Published in final edited form as:
Sci Stud Read. 2018 ; 22(2): 152–166. doi:10.1080/10888438.2017.1398259.

Text Comprehension and Oral Language as Predictors of Word-Problem Solving: Insights into Word-Problem Solving as a Form of Text Comprehension

Lynn S. Fuchs, **Jennifer K. Gilbert**, **Douglas Fuchs**, **Pamela M. Seethaler**, and **BrittanyLee N. Martin**
Vanderbilt University

Abstract

This study was designed to deepen insights on whether word-problem (WP) solving is a form of text comprehension (TC) and on the role of language in WPs. A sample of 325 second graders, representing high, average, and low reading and math performance, was assessed on (a) start-of-year TC, WP skill, language, nonlinguistic reasoning, working memory, and foundational skill (word identification, arithmetic) and (b) year-end WP solving, WP-language processing (understanding WP statements, without calculation demands), and calculations. Multivariate, multilevel path analysis, accounting for classroom and school effects, indicated that TC was a significant and comparably strong predictor of all outcomes. Start-of-year language was a significantly stronger predictor of both year-end WP outcomes than of calculations, whereas start-of-year arithmetic was a significantly stronger predictor of calculations than of either WP measure. Implications are discussed in terms of WP solving as a form of TC and a theoretically coordinated approach, focused on language, for addressing TC and WP-solving instruction.

Based on theories of reading comprehension, discourse processing, and word-problem (WP) processing (van Dijk & Kintsch, 1993; Perfetti, Yang, & Schmalhofer, 2008; Rapp, Van Den Broek, McMaster, Kendeou, & Espin, 2007; Verschaffel & De Corte, 1997), text comprehension (TC) and WP solving may be defined in integrated fashion. This definition assumes that representations of text have three components. The first involves constructing a coherent microstructure and deriving a hierarchical macrostructure to capture the text's essential ideas. The second, the situation model, requires supplementing the text with inferences based on the reader's world knowledge; for WPs, this includes informal and formal knowledge about quantities. With the third component, the reader derives problem models or schema; for WPs, this focuses on relations among quantities. WP solving does, however, differ from other forms of TC in that identification of the schema triggers and guides strategies for answering the WP's question. This second, WP solution phase involves representing the problem model with numerals and deriving the mathematical result, evaluating whether the mathematical outcome is computationally correct and reasonable, and communicating the solution (Jiménez & Verschaffe, 2014).

CONTACT Lynn S. Fuchs, lynn.fuchs@vanderbilt.edu, Vanderbilt University, 228 Peabody, Nashville, TN 37203.

In line with an integrated definition, Kintsch and Greeno (1985) posited that the general features of the reading comprehension process apply across stories, informational passages, and WP statements but that the comprehension strategies, the nature of required knowledge structures, and the form of resulting structures, inferences, and problem models differ by task. Yet few studies have investigated connections or distinctions between WP solving and other forms of TC.

Prior work involving concurrent data collection suggests an association between TC and WP solving. Vilenius-Tuohimaa, Aunola, and Nurmi (2008) reported substantial shared variance across TC and WP solving when controlling for foundational reading skill. Swanson, Cooney, and Brock (1993) identified TC as a correlate of WP solving while controlling for working memory, knowledge of operations, WP propositions, and calculation skill. Boonen, Van Der Schoot, Florytvan, De Vries, and Jolles (2013) found that TC had medium to large relations with WP solving. This relation was not evident at the WP item-level in Boonen, Van Wesel, Jolles, and Van Der Schoot (2014); however, the authors indicated that their WP items did not involve the semantic complexity that warrants strong reliance on TC.

L. S. Fuchs, Fuchs, Compton, Hamlett, and Wang (2015) extended this concurrent literature by relying on a longitudinal design and by examining sources of individual differences in TC versus WP outcomes. Early in the school year, children were assessed on general language comprehension, working memory, nonlinguistic reasoning, processing speed, and foundational reading and math skill; at year's end, on WP-specific language comprehension, TC, and WPs. Path analytic mediation analysis indicated the effect of general language comprehension on TC was entirely direct, whereas the effect on WPs was partially mediated by WP-specific language. Yet, across both domains, effects of working memory and reasoning operated in parallel ways. These findings are in line with Kintsch and Greeno (1985), who suggested that WP solving is a form of TC involving language comprehension, working memory, and reasoning, but differing from other forms of TC by requiring WP-specific as well as general language comprehension.

With the present study, we extended studies employing concurrent designs as well as L. S. Fuchs et al. (2015) by testing effects of initial TC on year-end WP solving. If WP solving is a form of TC, then start-of-year TC should support (uniquely account for) individual differences in year-end WP solving. To create a stringent test, we examined the specificity of effects by contrasting the contribution of TC to year-end WP solving against TC's effects on year-end calculations. Based on studies indicating (a) shared concurrent variance between TC and WP solving (Boonen et al., 2013, 2014; Swanson et al., 1993; Vilenius-Tuohimaa et al., 2008), (b) substantially similar patterns of cognitive and linguistic predictors across TC and WP solving (L. S. Fuchs et al., 2015), and (c) shared but some distinctive predictive patterns for WP solving versus calculations (L. S. Fuchs et al., 2008; Swanson & Beebe-Frankenberger, 2004), we hypothesized that TC's effects are stronger on WP solving than on pure calculations. We also included a measure of WP-language processing to represent WP solving without calculation demands. Our hypothesis concerning the role of TC in WP-language processing mirrored the prediction for WP solving. Conversely, we expected initial arithmetic skill (basic facts) to predict year-end calculations (two-digit problems with and without regrouping) more strongly than either of the year-end WP measures.

In addition, we targeted general oral language (in this article, referred to as *language*) as a strategy for deepening insight into whether WP solving is a form of TC, by exploring the role of start-of-year language in WPs while controlling for start-of-year TC. Consistent with Kintsch and Greeno (1985) and given studies documenting connections between language and TC (Catts, Hogan, & Adolf, 2005; Gough & Tunmer, 1986) and between language and WPs (Bernardo, 1999; L. S. Fuchs et al., 2008, 2010; Van Der Schoot, Bakker Arkema, Horsley, & Van Lieshout, 2009), we expected the effects of start-of-year language to be stronger on both WP outcomes than on calculations. Finding a stronger role for language in WPs than in calculations, while controlling for effects of TC (which is expected to share variance with language and therefore compete with language as a predictor of WPs), would strengthen existing evidence for the importance of language within WP solving. A common role for language across WPs and TCs would represent an important connection between the two academic domains and would thus strengthen the basis for thinking that WP solving may be conceptualized as a form of reading comprehension.

