Exceptional Children

Vol. 80, No. 2, pp. 149-168.

The Effects of Mathematics Strategy Instruction for Children With Serious Problem-Solving Difficulties

H. LEE SWANSON

MICHAEL J. OROSCO

CATHY M. LUSSIER

University of California, Riverside

ABSTRACT: *This study investigated the role of strategy instruction on solution accuracy in children with and without serious math difficulties (MD) in problem solving. Children's posttest solution accuracy was compared on standardized and experimental measures as a function of strategy conditions. Strategy conditions included curriculum materials that gradually increased the number of irrelevant propositions within word problems. Children in Grade 3 (*N *= 193) were randomly assigned to one of five conditions: materials + verbal strategies (e.g., underlining the question), materials + verbal + visual strategies, materials + visual strategies (e.g., correctly placing numbers in diagrams), materials only-no overt strategies, and an untreated control. Compared to children with MD in the control condition, posttest outcomes for children with MD on standardized measures improved significantly under verbal + visual conditions, whereas posttest scores on the experimental problem-solving measures improved under the materials-only condition. Those strategy conditions found least effective made substantial demands on children's working memory capacity. The authors discuss benefits and limitations of strategy instruction.*

Recent intervention studies directed towards improving problem-solving accuracy in children with serious math difficulties (MD) have found support for teaching cognitive strategies. Several studies have found that verbal strategy instruction (e.g., Montague, 2008; Montague, Warger, & Morgan, 2000; Xin, 2008) and visual-spatial strategies (e.g., Kolloffel, Eysink, de Jong, & Wilhelm, 2009; van Garderen, 2007) enhance children's math performance relative to control conditions (see Baker, Gersten, & Lee, 2002; Gersten et al., 2009 for reviews). Well-designed intervention studies (randomized clinical trials) have focused on high-risk samples with MD. For example, Jitendra and colleagues (1998) used a visual categorization method to cluster arithmetic word problems (e.g., change, compare) that significantly improved problem-solving accuracy as compared to the control condition (effect size = 0.45). Likewise, in a randomized control group design, Fuchs et al. (2003) taught problem-solving methods to children with MD and found that cognitive strategies (schema-based instructions) improved problem-solving accuracy (effect sizes ranged 1.16–1.18 depending on the transfer measure). Additional successful strategy models have included diagramming (van Garderen, 2007), identifying key words (e.g., Mastropieri, Scruggs, & Shiah, 1997), and meta-cognitive strategies (e.g., Montague, 1997; see Gersten et al., 2009; Xin & Jitendra, 1999, for reviews).

The majority of studies validating the use of cognitive strategies suggest that such training facilitates problem-solving accuracy because it reduces or compensates for the demands placed on the limited cognitive processes of children with MD. Despite the overall benefits of strategy instruction in remediating word problem-solving difficulties, the use of strategies for some children with MD may not always be advantageous. Because children with MD experience working memory difficulties (e.g., Fuchs et al., 2010; Swanson, Jerman, & Zheng, 2008), their poor problem-solving skills plus their low working memory may have direct consequences on the effectiveness of cognitive strategy interventions. Because strategies make demands on mental resources, we hypothesized that the availability of ample working memory resources is an important precondition for strategy training to be successful. Children with relatively smaller working memory capacity (WMC) may be easily overtaxed by certain strategies, which may lead to poor learning outcomes even after training. This hypothesis is in line with cognitive load theory (e.g., Sweller, 2005) whose central tenet is that instruction should be designed in alignment with the learner's cognitive architecture, which consists of a limited-capacity working memory system.

In the present study, we assessed whether strategy instruction that teaches children to focus on the relevant propositional structures of word problems improves solution accuracy. The interventions provided explicit instruction related to (a) verbal strategies that directed children to identify (e.g., via underling and circling) relevant or key propositions within the problems, (b) visual strategies that required children to place numbers into diagrams, or (c) a strategy condition that combined both verbal and visual strategies. Also, because warm-up activities related to calculation have been found effective in problem-solving interventions, this component was also included in all strategy training sessions (e.g., Fuchs et al., 2003). In addition, consistent with literature reviews that have identified key components related to treatment effectiveness (Gersten et al., 2009; Xin & Jitendra, 1999), each strategy training session involved explicit practice and feedback related to strategy use and performance.

