Addressing the Role of Working Memory in Mathematical Word-Problem Solving When Designing Intervention for Struggling Learners

Lynn Fuchs, Douglas Fuchs, Pamela M. Seethaler, and Marcia A. Barnes

Vanderbilt University

Publication Date: In Press/ Accepted for Publication in ZDM Mathematics Education

Document Completion Date: June 26, 2019

Authors' address: 228 Peabody, Vanderbilt University, Nashville, TN 37203

Corresponding author: lynn.fuchs@vanderbilt.edu; 615-343-4782

This research was supported by Grant R305A150200 from the U.S. Department of Education's Institute of Education Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences or the U.S. Department of Education.

Abstract

The focus of this article is the well documented association between low working memory capacity and difficulty with mathematical word-problem solving. We begin by describing a model that specifies how various cognitive resources, including working memory, contribute to individual differences in word-problem solving and by then summarizing findings on the relation between working memory and word-problem solving. This sets the context for the article's main purpose and major section: to describe the findings of research studies that take one of two approaches for addressing the needs of students with low working memory within word-problem solving intervention. One approach focuses on *compensating* for working memory limitations; the other on *building* working memory capacity. We then suggest the need for research on integrating the two approaches by embedding working memory training within explicit word-problem solving intervention.

Key words: mathematics word problems; word-problem solving; working memory; working memory training

The Role of Working Memory within Mathematical Word-Problem Solving:

Implications for Intervention

Difficulty with word problems can occur even when other forms of mathematical cognition are intact (Cummins et al., 1988; Koedinger & Nathan, 2004). Specific word-problem difficulty occurs in part because the cognitive resources involved when solving word problems differ from and are more numerous than those underlying number knowledge or calculation skill (e.g., Swanson & Beebee-Frankenberger 2004). Thus word-problem difficulty is multiply determined and may be difficult to overcome.

It is therefore important that schools assume a proactive stance to prevent such difficulty by reaching all students via effective word-problem classroom instruction. Moreover, for students who struggle with mathematics, schools must provide specially designed word-problem intervention. (In this article, we refer to this population as *struggling learners* to mean students with a history of low performance on number, calculations, or word problems. In the studies we consider, researchers used a variety of cut-points to operationalize *low* math performance. If the study's authors indicated their sample had low math performance, we included it even though some cut-points were as high as the 50^{th} percentile.)

The need for word-problem intervention for struggling learners is illustrated in a randomized control trial investigating effects of number knowledge intervention for struggling learners (Fuchs et al. 2013). Intervention produced stronger learning than occurred for struggling learners receiving their typical school program, including simple arithmetic, multi-digit calculations, number knowledge, and word problems. Yet, while the posttest arithmetic achievement gap (between struggling and other classmates) narrowed substantially for struggling learners who received intervention, their posttest word-problem achievement gap increased dramatically.

This widening word-problem achievement gap in the face of a narrowing arithmetic achievement gap is alarming because word-problem solving (WPS) may reflect understanding of and the capacity to apply mathematical ideas in everyday life and in the service of STEM learning (Foreman-Murray and Fuchs 2019). WPS is a strong school-age predictor of employment and wages in adulthood (Every Child a Chance Trust 2009; Parsons and Bynner 1997), and word problems represent a major emphasis in almost every strand of the mathematics curriculum. It is also alarming because the WPS literature provides the basis for only limited understanding about the nature of WPS difficulty or effective intervention practice.

The focus of this article is the well documented association between low working memory capacity (working memory) and difficulty with WPS and how recent studies have addressed this association within WPS intervention. We define *working memory* as a resource-limited capacity that allocates attention and plans, sequences, and maintains information in short-term memory while processing the same or other information (Baddeley and Logie 1999; Unsworth and Engle 2007). Working memory is often operationalized in assessments and training activities as complex working memory span or updating tasks. Complex working memory span tasks combine short-term memory demands with unrelated secondary tasks. For example, the student is presented with a series of sentences, listens to each and immediately decides if it is true or false; at the end of a string of sentences, the student recalls the last word in each sentence. The number of sentences and words to remember (working memory span) gradually increases. Updating requires continuous and simultaneous refreshing of several items. For example, words from a category are presented; after a variable and unpredictable number of updating steps, students name the last word in each category. Studies indicate these tasks index the same general construct of working memory capacity (Schmiedek et al. 2009).

We begin this article by describing a model that specifies how cognitive resources, including working memory, contribute to individual differences in WPS and then summarizing findings on the relation between working memory and WPS. This sets the context for the article's main purpose and major section: to describe the findings of research studies that take one of two approaches for addressing the needs of students with low working memory within word-problem solving intervention. One approach focuses on *compensating* for working memory limitations; the other on *building* working memory capacity. We then suggest the need for research on integrating the two approaches by embedding working memory training within explicit WPS intervention.