To contextualize our focus on language, we note that the link between language and mathematical development has been established (e.g., Aiken, 1972; Powell, Driver, Roberts, & Fall, 2017; Purpura & Ganley, 2014). Moreover, language has been specifically identified as active in the development of exact calculation skill (Dehaene, Spelke, Pinel, Stanescu-Cosson, & Tsivkin, 1999; Gordon, 2004), and other studies suggest a connection among linguistic abilities and early math and reading. In Purpura, Logan, Hassinger-Das, and Napoli (2017), mathematics-specific language mediated the relation between preschool literacy and numeracy. LeFevre et al. (2010) found that a composite across vocabulary and elision predicted number naming, calculation skill, and word reading in 4.5- to 7-year-olds, and vocabulary specifically predicted number naming (*non*linguistic arithmetic), but the relation between vocabulary and calculations was not reported. In Harlaar, Kovas, Dale, Petrill, and Plomin (2012), phenotypic and genetic correlations between a math composite (including but not limited to calculations) and reading comprehension were stronger than between math and word decoding. None of these studies addressed WP solving.

Given that language has been identified as active in the development of exact calculation skill, it would not be surprising to find a relation between language and calculations. By contrast, the hypothesis in the present study is that language exerts a *stronger effect on WPs than on calculations*, even when controlling for effects of TC, which is expected to share variance with language and therefore compete with language as a predictor of WPs. The present study thus addressed two related issues: whether WP-solving may be considered a form of TC, and the role of language in WPs. Our methods submit the question of whether WP solving is a form of TC to a more stringent test than in previous studies. This includes a final, methodological extension: use of multivariate, multilevel path analysis to account for classroom and school effects, which increases precision in estimating effects and eliminates the nature of school context and classroom instruction as competing explanations for findings.

Findings, if supportive, would raise the possibility that TC and WP performance may both be improved with some efficiency using a theoretically coordinated approach that integrates language instruction into TC and WP instruction. Such an approach would address the needs of an especially vulnerable subset of the population: students with comorbid learning disorders across TC and WP solving. These students, who perform lower in each domain than do students with difficulty in one area (Willcutt et al., 2013), are at risk of poor long-term outcomes in and out of school (Batty, Kivimäki, & Deary, 2010; Every Child a Chance Trust, 2009; Meneghetti, Carretti, & De Beni, 2006; Murnane, Willett, Braatz, & Duhaldeborde, 2001; Ritchie & Bates, 2013). We return to this point in the discussion.

In the present study's design, we simultaneously modeled the effects of eight predictors on three outcomes while controlling for nesting at the school and classroom levels. As outlined, the outcomes were end-of-second-grade WP solving, WP-language processing, and calculations. Among the eight predictors were four cognitive and linguistic measures. These represent the constructs addressed in Kintsch's model of WPs as a form of TC and for which the literature indicates a role in TC and WPs: language (Catts et al., 2005; L. S. Fuchs et al., 2010; Gough & Tunmer, 1986); nonlinguistic reasoning (Chase, 1969; L. S. Fuchs et al., 2015; Geary & Widaman, 1991); and working memory (Carretti, Borella, Cornoldi, & De Beni, 2009; Miyake, Just, & Carpenter, 1994; Siegel & Ryan, 1989). To avoid biasing predictive relations, we operationalized working memory span with two measures, one involving words and the other numerals. The other four predictors were academic. To address the study's major questions, TC was included. We controlled for arithmetic, a foundational skill for WPs (L. S. Fuchs et al., 2006); word reading, a foundational skill for TC (Gough & Tunmer, 1986); and start-of-year WP performance (in this article, referred to as *early word-problems*). Note that in the district where the study took place, the school year starts in August and ends in May. Predictors were assessed in September-October and outcomes in April.

Method

Participants

A sample of 325 children was selected to represent high, average, and low reading and mathematics performance (as indexed on the Wide Range Achievement Test [WRAT; Wilkinson, 1993] Reading and Arithmetic) from 133 second-grade classrooms in 24 schools in an urban district in the United States. (We relied on selection to ensure a representative sample, because a high proportion of participants came from high-poverty backgrounds.) Children scoring below the ninth percentile on both subtests of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) were excluded to ensure that participants had access to the study measures. (One third of participants in this analysis were participants in the L. S. Fuchs et al., 2015, analyses. In this line of studies, data were collected in waves, with overlapping but not identical measures across the waves. The questions asked in the two papers articles differ.)

In August, the sample's mean age was 7 years 6 months; 50% were female; 78% were from families with low socioeconomic status (as revealed in eligibility for the U.S. school subsidized lunch program); 48% were African American, 31% non-Hispanic White, 15% White Hispanic, and 7% other; 6% had an identified disability; and 11% were learning English as a second language. All tests were administered in English.

In September-October, the percentile score equivalent for the sample's raw score mean on Gates–MacGinitie Reading Comprehension (MacGinitie, MacGinitie, Maria, & Dreyer, 2002) was 43. On KeyMath–Revised–Problem Solving (KeyMath; Connolly, 1998), the mean standard score was 106.06 (*SD* = 9.37); on WRAT-Arithmetic (Wilkinson, 1993), 94.72 (*SD* = 13.26); on WRAT-Reading, 103.74 (*SD* = 14.53); and on WASI (Wechsler, 1999), 96.11 (*SD* = 12.75).

Measures

Screening and descriptive pretests used to select and describe sample—With WRAT-Arithmetic (Wilkinson, 1993), children have 10 min to write answers to calculation problems of increasing difficulty. With WRAT-Reading (Wilkinson, 1993), children identify up to 15 letters and up to 42 words without a time limit. Testing is discontinued after 10 errors. At ages 7–9, the alpha is .90 to .92 on these WRAT subtests. The two-subtest WASI (Wechsler, 1999), which comprises Vocabulary and Matrix Reasoning, is an individually administered measure of general cognitive ability (reliability = .92). Gates (MacGinitie et al., 2002) is described under academic predictors; KeyMath (Connolly, 1998) under outcomes.

Cognitive and linguistic predictors: Language, nonlinguistic reasoning, and working memory—Language was measured with WASI–Vocabulary (Wechsler, 1999), with which children identify objects in pictures (first four items) and construct verbal definitions of words (remaining items). At 6–11 years, test stability is .85.

Nonlinguistic reasoning was measured with WASI–Matrix Reasoning (Wechsler, 1999), which assesses pattern completion, classification, analogy, and serial reasoning. For each item, children select one of five options that best completes a visual pattern. At 6–11 years, test stability is .79.

Working memory was assessed with Working Memory Test Battery for Children–Counting Recall and Listening Recall (Pickering & Gathercole, 2001). Each subtest has six items at span Levels 1–6 to 1–9. Passing four items at a level moves the child to the next, increasing the number of items to be remembered by one. Failing three items within a level terminates the subtest. The score is trials correct. For Listening Recall (WM-words), children determine if each sentence in a series is true; then they recall the last word of each sentence. For Counting Recall (WM-numerals), children count four, five, six, or seven dots on a series of cards; then they recall the numerals of the counted sets. At Grade 2, subtest stability is .83–.85.