Our study was unique in that we directed children's attention to the relevant propositions within the context of interference (irrelevant propositions). That is, instructions were embedded within each lesson that directed children's attention to relevant propositions within word problems while concurrently increasing the number of irrelevant propositions. Our rationale for this activity was as follows. First, children with MD have difficulties with controlled attention (inhibiting irrelevant information; see Censabella & Noël, 2008; Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002; Ng & Lee, 2009; Passolunghi, Cornoldi, & De Liberto, 2001; Passolunghi & Siegel, 2001); therefore, strategies must be directly taught to help such children attend to relevant information within the context of irrelevant information. Second, previous research from national assessments has indicated that elementary children find it extremely difficult to discriminate relevant from irrelevant facts in word problems (see Cook & Rieser, 2005, for a review); this discrimination problem would be especially acute in children with MD. Third, differentiating information that is relevant to a given task from irrelevant information is a fundamental property of learning. The mechanisms that underlie such differentiation play a key role in mental problem representation (e.g., Kintsch & Greeno, 1985). Thus, merely providing problems with only relevant propositions for solution (as is done in the majority of school curriculum materials) does not teach children with MD to discriminate the relevant parts of a problem. Previous work shows children with MD engage in *number grabbing*, selecting numbers from text without regard to the number's relationship to the problem's meaning (see Cook & Riser, 2005, for discussion of this strategy). To avoid number grabbing, children must be directed to consider relevant information in the context of irrelevant information (i.e., competing semantic and numerical information). Fourth, word problem solving is significantly predicted by text comprehension (e.g., Swanson, Cooney, & Brock 1993), and to create an accurate mental model of text comprehension both relevant and incidental propositions must be considered. Finally, training that includes gradual increases in competing information within the context relevant information has been suggested as one means to improve controlled attention in working memory (e.g., Holmes, Gathercole & Dunning, 2009). Our approach is also based on research showing that a key mechanism that underlies WMC is *controlled attention*, an individual's ability to access and process relevant information in the context of interfering information (Engle, Tuholski, Laughlin, & Conway, 1999; Unsworth, 2010; Unsworth & Engle, 2007).

Elementary children find it extremely difficult to discriminate relevant from irrelevant facts in word problems.

To this end, children with MD and without MD were randomly assigned to one of four treatment groups: verbal strategies, visual-spatial strategies, combination of both verbal and visual-spatial strategies, and the curriculum without explicit strategy instruction (materials-only). The materials-only condition was intended to test whether training related to increasing irrelevant propositions within the instructional materials had a unique contribution to solution accuracy independent of explicit strategy instruction. Two questions directed this exploratory study:

1. Are some cognitive strategies more effective than others for children with MD?
2. Do cognitive strategies place different demands on cognitive processes in children with MD?

Regarding Question 1, although several strategy conditions may improve solution accuracy relative to the control condition, some strategies may play a more important role for children with MD to catch up to their average-achieving peers. This may occur for several reasons. One possibility that we explored is that training in the allocation of attention within materials can positively influence treatment outcomes. Practice in controlled attention in the present study was achieved by increasing the number of irrelevant propositions during instruction. The irrelevant propositions were simple in the early stages of training (one or two irrelevant sentences) and were increased at fixed levels across later intervention sessions (three to eight irrelevant sentences). However, these general effects may have also influenced whether some specifically taught strategies were more effective than others. For example, we explored whether a combination of both verbal and visual-spatial strategies might be beneficial for enhancing problem-solving accuracy relative to strategy conditions that focus on verbal or visual-spatial strategies in isolation. The underlying assumption is that because the combined strategy draws upon separate verbal and visual-spatial storage capacities, the combination of these storage systems opens up the possibility for more information to be processed (Mayer, 2005).

