language classroom discussions designed to sustain inquiry-based discussion and argumentation for students across the spectrum of WPS ability (e.g., Yackel & Cobb, 1996), as well as structured mathematics intervention shown to support low-level problem solvers' development (e.g., Fuchs et al., 2014).

We conclude that this study extends the individual differences literature on the role of component cognitive processes and academic skills in WPS by asking whether and if so how the engagement of these processes differs as a function of second-grade WPS ability, while controlling for variance attributable to schools, classrooms, and teachers. We found evidence that three component processes, working memory, language comprehension, and calculation fluency, are engaged differentially depending on students' WP skill. For calculation fluency and language comprehension, the role of the component process was stronger with more competent WPS ability. By contrast, the role of working memory was stronger with intermediate- than for strong-level problem solving. Results deepen insight into the role of these processes in WPS and provide the basis for hypothesizing how instructional strategies may be differentiated depending on students' overall level of WPS competence.

These conclusions must, however, be understood within the constraints of three study limitations. First, considering additional cognitive processes, such as other forms of working memory (inhibition and updating; Miyake & Shah, 1999), 3-D spatial visualization (Tolar, Lederberg, & Fletcher, 2009), and analogical or inferential reasoning (e.g., Holyoak & Thagard, 1997), may produce insights into how additional processes may be engaged differentially as a function of WPS ability. Second, we measured each construct with a particular measure. Including multiple measures to permit use of latent constructs is preferable, and this should be pursued. Third, conclusions about causality should be avoided because our methods were correlational.

Acknowledgements

This research was supported by 2 P20 HD075443 and R01 HD059179 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to Vanderbilt University and Core Grant HD15052 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to Vanderbilt University and by R324A150078 from the Institute of Education Sciences in the U.S. Department of Education to the University of Texas at Austin. 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, the National Institutes of Health, the Institute of Education Sciences, or the U.S. Department of Education.

References

Anderson, U. (2007). The contribution of working memory to children's mathematics word problem solving. *Applied Cognitive Psychology, 21*, 1201–1216. https://doi.org/10.1002/acp.1317.

Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., & Olson, R. K. (2013). The SWAN captures variance at the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders, 17*, 152–162. https://doi.org/10.1177/1087054711427399.

Bernardo, A. B. I. (1999). Overcoming obstacles to understanding and solving word problems in mathematics. *Educational Psychology, 19*, 149–163. https://doi.org/10.1080/0144341990190202.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology, 114*, 375–388. https://doi.org/10.1016/j.jecp.2012.09.015.

Bynner, J. M. (1997). Basic skills in adolescents' occupational preparation. *The Career Development Quarterly, 45*, 305–321.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education, 28*, 258–277. https://www.jstor.org/stable/749781.

Cummins, D. D. (1991). Children's interpretations of arithmetic word problems. *Cognition and Instruction, 8*, 261–289. https://doi.org/10.1207/s1532690xci0803_2.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*, 405–438. https://doi.org/10.1016/0010-0285(88)90011-4.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83*, 61–68. https://doi.org/10.1037//0022-0663.83.1.61.

De Corte, E., Verschafel, L., & De Win, L. (1985). Infuence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology, 77*, 460–470. https://doi.org/10.1037/0022-0663.77.4.460.

Every Child a Chance Trust (2009). *The long-term costs of numeracy difculties*. http://www.everychildachancetrust.org/counts/index.cfm. Accessed 14 Aug 2009.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica, 77*, 953–973. https://doi.org/10.3386/t0339.

Fuchs, D., & Fuchs, L. S. (2016). Responsiveness-to-intervention: A "systems" approach to instructional adaptation. *Theory into Practice, 55*, 225–233.

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., ... Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psycholog y, 48*, 1315–1326. https://doi.org/10.1037/a0027475.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., ... Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29–43. https://doi.org/10.1037/0022-0663.98.1.29.

Fuchs, L. S., Fuchs, D., Seethaler, P. M., & Barnes, M. A. (2019). The role of working memory in mathematical word-problem solving: Implications for instruction and intervention. *ZDM Mathematics Education*. https://doi.org/10.1007/s11858-019-01070-8.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. E. (2008). Problem-solving and computation skills: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*, 30–47. https://doi.org/10.1037/0022-0663.100.1.30.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. V. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520–1533. https://doi.org/10.1111/j.1467-8624.2010.01489.x.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., ... Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical and general cognitive abilities? *Developmental Psychology, 46*, 1731–1746. https://doi.org/10.1037/a0020662.

Fuchs, L. S., Zumeta, R. O., Schumacher, R. F., Powell, S. R., Seethaler, P. M., Hamlett, C. L., & Fuchs, D. (2010). 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. *Elementary School Journal, 110*, 440–463. https://doi.org/10.1086/651191.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., ... Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*, 58–77. https://doi.org/10.1037/a0030127.

Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & Martin, B. (2018). Text comprehension and oral language as predictors of word-problem solving: Insights into word-problem solving as a form of text comprehension. *Scientific Studies of Reading, 22*, 152–166. https://doi.org/10.1080/10888438.2017.1398259.