Academic predictors: Word reading, TC, arithmetic, and early word-problem—Word reading was assessed with Word Identification Fluency (L. S. Fuchs, Fuchs, & Compton, 2004). Children have 1 min to read a single page of 50 high-frequency words randomly sampled from the Dolch preprimer, primer, and first-grade levels. If they hesitate on a word for 4 s, the tester tells them to proceed. If they finish in less than 1 min, the score is prorated. Test stability is .86.

TC was measured with Gates (Level 2, Form S; MacGinitie et al., 2002), which relies on code-related skill as well as language ability (Cutting & Scarborough, 2006). It includes 11 narrative or informational passages, each with four paragraphs (except the first, with three sentences). Three pictures are shown next to each paragraph; children construct the passage's explicit or implicit meaning to select the picture best representing the text. They have 35 min to read and select pictures. Alternate form reliability was .74–.92.

Arithmetic was assessed with Second-Grade Calculations Battery (SGCB)–Arithmetic (L. S. Fuchs, Hamlett, & Powell, 2003), including Sums to 12, Sums to 18, Minuends to 12, and Minuends to 18. On each subtest, students have 1 min to write answers to up to 25 problems. The score is number of correct answers across subtests (sample $\alpha = .95$).

Early word problems was assessed with Story Problems (Jordan & Hanich, 2000), comprising 14 combine, compare, and change WPs requiring single-digit addition or subtraction. Testers read items; children follow along on paper (sample $\alpha = .87$).

Calculations, WP solving, and WP language outcomes—Calculations was assessed with SGCB–2-Digit Addition and 2-Digit Subtraction (L. S. Fuchs et al., 2003). Testers provide students 3 min on each subtest to complete up to 20 problems requiring two-digit calculations with and without regrouping. The score is total correct across subtests (sample $\alpha = .85$).

WP solving was assessed with KeyMath (Connolly, 1998), which includes problems of increasing difficulty involving all four operations. Text is shown visually, while the examiner reads problems aloud. There is no time limit. Students can reread text, but the tester reads the problem aloud once. (We opted for a WP measure in which text is presented visually to students while testers read aloud. This ensures access to the WP content and reduces the likelihood students grab numbers from text without processing text, which is common in emerging readers.) Nine problems are in the range of second graders (within 2 *SD*s of the sample's mean performance). Seven assess combine, compare, and change WP types, eight requiring one-digit calculations and the others requiring two-digit calculations; one involves money; and one involves identifying a number series. Seven problems are routine, requiring a number answer; one involves identifying irrelevant information; and one involves connecting sentences to create a WP. Testing is discontinued after three consecutive errors. At Grade 2, split-half reliability is .67–.74.

WP language was assessed with WP-Language Assessment (L. S. Fuchs, DeSelms, & Deason, 2012). On each of two subtests, testers read WPs aloud while students follow along on paper and write responses on paper. The first subtest, Bigger/Smaller WP Language, assesses understanding of language that determines bigger and smaller quantities, with two item types. With the first (eight items), students identify whether the quantity referred to in the WP's question is the bigger number, smaller number, or difference between numbers (e.g., *Linda has 3 toys. She has 8 fewer toys than Jane. How many toys does Jane have?*). With the second type (eight items), students identify which of four sentences matches the meaning of a sentence describing a compare relationship (e.g., Sue has 4 fewer stickers than Jan, response options are: *Jan has 4 fewer stickers than Sue; Jan has 4 more stickers than*

Sue; Sue has 4 more stickers than Jan; None of the above). The second subtest (eight items), Compare/Change WP Language, assesses understanding of language that determines whether a WP compares two quantities or describes a change in quantity for one object. This involves use of *more* combined with *than* or *then* (e.g., *Sue had 4 pieces of candy. Then she went to the store and bought 8 more pieces. How many pieces of candy does she have now?* vs. *Sue had 4 more pieces of candy than Jim. Jim has 7. How many does Sue have?* The tester asks, *Does the problem tell us about the difference between two amounts of candy or about a starting amount of candy that changes?*) The score is number correct across subtests (sample $\alpha = .84$).

Procedure

Testers were trained to criterion on each measure and used standard administration directions. Procedural fidelity checks were conducted prior to data collection and were repeated until each tester scored 90% or higher. In September-October, testing occurred in large groups in classrooms on WRAT–Arithmetic, Gates, Story Problems, and SGCB–Arithmetic. Testing occurred individually in quiet school spaces on WASI, WRAT–Reading, Working Memory Test Battery for Children–Counting Recall and Listening Recall, and Word Identification Fluency. In April, KeyMath was administered individually, and the other two outcome measures in small groups of two to four. Interscorer agreement on each measure exceeded 98%. Individual sessions were audiotaped; 15% of tapes were selected randomly, stratified by tester, for accuracy checks by independent scorers. Agreement exceeded 98%. Data were double entered, with discrepancies resolved.

Data analysis and results

Nesting, skewness, multicollinearity, and model fit

Using multivariate, multilevel path analysis, students (Level 1) were nested in classrooms (Level 2), classrooms were nested in teachers (Level 3), and teachers were nested in schools (Level 4). (The study was run in cohorts, such that some teachers had different classrooms across years.) Multivariate outcome modeling in Mplus 7.4 (Muthén & Muthén, 1998–2015) does not accommodate four-level models. Thus, we first ran three separate unconditional multilevel models in Stata using the *mixed* command to calculate the intraclass correlation coefficient (ICC) for each nesting level (classrooms, teachers, and schools) by outcome. The purpose was to gauge whether all nesting levels were necessary to include in the baseline model. School ICCs were .06–.08; teacher ICCs were exactly .00 (there was no teacher-to-teacher variability in the outcomes when accounting for school and classroom membership); classroom ICCs were .00–.18. These results indicate that clustering at the teacher level accounted for no variance in the outcome. It was therefore omitted from further modeling.

After excluding the teacher level, the baseline model accounted for school and classroom membership to accurately estimate the Level-1 standard errors. The relatively small number of schools (24) compared to model parameters produced warning messages in MPlus. Therefore, clustering at the school level was handled by the TYPE = COMPLEX TWOLEVEL option in Mplus in which random intercepts were allowed for classrooms but not schools, although standard error and model chi-square computations were adjusted for school clustering via the COMPLEX option (see Muthén & Satorra, 1995). Due to some skewness in the three outcome variables, we used maximum likelihood estimation with robust standard errors to account for nonnormality. Prior to estimating the baseline model, multicollinearity was checked in Stata using the *collin* command; no variance inflation factor approached or exceeded the typical cutoff of 10.