1. A Framework for Understanding Mathematical Word-Problem Solving

Kintsch's model of WPS (Kintsch and Greeno 1985) suggests a complex undertaking. The model assumes that general features of the text comprehension process apply across stories, essays, and word-problem statements, but that the comprehension strategies, the nature of the required knowledge structures, and the form of the resulting macrostructures and the situation and problem models differ by text type. Based on theories of text comprehension and discourse processing (van Dijk and Kintsch 1983), the model specifies that word-problem representations have three components. First, the problem solver constructs a coherent microstructure and derives a hierarchical macrostructure to capture the text's essential ideas. Second, the problem solver supplements the text using inferences based on world and mathematical knowledge to build the situation model. The problem solver then identifies the problem model or schema, in which structural relations among quantities are formalized. This schema provides the structure that drives the problem solver's solution strategies.

The Kintsch model posits that building the problem model, via the propositional text structure, inferencing, and schema induction and then applying solution strategies makes strong demands on memory capacity and oral language ability. In this article, we expand the Kintsch model in two ways. First, we reframe short-term memory as working memory. This seems appropriate because WPS involves not only briefly storing information but more pertinently sequentially holding and updating chunks of information in memory as the problem solver processes subsequent segments of the WP statement. This revision to the Kintsch model is grounded in studies showing that working memory is actively involved in WPS (e.g., Anderson 2007; Lee et al. 2004; Raghubar et al. 2010; Swanson et al. 2008; Swanson and Sachse-Lee 2001; also see description of the Peng et al. meta-analysis [in press] at the end of this section).

The second departure from the Kintsch model is our inclusion of reasoning as a key ability involved in WPS. This is based on evidence that individual difference explaining development within representative samples (e.g., Fuchs et al. 2015). It is also supported by the finding that children with stronger reasoning are differentially responsive to WPS intervention (Fuchs et al. 2014).

With respect to demands on working memory and reasoning, consider the following problem-solving process for this second-grade combine word problem (Part 1 plus Part 2 equals Total or P1 + P2 = T): *Cleo has 3 sisters. Her cousin, Fu, has 5 girlfriends and 1 sister. How many sisters do the girls have in all?* A problem solver may process the first sentence's propositional text base to identify that the object is sisters, the quantity is 3, and the actor is Cleo. Cleo's role is to be determined. The problem solver stores this information in working memory. In the second sentence, the problem solver similarly codes the propositions and places them in memory, but *girlfriends* fails to match the object code in the first sentence, signaling 5 as possibly irrelevant. This is added to working memory. In the last sentence (the question), the problem solver is cued by the quantitative proposition *how many sisters* and the phrase *in all* to select the combine schema; assign the role of superset (Total) to the missing quantity; assign subset roles (Part 1 and Part 2) to the to-be-determined entities in working memory; and reject 5 girlfriends as irrelevant. The problem solver translates the numbers associated with the combine schema's slots to frame a number sentence with a missing quantity. Then the problem solver relies on this number sentence to calculate the solution. With typical school instruction, children gradually induce the combine schema as a “problem type” (this is rarely explicitly taught), just as they devise their own strategies for managing the working memory and reasoning demands associated with this problem-solving sequence.

In terms of language comprehension, children enter school understanding generic vocabulary and language constructions. However, as per Kintsch and Greeno (l985) and others (e.g., De Corte and Verschaffel 1985), with arithmetic and word-problem exposure, they learn to treat these words in a math-specific way (e.g., *more* becomes the more complicated construction *more than* involving sets). Cummins et al.’s (1988) computational simulation showed that problem representation depends heavily on language comprehension; De Corte et al. (1985) showed that altering wording in minor ways dramatically affects solution accuracy; Fuchs et al (2018) showed that embedding word-problem language instruction in word-problem intervention offers added value on WPS outcomes over word-problem intervention without embedded language instruction.

To illustrate how WPS depends on and taxes language comprehension, consider this next problem: *Jose, who is a farmer, has 7 chickens. His friend, Thom, has 1 tractor and 6 sheep. 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 stem from more sophisticated representations of vocabulary involving taxonomic relations at superordinate levels and distinctions among categories (chickens + sheep = animals; tractors are not animals).

2. Evidence on the Relation between Working Memory and Word-Problem Solving

This expanded Kintsch model, which departs from the original model by substituting working memory for short-term memory and by adding reasoning as a third cognitive resource, provides this paper’s context for how working memory operates to support WPS. Because working memory is our main focus in this article, we now turn our attention to the empirical evidence on the relation between working memory and WPS by summarizing the findings of a major recent meta-analysis.

Peng et al. (in press) estimated the relation between mathematics and working memory and tested potentially moderating influences on this relation. The potential influences were working memory domain (verbal, numerical, visuospatial), type of math skill (basic number knowledge, whole-number calculations, WPS, fractions, geometry, algebra), and type of mathematics difficulty (math disability with vs. without other co-occurring cognitive weaknesses or other disorders). The motivation for this undertaking was variation in estimates of the magnitude of the relation, ranging from .00 (e.g., Meyer et al., 2010; Tirre & Pena, 1993) to .70 (e.g., Andersson and Lyxell 2007; Passolunghi and Siegel 2004). The major extension over the previous meta-analysis (Friso-van den Bos et al. 2013) was the focus on moderators as a means of explaining such variation.