Fuchs, L. S., Gilbert, J. K., Powell, S. R., Cirino, P. T., Fuchs, D., Hamlett, C. L., ... Tolar, T. M. (2016). The role of cognitive processes, foundational math skill, and calculation accuracy and fluency in word-problem solving versus pre-algebraic knowledge. *Developmental Psychology, 52*, 2085–2098. https://doi.org/10.1037/dev0000227.

Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlett, C. L., ... Changas, P. C. (2014). Does calculation or word-problem instruction provide a stronger route to pre-algebraic knowledge? *Journal of Educational Psychology, 106*, 990–1006. https://doi.org/10.1037/a0036793.

Fuchs, L. S., Hamlett, C. L., & Powell, S. R. (2003). Computational battery. Available from L.S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

Geary, D. C., & Widaman, K. F. (1991). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence, 16*, 47–80. https://doi.org/10.1016/0160-2896(92)90025-M.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52*, 35–44. https://doi.org/10.1037/0003-066X.52.1.35.

Jordan, N. C., & Hanich, L. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33*, 567–578. https://doi.org/10.1177/002221940003300605.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109–129. https://doi.org/10.1037/0033-295x.92.1.109.

Killewald, A., & Bearak, J. (2014). Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review, 79*, 350–357. https://doi.org/10.1177/0003122414524574.

Koenker, R. (2015). Quantile regression. International encyclopedia of the social & behavioral sciences (2nd ed., pp. 712–718). https://doi.org/10.1016/B978-0-08-097086-8.42074-X.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50. https://doi.org/10.2307/1913643.

Lakes, K. D., Swanson, J. M., & Riggs, M. (2012). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal behavior rating scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders, 16*, 510–516. https://doi.org/10.1177/1087054711413550.

Lee, K., Ng, S. F., Ng, E. L., & Lim, Z. Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology, 89*, 140–158. https://doi.org/10.1016/j.jecp.2004.07.001.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270–291. https://doi.org/10.1037/a0028228.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge, UK: Cambridge University Press.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognitive and Instruction, 9*, 392–1389. https://doi.org/10.1207/s1532690xci0904_2.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2019). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000079 (in press).

Petscher, Y., & Logan, J. A. R. (2014). Quantile regression in the study of developmental sciences. *Child Development, 85*, 861–881. https://doi.org/10.1111/cdev.12190.

Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children.* London, England: Psychological Corporation.

Porter, S. R. (2015). Quantile regression: Analyzing changes in distributions instead of means. In M. B. Paulsen (Ed), Higher Education, Vol. 30: Handbook of Theory and Research (pp. 335–381). Cham, Switzerland: Springer.

Purpura, D. J., & Logan, J. A. R. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. *Developmental Psychology, 51*, 1717–1724. https://doi.org/10.1037/dev0000055.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110–122. https://doi.org/10.1016/j.lindif.2009.10.005.

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psycyhology, 98*, 265–281. https://doi.org/10.1037/0022-0663.98.2.265.

Swanson, H. L. (2016). Word problem solving, working memory and serious math difficulties: Do cognitive strategies really make a difference? *Journal of Applied Research in Memory and Cognition, 5*, 368–383. https://doi.org/10.1016/j.jarmac.2016.04.012.

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*, 471–491. https://doi.org/10.1037/0022-0663.96.3.471.

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*, 343–379. https://doi.org/10.1037/0022-0663.100.2.343.

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*, 294–321. https://doi.org/10.1006/jecp.2000.2587.

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., ... Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN rating scales. *The International Journal of Educational and Psychological Assessment, 10*, 51–69. https://doi.org/10.1037/1040-3590.15.1.3.

Swanson, H. L., Moran, A., Lussier, C., & Fung, W. (2014). The effect of explicit and direct generative strategy training and working memory on word problem-solving accuracy in children at risk for math difficulties. *Journal of Learning Disabilities, 37*, 111–123. https://doi.org/10.1177/0731948713507264.

Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: Computational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology, 29*, 239–266. https://doi.org/10.1080/01443410802708903.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Van der Schoot, M., Bakker Arkema, B. A. H., Horsley, T. M., & Van Lieshout, E. D. C. M. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: An eye movement study in primary school children. *Contemporary Educational Psychology, 34*, 58–66. https://doi.org/10.1016/j.cedpsych.2008.07.002.

Vicente, S., Orarntia, J., & Verschaffel, L. (2007). Influence of situational and conceptual rewording on word problem solving. *British Journal of Educational Psychology, 77*, 829–848. https://doi.org/10.1348/000709907X178200.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence.* San Antonio, TX: Psychological Corporation.

Wilkinson, G. S. (1993). *Wide range achievement test - 3.* Wilmington: Wide Range.

Woodcock, R. W. (1997). *Woodcock diagnostic reading battery.* Itasca, IL: Riverside.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research on Mathematics Education, 27*, 458–477. https://doi.org/10.2307/749877.