Prior to running the baseline model with all predictor variables included, we ran an unconditional multivariate model to record the amount of variance at Level 1 (student/residual). We then used these estimates to compute a Level-1 proportion of variance accounted for (R^2) for each outcome. For the baseline model, all indicators of model fit were in the acceptable range (Kline, 2011). Specifically, the chi-square test of model fit was nonsignificant, $\chi^2(3) = 3.47$, $p = .32$, the RMSEA of .02 fell below the .05 cutoff, and the CFI and TLI were above the .95 cutoff at 1.00 and 0.99, respectively (Figure 1).

Proportion of variance explained and predictor variable effects

Means, standard deviations, and correlations are shown in Table 1. Path estimates (in *z*-score units) are listed in Table 2. Together, the predictors accounted for 49% of Level-1 variance (Level-1 R^2) in calculations, 54% in WP solving, and 42% in WP language. Table 3 provides these estimates along with Level-1 variance components for all outcomes in the unconditional and baseline models. Also shown are Level-1 residual correlations among the outcome variables (.04–.13). This indicates little of the unexplained variance of one outcome was related to that of another. In other words, the predictor variables in our model accounted for the vast majority of shared variance among the outcomes.

Before discussing trends for predictor variables for the baseline model, we interpret one path as a guide to interpretation. The estimated path from TC to calculations was 0.16, $SE = 0.07$, $p = .030$. (Interpretations of standardized path coefficients parallel those of standardized regression coefficients.) Significant predictors of the calculation outcome in descending order of path estimates were as follows: arithmetic (0.44), early word-problems (0.19), TC (0.16), and language (−0.12); in predicting WP solving, early word-problems (0.36), working memory-words (0.17), TC (0.15), language (0.12), and arithmetic (0.11); in predicting WP language, early word-problems (0.27), TC (0.24), language (0.13), and working memory-words (0.12).

When considering results by predictors rather than outcomes, three predictors were significant for all outcomes when controlling for the other predictors: language, early word-problems, and TC. Three predictors were nonsignificant for all outcomes when controlling for the other predictors: nonlinguistic reasoning, working memory-numerals, and word reading. As expected, arithmetic was a significant predictor of both outcomes involving calculations but not of WP language. Working memory-words predicted the two WP outcomes but not calculations.

See Table 4 for tests of the six a priori hypotheses involving the equality of predictor effects between two outcomes. To be clear, the null hypothesis tested in each case was that the standardized path of one predictor to two different outcomes was the same. The false discovery rate was controlled by adjusting the alpha criterion for each test using the

Benjamini–Hochberg step-up procedure (Benjamini & Hochberg, 1995); all *p* values less than .05 listed in Table 4 are significant when applying the adjusted alpha. Estimates reported in this section are differences in standardized path values.

We found that, controlling for the other predictors, language was indeed less predictive of calculations than either WP outcome. The standardized path from language to calculations was 0.24 *SD* less and 0.25 *SD* less, respectively, than the path from language to WP solving and from language to WP language. The arithmetic variable also behaved as expected, predicting calculations more strongly than the WP outcomes (difference of 0.33 for WP solving and 0.45 for WP language). Surprisingly, TC's path to calculations was comparably strong as TC's effect on WP solving (difference = 0.01) and on WP language (difference = 0.09).

Discussion

The present study addressed two related issues: whether WP solving is a form of TC and the role of language in WPs. In considering the value of language in predicting individual differences in WPs, we employed a stringent model that simultaneously controlled for TC and early word-problem skill, both of which have been shown to share variance with language (Catts et al., 2005, L. S. Fuchs et al., 2008; L. S. Fuchs et al., 2010; Gough & Tunmer, 1986; Van Der Schoot et al., 2009), even as we controlled for other salient cognitive processes and academic skills as well as school and classroom nesting. Even with this stringent model, language emerged as uniquely significant in forecasting individual differences in both WP outcomes. This includes WP solving as well as WP language processing. Moreover, language's effect was not only strong (path coefficients of .12 and .13), its predictive value was also specific: Language predicted each of the WP outcomes substantially and significantly more strongly than the contrasting mathematics outcome involving pure calculations. This provides strong evidence for the role of language in WP solving.

At the same time, we further extended insights into the nature of WP solving and its potential connections to TC by assessing whether initial arithmetic skill (basic facts) predicts year-end calculations (multidigit problems with and without regrouping) more strongly than both year-end WP measures. Results supported this specificity hypothesis, with path coefficients of .44 on calculations, .11 on WP solving, and –.01 on WP language.

Together, language's differentially strong value in predicting WP over calculation outcomes as well as initial arithmetic's differentially strong role in predicting calculation over WP outcomes suggest that WPs and calculations may represent distinct domains of academic performance. This corroborates other correlational findings (L. S. Fuchs et al., 2008; Swanson & Beebe-Frankenberger, 2004). It also echoes an experimental study (L. S. Fuchs et al., 2014) in which calculation intervention did not enhance WP learning, and WP intervention did not improve calculation learning.

Finding that the cognitive, linguistic, and academic predictors of WP performance are separable from those involving pure calculations, while finding a differentially strong role for language in WPs over calculations, also speaks to the question of whether WP solving is a form of TC. Kintsch and Greeno (1985) suggested this possibility decades ago. Vilenius-Tuohimaa et al. (2008), Boonen et al. (2013), Boonen et al. (2014), and Swanson et al. (1993) provided evidence of concurrent relations between TC and WPs. In a longitudinal framework, L. S. Fuchs et al. (2015) found more commonalities than differences in the predictors of WP solving and TC, with effects of working memory and nonlinguistic reasoning operating in parallel ways on both outcomes and the effects of total (general and WP-specific) language comparably important in determining individual differences in both outcomes. The major distinction was that the effects of language on TC was entirely direct, whereas the effect of language on WPs was partially mediated by WP-specific language.

Present findings provide additional evidence that WP solving may productively be conceptualized as a form of TC. This growing literature also suggests that the instructional framework for addressing WP solving should not focus dominantly on calculation skill, as is often the case in school instruction. Instead, results suggest an important role for TC instruction within WP teaching.

This instructional approach involves a strong focus on language, including but not limited to WP-specific vocabulary and syntactic knowledge (e.g., understanding the distinction between *more than* and *then there were more*; that the cause and effect in change WPs may be presented in either order within WP statements). Such an approach is consistent with recent calls (Catts & Kamhi, 2017; Ukrainetz, 2017) to intimately connect an instructional focus on oral language to specific TC task demands, even as that instruction targets the subset of learners with language deficits for such embedded language instruction and relies on outcome measures reflecting the relevant TC tasks to index effects. An integrated approach with a deliberate focus on the TC demands of WP statements may also include methods to assist students in constructing explicit text-level representations, generating text-connecting inferences, retrieving general as well as math-specific background knowledge, and integrating that knowledge with information in text-level representations (Perfetti et al., 2008; Rapp et al., 2007; Verschaffel & De Corte, 1997), all in the service of building the WP situation and problem model.