The authors considered 100 studies with 829 effect sizes, while controlling for age. The overall estimate of the correlation between working memory and mathematics performance was a moderate .35. However, type of math skill and type of mathematics difficulty moderated this relation. With respect to the topic of the present article, the type of math skill with the largest estimate of effect was word-problem solving (r = .37), and this relation was comparable across verbal, numerical, and visuospatial forms of working memory tasks. This finding indicates that WPS taxes domain-general working memory resources that are not tied entirely to language-based processes. It suggests the importance of the central executive for directing attention to relevant information in the WP statement, for coordinating multiple cognitive resources, and for inhibiting irrelevant information. At the same time, the effect of type of mathematics difficulty revealed that the relation between working memory and mathematics was stronger for students who experienced a math disability along with another, complicating cognitive weakness or disorder. This suggests that students with more severe math difficulty lack appropriate strategies for managing the working memory demands involved in WPS.

Some may wonder about the role of math anxiety in math performance and how working memory may affect that relation. In a more recent meta-analysis, centrally focused on math anxiety, Namkung et al. (in press) addressed this question. They found that the relation between math anxiety and math performance was not moderated by working memory capacity. While the estimate of the correlation between math performance and working memory was .31, the correlation between math anxiety and working memory was only .08. The authors noted that the literature that simultaneously attends to math, working memory, and anxiety is small; thus, additional research on this question is warranted. Math anxiety aside, the extant literature clearly documents a role for working memory in WPS, and the effect between working memory and mathematics appears larger for individuals with complex mathematics difficulty simultaneously involving other neurocognitive difficulties. This provides the basis for considering intervention methods to address the working memory limitations of struggling learners.

3. Two Approaches for Addressing Students' Working Memory Limitations within Word-Problem Intervention

3.1 Approach 1: Compensating for Low Working Memory Capacity via Cognitive Strategy Intervention

In this section, we describe an approach, incorporated in three related lines of research, designed to compensate for low working memory capacity via cognitive strategy intervention. With this approach, students learn to conceptualize word problems as belonging to word-problem types; in the Kintsch and Greeno (1985) framework, this is referred to as the *problem model* derived from text processing. Students learn to represent each word-problem type with a diagram or equation that maps onto the word-problem type's central mathematical event. Once students identify the word problem's type, they execute the step-by-step solution strategy used to solve that problem type. This involves mapping the information in the word problem into that word-problem type's diagram or equation and then transforming that representation into a number sentence with a missing quantity to solve for the missing quantity. Jitendra and colleagues (e.g., Jitendra et al. 2011; Jitendra et al. 2009), Fuchs and colleagues (e.g., Fuchs, Zumeta et al. 2010; Fuchs et al. 2014), and Powell and colleagues (2019) refer to this approach as *schema-based instruction* (or word-problem type instruction). Swanson (Swanson et al. 2014, Swanson 2016) refers to this as *cognitive strategy instruction*.

In his 2016 study, Swanson investigated the effects of four cognitive strategy conditions (verbal or visual or verbal + visual emphasis or specific [materials-only] cognitive strategies) or a control group, with 30 classrooms randomly assigned to conditions. Each intervention condition ran 20 sessions over eight weeks, 30 min per session, three times per week in groups of 4-5. In a warm-up activity, children solved addition and subtraction problems with missing information in any of the three slots of equations. During instruction, tutors read or reviewed the relevant condition's strategy rule card. In guided practice, children completed three practice problems and reviewed problems from the instructional phase, as tutors assisted with applying the strategy's steps, finding the correct operation, and identifying key words. Independent practice followed on three problems, each with three parts: a question sentence, a number sentence, and irrelevant information. The number of irrelevant sentences during guided and independent practice gradually increased over sessions from one in lessons 1-7 to five in lessons 18-20, thus gradually increasing working memory demands.

In the verbal condition, steps are find and underline the question, circle numbers, put a square around the key word, cross out irrelevant information, determine the operation, and solve. In the visual condition, two types of diagrams are taught: for the combine problem type showing constituting the whole; for the compare problem type representing quantities being compared. Students enter quantities from the word-problem statement into the selected diagram, with a question mark signifying the missing quantity. The verbal + visual condition integrates the two conditions. These schema-induction methods in combination with the overt problem-solving steps are thought to decrease the working memory demands involved in WPS. The specific strategy condition relies on the same materials as other conditions but without overt problem-solving steps (e.g., underlining, diagramming). Control group students participated in the school math curriculum in their classroom.

On problem-solution accuracy, a significant interaction emerged between intervention condition effects and students' pretest working memory capacity. For students with high working memory, materials-only was superior to the visual (diagram) condition, but the visual, combined, control, and verbal conditions performed comparably. For students with low working memory, the materials-only and visual conditions were comparable to each other, but both were superior to the remaining conditions. Thus, via the specific strategy condition, Swanson (2016) found some support for the idea that cognitive strategies compensate for struggling learners' low working memory capacity, with additional evidence for the visual over the verbal or combined strategies for students with low working memory capacity. In a 2014 analysis, focused on a subset of students with more severe pretest math deficits, Swanson et al. (2014) found a similar pattern of effects.

The programmatic lines of research on schema-based cognitive strategy instruction provide more robust evidence favoring intervention over control group struggling learners at the elementary grades. For example, Jitendra et al. (2007) randomly assigned third-grade struggling learners to schema-based instruction or general strategies instruction in groups of 15-16 students. Reflecting common school word-problem instructional practice, the general strategies instruction condition taught students to use objects, draw a diagram that students thought depicted the word-problem narrative, write a number sentence, and use data from a graph.

Schema-based instruction comprised two phases. The first, schema-induction, helped students to consolidate the mathematical structure associated with three problem types (combine, compare, change) and to represent each structure with the researcher-designed schematic diagram representing that problem type (similar to Swanson's [2016] visual condition). The second phase, problem solution, taught students to solve problems with four steps: find the problem type, organize the information in the problem using the problem type's diagram, plan to solve the problem, and solve the program. The effect size (ES) favoring schema-based instruction I over general strategies instruction was 0.52 on a WPS measure mapping the 1- and 2-step problem types taught during intervention; 0.69 six weeks later on the same measure; and 0.65 on the state achievement test. The moderating effects of pretest working memory capacity have not been tested in this line of research.

The Fuchs et al. research group has also documented learning advantages for schema-based instruction over control, using a similar set of methods tested at the classroom (e.g., Fuchs et al. 2014) and intervention (e.g., Powell et al. 2015) levels. It explicitly teaches children the underlying structure of combine, compare, and change schema. The teacher begins by role playing the problem type's central mathematical event using intact number stories (no missing quantities), concrete objects, and the child's and teacher's names. The teacher then connects the central event to a visual schematic (into which story quantities can be written, similar to Jitendra et al. 2007 and Swanson 2016). However, with the Fuchs's methods, students immediately transition from the visual schematic, which is not ordinarily available to them, to first a hand gesture that quickly permits tutors to unobtrusively remind students of the schema's central mathematical event and then the problem model number sentence (combine: P1 + P2 = T; compare: B – s = D; change: St + C = E or St – C = E). Next, tutors introduce problems (with missing quantities) using role playing, the problem type's schematic and hand gesture, and the problem model number sentence.

Central to this article's focus and in line with Jitendra et al.'s (2007), Powell et al.'s (2019), and Swanson's (2016) intervention methods, schema-based instruction is intended to provide children with strategies that reduce working memory demands. First, when introducing problem types, the teacher makes connections among the situation model, schema, and productive solution strategies explicit. Second, the teacher models step-by-step strategies for identifying problem statements as combine, compare, or change schema and for building the propositional text structure. Children learn to begin with an attack strategy, called *RUN*: Read the problem, Underline the word representing the problem's object code (to anchor the problem's central focus and provide a label for the numerical answer without taxing working memory), and Name the schema by writing the first letter of the problem type to make it available, rather than holding it in working memory (T for total [combine], D for difference [compare], and C for change). Children then write the problem model number sentence and then re-read the problem statement. While re-reading, children replace letters in the problem model number sentence with relevant numerals, cross out irrelevant numerals from the word-problem statement, and insert a blank in the number sentence to signify the missing quantity. Relying on these strategies results in the number sentence for problem solution.

This form of schema-based instruction, designed to address struggling learners' limitations in working memory, includes five units: (a) foundational skills (equal sign as a relational term; efficient counting strategies to add and subtract; methods to solve 2-digit calculation problems; strategies to find a missing quantity when it occurs in any of the three slots of standard addition or subtraction problems; strategies for checking work); (b) combine problems; (c) compare problems; (d) change problems; and (e) review. Across multiple studies, the ES for this form of schema-based instruction over control is 0.68. As with Jitendra's line of work, the moderating effects of students' pretest working memory capacity have not been tested. Clearly, research testing the compensatory hypothesis is warranted.

Across the various intervention research groups, schema-based instruction is viewed as a validated form of WPS intervention. The 2008 U.S. Department of Education's responsiveness-to-intervention practice guide (http://ies.ed.gov/ncee/wwc/pdf/practice_guides/rti_math_pg_042109.pdf) recommended that WPS intervention for struggling learners should include instruction on solving word problems based on common underlying structures, citing eight randomized controlled trials to support the panel's strong level of evidence for this recommendation.

3.2 Approach 2: Building Working Memory Capacity

Like all validated interventions, schema-based instruction does not address the needs of *all* struggling learners. Randomized controlled trials (e.g., Fuchs et al. 2014; Fuchs et al. 2004) suggest a response rate of 70%. One viable approach for expanding schema-based instruction's framework is to *build* working memory capacity while *compensating* for students' low working memory capacity via schema-based instruction.