Moreover, an integrated instructional framework may ultimately provide direction for a theoretically coordinated approach for simultaneously improving performance across TC and WPs in relevant ways. This might include, for example, teaching cause–effect informational text structure in conjunction with change WPs (in which an event serves to increase or decrease a starting amount, thereby creating a new ending amount) or connecting compare–contrast informational text structure with WPs that compare quantities.

Testing effects of a more inclusive approach to TC intervention, designed to focus on TC and WPs in coordinated fashion, would provide experimental evidence on whether WP solving is a form of TC, even as it would extend theoretical understanding of both domains. This line of work is potentially important for three additional reasons. First, students with comorbid learning disorders across TC and WPs represent an especially vulnerable subset of the population, because TC is a strong predictor of quality of life, financial security, and life expectancy (Batty et al., 2010; Meneghetti et al., 2006; Ritchie & Bates, 2013), even as WP solving is the best school-age predictor of later employment and wages (Every Child a Chance Trust, 2009; Murnane et al., 2001). Second, students with concurrent difficulty perform lower in each domain than students with difficulty in only one domain (Willcutt et al., 2013). Third, schools are challenging environments for providing students with intervention on more than one academic domain. A coordinated approach for addressing TC and WPs would alleviate this logistical problem.

We also note that WP solving, as operationalized in the school curriculum, involves word reading. Yet, in the present study, both WP measures (WP solving and WP language) were read aloud to children (although with both measures, each WP is presented visually at the same time, and the text remains available after the tester's reading ends, until the child has completed the item). When students are required to independently read text for accessing WP statements without the reading-aloud support, as is typically the case in school generally and on high-stakes assessments, connections between WP solving and TC may increase further. Similarly, differences in assessment modality (visual for TC but oral-visual for WPs) may have affected findings. These possibilities should be addressed empirically in studies that systematically vary calculation demands within the WP process (as in the present study, by including measures of WP solving as well as WP language without calculation demands) while varying reading demands and modalities in assessing WP solving (including measures that require reading vs. listening to WP statements, a contrast not included in the present study).

This brings us to the issue of whether the present study provides support for TC as a predictor of the WP outcomes. On one hand, in the context of our stringent model that controlled for initial word reading, language, nonlinguistic reasoning, two forms of working memory, arithmetic, and early WP skill, TC emerged as a significant predictor of and accounted for a substantial portion of variance in year-end WP solving. The path coefficient of .15 was the largest predictor of this outcome, after controlling for early WP skill (with a coefficient of the same magnitude for working memory with words). A similar pattern emerged for the WP-language processing outcome, with a path coefficient of .24 for this WP outcome absent the calculation demands.

On the other hand, despite TC's strong value in predicting the WP outcomes, TC was not a specific predictor. Contrary to our hypothesis, the magnitude of the TC's path coefficient to WP solving and WP-language processing was comparable to its predictive value for end-of-second-grade calculations. In fact, the point estimates for TC's predictive value on year-end calculations and WPs were nearly identical (.16 and .15). The strong predictive value of TC in explaining individual differences in calculations is surprising.

Three explanations appear plausible. First, because we designed the study with WPs as the major outcome, we did not control for at least two viable predictors of the calculation outcome, processing speed (L. S. Fuchs et al., 2006; L. S. Fuchs et al., 2008; Geary, 2011) and visual-spatial ability (Geary et al., 2009; Holmes, Adams, & Hamilton, 2008; Zhang & Lin, 2015). Including those predictors in the model may have reduced the estimate of TC's role in explaining the calculation outcome.

The second explanation is that reading exerts an effect on math. Some prior provides such evidence. For example, L. S. Fuchs, Geary, Fuchs, Compton, and Hamlett (2016) found that the effects of reading competency at the start of first grade on end-of-third-grade calculation performance were sizeable and significant (the reverse was not true), but findings on this point are mixed (e.g., Duncan et al., 2007). Additional research that systematically varies effects on higher versus lower order academic processes as well as reading versus mathematics (word-level reading skill vs. calculations vs. TC vs. WP solving) is needed.

The third potential explanation raises questions about what broad, multi-component TC measures, such as the Gates, actually assess. Such measures transparently tap an array of linguistic reasoning abilities as well as depth of background knowledge, even as they may index the strategic behavior involved in multiple-choice testing. Such measures have also demonstrated poor sensitivity to the effects of intervention within experimental investigations, even when measures aligned to the instructional targets and considered essential to the TC process reveal responsiveness (Alfassi, 1998; Catts & Kamhi, 2017; D. Fuchs et al., in press; Jitendra, Hoppes, & Xin, 2000). Present findings, in which TC predicted performance on later calculation performance, also raise the possibility that broad, multicomponent TC measures may operate as indicators of overall cognitive ability, raising questions about the tenability of broad, multicomponent TC measures as outcomes of TC instructional studies. In fact, discarding TC instructional procedures on the basis of low effect sizes on such broad measures, in the face of demonstrated effects on more narrow measures, may represent Type II error. At the same time, finding that TC is broadly predictive of three mathematics outcomes indicates the need to study the role of an array of TC assessments in predicting WP solving, while incorporating other forms of mathematics outcomes to assess specificity, as in the present. Such a research program can provide insight not only into connections between TC and WP solving but also the very structure of academic performance.

Before closing, we note two important study limitations. First, we relied exclusively on a measure of vocabulary as the indicator of language ability, without indexing syntax or sentence processing. Although WASI-Vocabulary requires children to construct verbal definitions of words and thus reflects broader language ability, the WP-specific language comprehension task highlights how sentence structure operates within WP solving. Exclusive reliance on vocabulary, even using a measure that reflects broader language ability by requiring students to construct verbal definitions, runs the risk of overestimating the importance of TC and underestimating the importance of language in WP performance. Second, the present study's sample was largely from families of low socioeconomic status (78% of participants met eligibility for the U.S. school subsidized lunch program). Thus, it is important that future studies index language ability more broadly and investigate the robustness of findings for diverse samples.

Acknowledgments

Funding

This research was supported by P20 HD075443 and R01 HD053714 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

Aiken LR. Language factors in learning mathematics. Review of Educational Research. 1972; 42:359–385. DOI: 10.3102/00346543042003359

Alfassi M. Reading for meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. American Educational Research Journal. 1998; 35:309–332. DOI: 10.3102/00028312035002309

Batty GD, Kivimäki M, Deary IJ. Intelligence, education, and mortality. BMJ: British Medical Journal. 2010; 340:1–2. DOI: 10.1136/bmj.c563

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B. 1995; 57:289–300. DOI: 10.2307/2346101

Bernardo ABI. Overcoming obstacles to understanding and solving word problems in mathematics. Educational Psychology. 1999; 19:149–163. DOI: 10.1080/0144341990190203

Boonen AJH, Van Der Schoot M, Florytvan W, De Vries H, Jolles J. What underlies successful word problem solving? A path analysis in sixth grade students. Contemporary Educational Psychology. 2013; 38:271.doi: 10.1016/j.cedpsych.2013.05.001

Boonen AJH, Van Wesel F, Jolles J, Van Der Schoot M. The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. International Journal of Educational Research. 2014; 68:15–26. DOI: 10.1016/j.ijer.2014.08.001

Carretti B, Borella E, Cornoldi C, De Beni R. Role of working memory in explaining the performance of individuals with specific comprehension difficulties: A meta-analysis. Learning and Individual Differences. 2009; 19:246–251. DOI: 10.1016/j.lindif.2008.10.002

Catts, HW., Hogan, TP., Adolf, SM. Developmental changes in reading and reading disabilities. In: Catts, HW., Kamhi, AG., editors. The connections between language and reading disabilities. Mahwah, NJ: Erlbaum; 2005. p. 25-40.

Catts HW, Kamhi AG. Prologue: Reading comprehension is not a single ability. Language, Speech, and Hearing Services in Schools. 2017; 48(2):73–76. DOI: 10.1044/2017_LSHSS-16-0033

Chase CI. Concept formation as a basis for measuring developmental reading. The Elementary School Journal. 1969; 69:210–214. DOI: 10.1086/460501

Connolly, AJ. KeyMath–Revised. Circle Pines, MN: American Guidance; 1998.

Cutting LE, Scarborough HS. Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. Scientific Studies of Reading. 2006; 10:277–299. DOI: 10.1207/s1532799xssr1003_5

Dehaene S, Spelke ES, Pinel P, Stanescu-Cosson R, Tsivkin S. Sources of mathematical thinking: Behavioral and brain-imaging evidence. Science. 1999; 284(5416):970–974. DOI: 10.1126/science.284.5416.970 [PubMed: 10320379]

Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, Japel C. School readiness and later achievement. Developmental Psychology. 2007; 43:1428–1446. DOI: 10.1037/0012-1649.43.6.1428 [PubMed: 18020822]

Every Child a Chance Trust. The long-term costs of numeracy difficulties. 2009. Retrieved from http://www.everychildachancetrust.org/counts/index.cfm

Fuchs D, Hendricks E, Walsh ME, Fuchs LS, Gilbert JD, Tracy W, Kim W. Evaluating the efficacy of a multidimensional reading comprehension program for at-risk students and reconsidering the lowly reputation of tests of near transfer. Learning Disabilities Research and Practice. (in press).

Fuchs, LS., DeSelms, J., Deason, JL. Word problem-specific language assessment. Nashville, TN 37203: Vanderbilt University; 2012. Available from L.S. Fuchs

Fuchs LS, Fuchs D, Compton DL. Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. Exceptional Children. 2004; 71:7–21. DOI: 10.1177/001440290407100101

Fuchs LS, Fuchs D, Compton DL, Hamlett CL, Wang AY. Is word-problem solving a form of text comprehension? Scientific Studies of Reading. 2015; 19:204–223. DOI: 10.1080/10888438.1005745 [PubMed: 25866461]

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, 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, Seethaler PM, Schatschneider C. Do different types of school mathematics development depend on different constellations of numerical and general cognitive abilities? Developmental Psychology. 2010; 46:1731–1746. DOI: 10.1037/a0020662 [PubMed: 20822213]

Fuchs LS, Geary DC, Fuchs D, Compton DL, Hamlett CL. Pathways to third-grade calculation versus word-reading competence: Are they more alike or different? Child Development. 2016; 87:558–567. DOI: 10.1111/cdev.12474 [PubMed: 26700885]

Fuchs, LS., Hamlett, CL., Powell, SR. Second-grade calculations battery. Vanderbilt University; Nashville, TN 37203: 2003. Available from L. S. Fuchs

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. DOI: 10.1037/a0036793 [PubMed: 25541565]

Geary DC. Cognitive predictors of achievement growth in mathematics: A five year longitudinal study. Developmental Psychology. 2011; 47:1539–1552. DOI: 10.1037/a0025510 [PubMed: 21942667]

Geary DC, Bailey DH, Littlefield A, Wood P, Hoard MK, Nugent L. First-grade predictors of mathematical learning disability: A latent class trajectory analysis. Cognitive Development. 2009; 24:411–429. DOI: 10.1016/j.cogdev.2009.10.001

Geary DC, Widaman KF. Numerical cognition: On the convergence of componential and psychometric models. Intelligence. 1991; 16:47–80. DOI: 10.1016/0160-2896(92)90025-M

Gordon P. Numerical cognition without words: Evidence from Amazonia. Science. 2004; 306(5695): 496–499. DOI: 10.1126/science.1094492 [PubMed: 15319490]

Gough PB, Tunmer WE. Decoding, reading, and reading disability. Remedial and Special Education. 1986; 7:6–10. DOI: 10.1177/074193258600700104

Harlaar N, Kovas Y, Dale PS, Petrill SA, Plomin R. Mathematics is differentially related to reading comprehension and word decoding: Evidence from a genetically sensitive design. Journal of Educational Psychology. 2012; 104:622N635.doi: 10.1037/a0027646

Holmes J, Adams JW, Hamilton CJ. The relationship between visuospatial sketchpad capacity and children's mathematical skills. European Journal of Cognitive Psychology. 2008; 2:272–289. DOI: 10.1080/09541440701612702

Jiménez L, Verschaffe L. Development of children's solutions of non-standard arithmetic word problem solving. Revista De Psicodidáctica. 2014; 9(1):93–123. DOI: 10.1387/RevPsicodidact.786

Jitendra AK, Hoppes MK, Xin YP. Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. Journal of Special Education. 2000; 34:127–139. DOI: 10.1177/002246690003400302

Jordan NC, Hanich L. Mathematical thinking in second-grade children with different forms of LD. Journal of Learning Disabilities. 2000; 33:567–578. DOI: 10.1177/002221940003300605 [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. 3. New York, NY; Guilford: 2011.

LeFevre JA, Fast L, Skwarchuk SL, Smith-Chant BL, Bisanz J, Kamawar D, Penner-Wilger M. Pathways to mathematics: Longitudinal predictors of performance: Pathways to mathematics.

Child Development. 2010; 81:1753–1767. DOI: 10.1111/j.1467N8624.2010.01508.x [PubMed: 21077862]

MacGinitie, WH., MacGinitie, RK., Maria, K., Dreyer, LG. Gates-MacGinitie reading tests. 4. Itasca, IL: Riverside; 2002.