Good problem solvers have greater working memory capacity than do poor problem solvers, and individual differences in working memory account for variance in WPS when controlling for other cognitive resources (e.g., LeBlanc and Weber-Russell 1996; Passolunghi and Siegel 2004; Swanson & Sachse-Lee 2001). Moreover, Barnes et al.'s (2014) longitudinal mediation modeling with young children with and without spina bifida myelomeningocele, a neurodevelopmental disorder associated with mathematics difficulty, found that visuospatial working memory at 36 months fully mediated the effect of group (with vs. without the disorder) on WPS at 8-9 years. As mentioned earlier, Peng et al.'s (in press) meta-analysis showed that the effect of working memory on WPS is higher than for other types of math. Such observational evidence provides the basis for hypothesizing that improving working memory will strengthen WPS.

The research base testing this hypothesis is, however, inconsistent and controversial. According to a 2013 meta-analysis (Melby-Lervag and Hulme), working memory training produces reliable short-term improvements in working memory. However, near-transfer effects do not generally sustain at follow-up, and transfer to academic skills is limited. On measures of mathematics calculations, the ES approximated 0.25 *standard deviations* for the subset of studies *not* targeting students with IQ between 55 - 85 (it approached zero without this exclusion). In this meta-analysis, no studies indexing transfer to math WPS were identified. (Sala and Gobet's 2017 meta-analysis, which only considered effects on typical learners, also did not identify studies that disaggregated effects on WPS.)

The issue of transfer to academic outcomes is critical for the population of students with histories of poor mathematics learning. This population is especially vulnerable to transfer difficulty, because these students fail to recognize novel stimuli as related to tasks on which they have received instruction (e.g., Haskell 2001; National Research Council 2000). Randomized controlled trials on schema-based instruction show that transfer distance (i.e., the alignment of outcome measures with the training's content) has a more deleterious effect on WPS outcomes for struggling learners than for typically developing learners (e.g., Fuchs et al. 2008). In working memory training, struggling learners may suffer serious transfer problems because they do not have the reservoir of mathematics skill onto which enhanced working memory capacity, derived from working memory training, may be applied.

We next consider results of six randomized controlled trials, conducted since publication of Melby-Lervag and Hulme's (2013) meta-analysis, examining working memory training effects on mathematics outcomes of students with mathematics difficulties. (We did not consider studies targeting students with attention deficit hyperactivity.) Three of the six studies examined working memory training's effects against one or more control group, while assessing transfer to mathematics calculations, without any attention to mathematics within working memory training or as a supplement to working memory training.

Partanen et al. (2015) included primary-grade special education students performing below the 10^{th} percentile on forward or backward digit span, arithmetic, reading, and writing. Students were randomly assigned, within school, to CogMed (a computerized adaptive program with practice on simple span tasks) or a non-active control group. In five schools, students received CogMed alone; in another five schools, CogMed was used in conjunction with metacognitive strategies. They found no significant effects on working memory (auditory and visuospatial backward digit span and spatial span) and no significant transfer effect to calculations.

Ang et al. (2015), by contrast, obtained some support for enhancing working memory, but again without transfer to calculations. They included 7-year-olds with working memory scores below the 25^{th} percentile; teachers had identified 23% of these children as performing poorly in mathematics, and the majority received school-delivered supplemental math support. Students were randomly assigned to four conditions: an author-developed computerized program (updating practice), CogMed, active control (same games as the author-developed program but without updating), and non-active control. Working memory training had a significant effect on working memory tasks similar to those trained at 6-month follow-up but not immediate posttest (1-2 weeks after training ended); CogMed produced significant effects only at immediate posttest. Neither form of working memory training produced transfer to untrained working memory tasks or addition or subtraction fluency.

Zhan et al. (2018) randomly assigned students with learning difficulties in reading and math to working memory training on an updating task or non-active control. The process for identifying learning difficulties was unclear; we infer students with Chinese language and math scores below the 25^{th} percentile but with their fluid intelligence scores above the 49^{th} percentile.

Working memory training outperformed control on a measure of updating accuracy (not on response time), and effects transferred to performance on a broad-based mathematics achievement test. On this test, 60% of items were pure calculations; the others assessed some form of problem solving. We assume some were word problems, but only total test score was analyzed. Also, working memory training was also associated with changes in brain activity indicating improved cognitive control and working memory updating capacity.

Although these studies differed in a variety of ways, we observe that severity of students' learning problems appears associated with working memory training's efficacy. In Partanen et al. (2015), whose sample performed below the 10th percentile in multiple areas, working memory training did not improve working memory or math. Ang et al.'s (2015) sample performed below the 25th percentile on working memory but with only 23% scoring low in math. Here, effects on near-transfer working memory were evident, but without transfer to math or other working memory measures. In Zhan et al. (2018), where we infer participants scored below the 25th percentile in mathematics but above the 49th percentile on reasoning, effects on working memory and transfer to mathematics were realized.