Meneghetti C, Carretti B, De Beni R. Components of reading comprehension and scholastic achievement. Learning and Individual Differences. 2006; 16:291–301. DOI: 10.1016/j.lindif.2006.11.001

Miyake A, Just MA, Carpenter PA. Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. Journal of Memory and Language. 1994; 33:175–202. DOI: 10.1006/jmla.1994.1009

Murnane RJ, Willett JB, Braatz MJ, Duhaldeborde Y. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. Economics of Education Review. 2001; 20:311–320. DOI: 10.1016/S0272-7757(00)00056-X

Muthén, LK., Muthén, BO. Mplus user's guide. 7. Los Angeles, CA: Muthén & Muthén; 1998–2015.

Muthén BO, Satorra A. Complex sample data in structural equation modeling. Sociological Methodology. 1995; 25:267–316. DOI: 10.2307/271070

Perfetti C, Yang C, Schmalhofer F. Comprehension skill and word to text integration processes. Applied Cognitive Psychology. 2008; 22:303–318. DOI: 10.1002/(ISSN)1099-0720

Pickering, S., Gathercole, S. Working memory test battery for children. London, UK: The Psychological Corporation; 2001.

Powell SR, Driver MK, Roberts G, Fall AM. An analysis of the mathematics vocabulary knowledge of third- and fifth-grade students: Connections to general vocabulary and mathematics computation. Learning and Individual Differences. 2017; 57:22–32. DOI: 10.1016/j.lindif.2017.05.011

Purpura DJ, Ganley CM. Working memory and language: Skill-specific or domain-general relations to mathematics? Journal of Experimental Child Psychology. 2014; 122:104–121. DOI: 10.1016/j.jecp.2013.12.009 [PubMed: 24549230]

Purpura DJ, Logan JAR, Hassinger-Das B, Napoli AR. Why do early mathematics skills predict later reading? The role of mathematical language. Developmental Psychology. 2017; 53:1633–1642. DOI: 10.1037/dev0000375 [PubMed: 28758785]

Rapp DN, Van Den Broek P, McMaster KL, Kendeou P, Espin CA. Higher-order comprehension processes in struggling readers: A perspective for research and intervention. Scientific Studies of Reading. 2007; 11:289–312. DOI: 10.1080/10888430701530417

Ritchie SJ, Bates TC. Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. Psychological Science. 2013; 24:1301–1308. DOI: 10.1177/0956797612466268 [PubMed: 23640065]

Siegel LS, Ryan EB. The development of working memory in normally achieving and subtypes of learning disabled children. Child Development. 1989; 60:973–980. DOI: 10.2307/1131037 [PubMed: 2758890]

Swanson HL, Beebe-Frankenberger M. The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious mathematics difficulties. Journal of Educational Psychology. 2004; 96:471–491. DOI: 10.1037/0022-0663.96.3.471

Swanson HL, Cooney JB, Brock S. The influence of working memory and classification ability on children's word problem solution. Journal of Experimental Child Psychology. 1993; 55:374–395. DOI: 10.1006/jecp.1993.1021

Ukrainetz TA. Commentary on "Reading Comprehension Is Not a Single Ability": Implications for child language intervention. Language, Speech, and Hearing Services in Schools. 2017; 48(2):92–97. DOI: 10.1044/2017_LSHSS-16-0031

Van Der Schoot M, Bakker Arkema BAH, Horsley TM, Van Lieshout EDCM. 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. 2009; 34:58–66. DOI: 10.1016/j.cedpsych.2008.07.002

van Dijik, TA., Kintsch, W. Strategies of discourse comprehension. New York: Academic Press; 1993.

Verschaffel L, De Corte E. Teaching realistic mathematical modeling in the elementary school: A teaching experiment with fifth graders. Journal for Research in Mathematics Education. 1997; 28:577–601. DOI: 10.2307/749692

Vilenius-Tuohimaa PM, Aunola K, Nurmi JE. The association between mathematical word problems and reading comprehension. Educational Psychology. 2008; 28:409–426. DOI: 10.1080/01443410701708228

Wechsler, D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corporation; 1999.

Wilkinson, GS. Wide Range Achievement Test 3. Wilmington, DE: Wide Range; 1993.

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]

Zhang X, Lin D. Pathways to arithmetic: The role of visual-spatial and language skills in written arithmetic, arithmetic word problems, and nonsymbolic arithmetic. Contemporary Educational Psychology. 2015; 41:188–197. DOI: 10.1016/j.cedpsych.2015.01.005

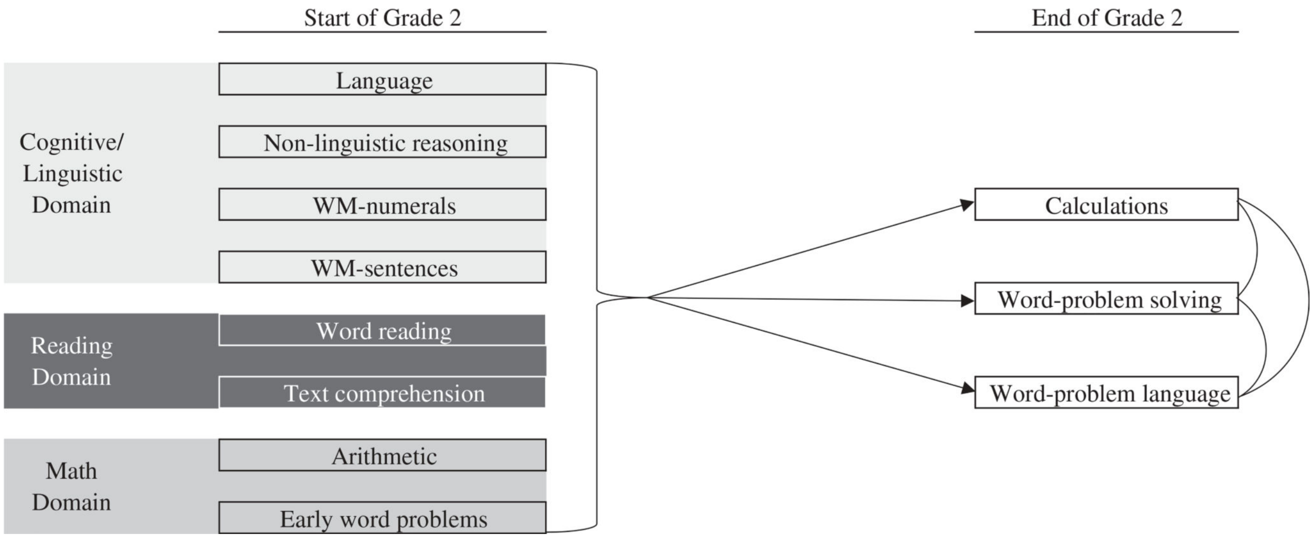

Figure 1.
Baseline model in which all end-of-second-grade outcomes were regressed on all start-of-second-grade predictors.