Yet, across the next three studies, results were modest even though math difficulty was defined leniently: below the 50th percentile. What distinguishes this set of studies from the previous set is that each included features designed to strengthen working memory training's capacity for improving mathematics. Nelwan and colleagues addressed the transfer challenge raised above: that transfer from working memory training to math may be especially difficult for struggling learners because they lack a reservoir of mathematics skill. Nelwan and Kroesbergen (2016) investigated effects of a commercial working memory training program in 9-12-year-olds with attention and mathematics difficulties as rated by teachers and with scores below the 50th percentile on math tests. They randomly assigned students to three conditions, each comprising two 8-week periods: Working memory training followed by adaptive computerized arithmetic training, this math training followed by working memory training, and non-active control followed by this math training. Although short-term effects on verbal updating were revealed, none was found on short- or long-term visuospatial updating. Moreover, there was little evidence of transfer to fluency on the four operations and no evidence that working memory training or math training improved number sense. The authors speculated that working memory training, although implemented as the program developers intended, may have been compromised by inadequate guidance and feedback to students.

Nelwan et al. (2018) pursued this idea with children identified for participation in the same way, trained with the same programs, and measured with the same assessments as the 2016 study. Children were randomly assigned to working memory training followed by math training or non-active control followed by math training. In working memory training, however, research staff closely monitored each student's progress and provided weekly feedback designed to overcome difficulties students experienced (e.g., verbalize visual stimuli or subvocally repeat solutions to arithmetic problems). Analyses incorporated students in Nelwan and Kroesbergen's (2016) low-coaching group as a second contrast condition. Only weak support was found for the idea that high-coaching working memory training improves working memory outcomes more than low-coaching working memory training or control. "Minimal" effects were documented on the first short-term visuospatial working memory outcome, but without effects on verbal working memory. Although math performance across the first training period increased more in highly-trained working memory training students than in other conditions, working memory training did not build working memory capacity in ways that differentially benefited calculations learning during the second period's math training.

Kroesbergen et al. (2014) evaluated a different approach for promoting transfer to math by testing effects of two versions of working memory training: Exercises were constructed in highly similar ways except that one version relied on numerical stimuli; the other on domain-general stimuli. Five-year old children with mathematics performance below the 50th percentile were randomly assigned to three groups: two active conditions and a non-active control group. There were no significant differences among conditions on phonological working memory, but the two working memory training conditions (which performed comparably) scored significantly higher than control on visuospatial working memory. In terms of transfer, both working memory training conditions outperformed control on nonsymbolic quantity discrimination, which involves visuospatial perception. More interestingly, on counting skills, domain-specific working memory training but not domain-general working memory training outperformed control (the two working memory training conditions performed comparably). Results suggest that domain-specific working memory training improves transfer to math performance, although this effect was revealed with a lenient criterion for mathematics difficulty.

4. Integrating Schema-Based Instruction with Domain-Specific Working Memory Training

These six studies in combination with Melby-Lervag and Hulme's meta-analysis (2013) and other analyses of the working memory training literature (Hulme and Melby-Lervag 2012; Sala and Gobet 2017; Shipstead et al. 2012) reveal only mixed support for working memory training's effects on working memory capacity and with only limited transfer to mathematics skill. They also reveal notable gaps in the literature. Few studies have assessed working memory improvement as a mediator of working memory training transfer; none has investigated WORKING MEMORY improvement as a mediator of transfer to math outcomes in students with math difficulties. We also identified no prior studies assessing working memory training's transfer effects to WPS. These gaps in the extant literature should be addressed in future research.

Further, across the six recent studies focused on students with mathematics difficulties, effects appear more encouraging when studies apply more lenient cut-points for identifying samples of students with mathematics difficulties. Studies are needed that focus on students with substantial deficits because school intervention in many countries is reserved for such students.

At the same time, no studies have evaluated working memory training's added value over a *standard-of-practice* condition, a point raised by Kroesbergen et al. (2014). Given intervention costs along with limited school time and complicated schedules for providing intervention, innovative interventions must prove their value against standard-of-practice programs. Otherwise, validated interventions, involving explicit instruction, are preferred.

Thus, need exists for studies examining working memory training's transfer effects to WPS on students who begin intervention with clear math deficits, that assess WORKING MEMORY improvement as a mediator of working memory training's transfer effects to WPS, and that include a standard-of-practice contrast condition. In considering potentially productive directions for enhancing working memory training's transfer effects to math, Nelwan et al. (2018) provides rationale for supplementing working memory training with math training during or immediately following working memory training. Kroesbergen et al. (2014) provides rationale for conducting working memory training with domain-specific tasks to prime struggling learners to transfer enhanced working memory capacity to math.