Note. In the model, there was a path from each predictor to each outcome; the brace in this figure is for visual clarity only.

Table 1

Means, Standard Deviations, and Correlations Among Variables.

			Correlations										
Domain/Variable	**M**	**SD**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Start-of-Grade 2 Predictors													
Cognitive & Linguistic Domain													
1. Language	24.40	6.59	—										
2. Nonlinguistic reasoning	13.76	5.91	.17	—									
3. WM–Numerals	14.77	5.15	.21	.35	—								
4. WM–Words	7.53	4.23	.43	.32	.39	—							
Reading Domain													
5. Word reading	51.61	25.07	.38	.22	.34	.34	—						
6. Text comprehension	26.98	7.44	.53	.33	.43	.50	.71	—					
Math Domain													
7. Arithmetic	6.52	3.25	.21	.31	.38	.36	.46	.41	—				
8. Early word problems	8.10	3.99	.44	.43	.49	.54	.53	.60	.51	—			
End-of-Grade 2 Outcomes													
9. Calculations	18.68	9.13	.18	.24	.41	.35	.40	.43	.63	.52	—		
10. Word-problem solving	4.52	2.13	.46	.40	.41	.55	.44	.57	.46	.68	.53	—	
11. Word-problem language	19.84	6.07	.44	.30	.36	.46	.42	.56	.32	.57	.37	.61	—

Note. N = 325. Language = WASI Vocabulary (Wechsler, 1999); Nonlinguistic reasoning = WASI Matrix Reasoning; WM-numerals = Working Memory Test Battery for Children (WMTB)–Counting Recall (Pickering & Gathercole, 2001); WM-words = WMTB-Listening Recall; Word reading = Word Identification Fluency (Fuchs et al., 2004); Text comprehension = Gates-MacGinitie–Comprehension (MacGinitie et al., 2002); Arithmetic = Second-Grade Calculations Battery (SGCB)–Arithmetic (Fuchs et al., 2003); Early word problems = Story Problems (Jordan & Hanich, 2000); Calculations = SGCB–2-Digit Addition and 2-Digit Subtraction (Fuchs et al., 2003); Word problem solving = KeyMath-Problem Solving (Connolly, 1998); Word-problem language = Word Problem-Specific Language Assessment (Fuchs et al., 2012).

Table 2

Standardized Results From Baseline Model.

	Outcome								
	Calculations			**Word-Problem Solving**			**Word-Problem Language**		
Domain/Predictor	**Est.**	***SE***	***p***	**Est.**	***SE***	***p***	**Est.**	***SE***	***p***
Cognitive & Linguistic									
Language	−0.12	0.04	.001	0.12	0.04	.001	0.13	0.06	.019
Nonlinguistic reasoning	−0.06	0.05	.212	0.09	0.05	.065	0.03	0.05	.580
WM–Numerals	0.10	0.06	.073	0.02	0.04	.696	0.05	0.05	.342
WM–Words	0.05	0.05	.275	0.17	0.03	<.001	0.12	0.05	.020
Reading									
Word reading	−0.03	0.06	.652	−0.03	0.05	.578	0.00	0.05	.945
Text comprehension	0.16	0.07	.030	0.15	0.06	.007	0.24	0.06	<.001
Math									
Arithmetic	0.44	0.04	<.001	0.11	0.05	.014	−0.01	0.06	.918
Early word problems	0.19	0.06	.001	0.36	0.06	<.001	0.27	0.08	.001
Random effects									
Classroom	0.01	0.06	.906	0.00	0.03	.995	0.00	0.03	.996
Level 1 residual	0.38	0.03	<.001	0.44	0.05	<.001	0.58	0.05	<.001

Note. N = 325. Calculations = Second-Grade Calculations Battery–2-Digit Addition and 2-Digit Subtraction (Fuchs et al., 2003); Word-problem solving = KeyMath-Problem Solving (Connolly, 1998); Word-problem language = Word Problem-Specific Language Assessment (Fuchs et al., 2012); Est. = model estimate; Language = WASI Vocabulary (Wechsler, 1999); Non-linguistic reasoning = WASI Matrix Reasoning; WM-numerals = WMTB–Counting Recall (Pickering & Gathercole, 2001); WM-words = WMTB–Listening Recall; Word reading = Word Identification Fluency (Fuchs et al., 2004); Text comprehension = Gates-MacGinitie–Comprehension (MacGinitie et al., 2002); Arithmetic = SGCB–Arithmetic (Fuchs et al., 2003); Arithmetic = Second-Grade Calculations Battery (SGCB)-Arithmetic (Fuchs et al., 2003); Early word problems = Story Problems (Jordan & Hanich, 2000).

Table 3

Level 1 Variance Components.

				Residual Correlation		
Outcome	**Unconditional Model Variance**	**Base Model Variance**	R^2	**1**	**2**	**3**
1. Calculations	0.76	0.39	.49	—		
2. Word-problem solving	0.96	0.44	.54	.09	—	
3. Word-problem language	0.99	0.58	.42	.04	.13	—

Note. Calculations = Second-Grade Calculations Battery–2-Digit Addition and 2-Digit Subtraction (Fuchs et al., 2003); Word-problem solving = KeyMath-Problem Solving (Connolly, 1998); Word-problem language = Word Problem-Specific Language Assessment (Fuchs et al., 2012).

Table 4

Tests of Relative Strength of Predictors Across Outcomes.

Domain/Predictor	Hypothesis Tested	Est.	*SE*	Est./ *SE*	*p*	Data Support Hypothesis?
Cognitive & Linguistic						
Language	*b* → calculations < *b* → WP solving	−0.24	0.04	−5.51	<.001	Yes
	b → calculations < *b* → WP language	−0.25	0.07	−3.52	<.001	Yes
Reading						
Text comprehension	*b* → calculations < *b* → WP solving	0.01	0.10	0.11	.916	No
	b → calculations < *b* → WP language	−0.09	0.10	−0.85	.397	No
Math						
Arithmetic	*b* → calculations (two-digit) > *b* → WP solving	0.33	0.05	6.34	<.001	Yes
	b → calculations (two-digit)> *b* → WP language	0.45	0.07	6.13	<.001	Yes

Note. N = 325. Language = WASI Vocabulary (Wechsler, 1999); Calculations = Second-Grade Calculations Battery–2-Digit Addition and 2-Digit Subtraction (Fuchs et al., 2003); WP = word problems; WP solving = KeyMath-Problem Solving (Connolly, 1998); WP language = Word Problem-Specific Language Assessment (Fuchs et al., 2012); Text comprehension = Gates-MacGinitie–Comprehension (MacGinitie et al., 2002); Arithmetic = SGCB-Arithmetic (Fuchs et al., 2003).