This leads us to suggest the following directions for future research. First, randomized control trials are needed to contrast domain-general working memory training to standard-of-care WPS interventions for students with meaningful mathematics deficits, at below the 25th percentile. Second, innovative forms of validated WPS, perhaps schema-based instruction, might be extended by embedding domain-general or domain-specific working memory training within the validated WPS program. This might involve complex span games conducted adaptively by live tutors on calculations and word-problem stimuli. (For example, in a 2-span task, the tutor reads the first problem aloud: "There were 6 donuts in the box. Then, Chris ate 1. How many donuts are left?" The student says, "Donuts, 6 minus 1 equals five." The tester then reads the second problem out loud: “Blair read 4 books. Then, she read 2 more. How many books has Blair read now?” The student says, “Books, 4 plus 2 equals 6.” Then, the student recalls: “5, 6.”) Such innovation might thereby succeed in developing WPS even as it builds working memory. Testing such an embedded approach in the same study that incorporates a domain-general working memory condition and a standard-of-care WPS intervention might deepen understanding of working memory training’s potential in valuable ways.

5. **Summary**

WPS is a complex form of mathematical cognition, which makes strong demands on cognitive resources. Standard-of-practice instructional classroom programs and interventions for students with mathematics difficulties include strategies for helping students *compensate* for students’ cognitive processing limitations, including working memory. By contrast, working memory training is designed to *build* working memory capacity. Yet, standard-of-practice programs designed to compensate for working memory limitations fail to meet the needs of all struggling learners, while working memory training has yet to demonstrate convincing transfer to math outcomes for the struggling population. One approach for extending intervention methods to address the needs of struggling learners is to merge validated explicit instructional WPS programs with cognitive training.

The hope explored in this article is that integrating schema-based instruction with domain-specific versions of working memory training may create synergy for promoting stronger WPS learning in struggling learners. By building capacity for *future* learning through working memory training, while increasing mathematics performance on the school’s curriculum, such a merger may at least in part address intervention fade-out effects over time (e.g., Bailey et al. in press; Clarke et al. 2016; Smith et al. 2013). That is, interventions that simultaneously forge more solid foundational mathematics skills while building capacity to support future learning create synergy by which struggling learners profit from the general education’s program in future grades.

Acknowledgement

This research was supported by Grant R305A150200 from the U.S. Department of Education's Institute of Education Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences or the U.S. Department of Education.

References

Andersson, U., & Lyxell, B. (2007). Working memory deficit in children with mathematical difficulties: A general or specific deficit? *Journal of Experimental Child Psychology, 96,* 197–228. DOI: 10.1016/j.jecp.2006.10.001

Ang, S. Y., Lee, K., Cheam, F., Poon, R., & Koh, J. (2015). Updating and working memory training: Immediate improvement, long-term maintenance, and generalizability to non-trained tasks. Journal of *Applied Research in Memory and Cognition, 4*(2), 121-128. DOI: 10.1016/j.jarmac.2015.03.001

Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.). *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge, U.K.: Cambridge University Press.

Bailey, D., Fuchs, L. S., Gilbert, J. K., Geary, D. C., & Fuchs, D. (in press). Prevention: Necessary but insufficient? A two-year follow-up of effective first-grade mathematics intervention. *Child Development*. DOI: 10.1111/cdev.13175

Barnes, M. A., Raghubar, K. P., English, L., Williams, J. M., Taylor, H., & Landry, S. (2014). Longitudinal mediators of achievement in mathematics and reading in typical and atypical development. *Journal of Experimental Child Psychology, 119*, 1-16. DOI: 10.1016/j.jecp.2013.09.006.

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

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

De Corte, E., & Verschaffel, L. (1985). Beginning first graders' initital representations of arithmetic word problems. *The Journal of Mathematical Behavior, 4,* 3-21. DOI: 10.1007/BF01323046

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

Every Child a Chance Trust (2009). *The long-term costs of numeracy difficulties*. Retrieved. August 14, 2009, http://www.everychildachancetrust.org/counts/index.cfm.

Foreman-Murray, L., & Fuchs, L. S. (2019). Quality of explanation as an indicator of fraction magnitude understanding. *Journal of Learning Disabilities, 52*, 181-191. DOI: 10.1177/0022219418775120

Friso-van den Bos, I., Kroesbergen, E. H., & van Luit, J. E. (2014). Number sense in kindergarten children: Factor structure and working memory predictors. *Learning and Individual Differences, 33,* 23–29. DOI: 10.1016/j.lindif.2014.05.003

Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. (2015). Is word-problem solving a form of text comprehension? *Scientific Studies of Reading*, 19, 204–223. DOI: 10.1080/10888438.2015.1005745

Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology, 100*, 491-509. DOI: 10.1037/0022-0663.100.3.491

Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal of Educational Psychology, 96*, 635-647. DOI: 10.1037/0022-0663.96.4.635

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., Bryant, J. D., & 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. DOI: 10.1037/a0020662

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., DeSelms, J., Seethaler, P. M., Wilson, J., Craddock, C. F., Bryant, J. D., Luther, K., & Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*, 58-77. DOI: 10.1037/t05529-000

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. DOI: 10.1080/10888438.2017.1398259.

Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlett, C. L., Fuchs, D., Compton, D. 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. DOI: 10.1037/a0036793

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. DOI: 10.1086/651191

Haskell, R. E. (2001). *Transfer of learning: Cognition, instruction, and reasoning*. Academic Press.

Hulme, C., & Melby-Lervag, M. (2012). Current evidence does not support the claims made for CogMed working memory training. *Journal of Applied Research in Memory and Cognition, 1*, 197-200. DOI: 10.1016/j.jarmac.2012.06.006

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematics problem solving. *Journal of Educational Psychology, 99,* 115-127. DOI: 10.1037/0022-0663.99.1.115

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction, 21,* 731-745. DOI:10.1016/j.learninstruc.2011.04.002 732

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., Hughes, C. L., & Mack, T. R. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology, 34*, 250-264. DOI: 10.1016/j.cedpsych.2009.06.001

Kintsch, W., & Greeno, J. G. (l985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129. DOI: 10.1037/0033-295X.92.1.109

Koedinger, K. R., & Nathan, M. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences, 113*, 129-164. DOI: 10.1207/s15327809jls1302_1

Kroesbergen, E. H., van't Noordende, J. E., & Kolkman, M. E., (2014). Training working memory in kindergarten children: Effects on working memory and early numeracy. *Journal of Child Neuropsychology, 20*, 23-37. DOI: 10.1080/09297049.2012.736483

LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematics connections: A computer model of arithmetic work problem solving. *Cognitive Science, 20*, 357-407. DOI*:* 10.1016/S0364-0213(99)80010-X

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. *DOI*: 10.1016/j.jecp.2004.07.001

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

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., & Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences, 20,* 101–109. DOI: 10.1016/j.lindif.2009.08.004

Namkung, J. M., Peng, P., & Lin, X. (in press). The relation between mathematics anxiety and mathematics performance among school-aged students: A meta-analysis. *Review of Educational Research*. DOI: 10.3102/0034654319843494

National Research Council. (2000). *How people learn: Brain, mind, experience, and school*. The National Academies Press. DOI: 10.17226/9853

Nelwan, M., Vissers, C., & Kroesbergen, E. H. (2018). Coaching positively influences the effects of working memory training on visual working memory as well as mathematical ability. *Neuropsychologia, 113,* 140-149. DOI: 10.1016/j.neuropsychologia.2018.04.002

Nelwan, M., & Kroesbergen, E. H. (2016). Limited near and far transfer effects of Jungle Memory working memory training on learning mathematics in children with attentional and mathematics difficulties. *Frontiers in Psychology*. DOI: 10.3389/fpsyg.2016.01384

Parsons, S., & Bynner, J. (1997). Numeracy and employment. *Education and Training, 39,* 43-51. DOI: 10.1108/00400919710164125

Partanen, P. Jansson, B., Lisspers, J., & Sundin, O. (2015). Metacognitive strategy training adds to the effects of working memory training in children with special educational needs. *International Journal of Psychological Studies*, 7, 130-140. DOI: 10.5539/ijps.v7n3p130

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology 88,* 348-367. DOI: 10.1016/j.jecp.2004.04.002

Peng, P., Namkung, J., Barnes, M., & Sun, C. (in press). 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*. Advance online publication. DOI: 10.1037/edu0000079

Powell, S. R., Berry, K. A., & Barnes, M. A. (2019). *The role of algebraic reasoning with a word-problem intervention for third-grade students with mathematics difficulty*. Manuscript submitted for review.

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. DOI: 10.1016/j.lindif.2009.10.005

Sala, G., & Gobet, F. (2017). Working memory training in typically-developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*, 671-685. DOI: 10.1023/A:1011117103381

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1089-1096. DOI: 10.1037/a0015730

Shipstead, Z., Hicks, K. L. W., & Engle, R. (2012). CogMed working memory training: Does the evidence support the claims? *Journal of Applied Research in Memory and Cognition, 1*, 185-193. DOI: 10.1016/j.jarmac.2012.06.003

Smith, T., Cobb, P., Farran, D. C., Cordray, D. S., & Munter, C. (2013). Evaluating Math Recovery: Assessing the causal impact of a diagnostic tutoring program on student achievement. *American Educational Research Journal*, 50, 397-428. DOI: 10.3102/0002831212469045

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. DOI: 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. DOI: 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. DOI: 10.1037/0022-0663.100.2.343. 343

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. DOI: 10.1177/0731948713507264

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. DOI: 10.1006/jecp.2000.2587

Tirre, W. C., & Pena, C. M. (1993). Components of quantitative reasoning: General and group ability factors. *Intelligence, 17,* 501–521. DOI: 10.1016/0160-2896(93)90015-W

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104-132. DOI: 10.1037/0033-295X.114.1.104

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press. DOI: 10.3758/s13421-019-00926-4

Zhan, H., Chang, L., Chen, X., Ma, L., & Zhou (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*. DOI: 10.3389/fnhum.2018.00154