Regarding Question 2, strategies have different demands on children's cognitive abilities. This study explored whether the use of some strategies compensates for the excessive demands placed on the cognitive processes in children with MD. Thus, we explored the correlations between working memory and problem-solving performance within the different strategy conditions, investigating four hypotheses. One hypothesis tested whether correlations between cognitive processes and problem solving under strategy condition are nonsignificant. The basic skills model suggests that if declarative knowledge is intact (e.g., reading comprehension, computational knowledge), strategy instruction provides helpful information (possibly redundant or elaborative information) to solve word problems without making demands on cognitive processes. This model suggests that cognitive instruction provides additional information over control conditions when basic skills (e.g., calculation, reading) are intact. In contrast, the general resource model hypothesis suggests that individual differences in cognitive processes are related to solution accuracy, regardless of treatment conditions. The resource model predicts that because cognitive processes act as a general system that underlies several problem-solving tasks, it has a general (non-treatment-specific) effect on problem-solving outcomes. Another hypothesis, based on the high cognitive skills model, suggests that the children with MD that benefit most from strategy conditions are those with higher working memory skills. Thus, significant positive correlations between WMC and problem solving are more likely to occur in strategy conditions than in control conditions. Finally, a compensatory hypothesis views cognitive training as a method that helps reduce processing demands on children's problem solving by freeing additional mental resources to solve problems. The compensatory model predicts that children with lower cognitive abilities (e.g., WMC) are more likely to place a greater reliance on strategy conditions than those with relatively higher abilities. Thus, a negative correlation between cognitive processes and problem-solving accuracy occurs within strategy conditions.

METHOD

Participants

One hundred and ninety-three children from Grade 3 in a Southern California public school district participated in this study. The research included institutional approval (HS-06-099). The final selection was based on parent approval for participation and math achievement scores. Of the 193 children selected, 83 were females and 110 were males. The ethnic representation of the sample was 116 Anglo, 30 Hispanic, 13 African American, 12 Asian, and 22 mixed or other (e.g. Anglo and Hispanic, Native American). The mean socioeconomic status (SES) of the sample was primarily low SES to middle SES based on free lunch participation, parent education, or occupation. However, the sample varied from lower middle class to upper middle class.

Definition of Risk for MD

We were able to identify children at risk for persistent difficulties in problem-solving performance across 2 years of classroom performance (Grades 2 and 3, in this case). However, because the majority of children in our sample were not diagnosed with specific learning disabilities in math problem solving, we utilized the term *at risk for MD in problem solving* to operationally define our sample. It is important to note that the focus of this study was on word problem-solving deficits and not on calculation per se, and therefore we identified children who performed in the lower 25th percentile on norm-referenced word (story) problem-solving math tests and above the 25th percentile in calculation. We also sought to identify children whose skills in other areas (e.g., fluid intelligence, computation, reading comprehension) were within the normal range. Our attempts at classification are consistent with current federal categories of learning disabilities that distinguish between calculation and mathematical reasoning (34 C.F.R. § 300.541[a][2]); we assumed that word problem solving would capture mathematical reasoning processes. The 25th percentile cut-off score on standardized achievement measures has been commonly used to identify children at risk (e.g., Fletcher et al., 1989; Siegel & Ryan, 1989). We also administered tests to determine if the groups differed on cognitive processing measures. Because research has shown consistent differences between children with and without MD on the executive component of working memory (Swanson & Beebe-Frankenberger, 2004; Swanson et al., 2008), we administered three working-memory tasks (discussed below) and compared the composite scores (see Table 1). We also computed measures of visual-spatial working memory. These latter measures were administered because we assumed that two of our treatment conditions (i.e., visual-only, combined verbal + visual) would draw heavily on visual-spatial skills and we wanted to establish comparable abilities between the MD and non-MD (NMD) risk group.

Overall, our classification procedure separated the sample into 73 children with MD (38 females) and 120 children without MD (51 females). Table 1 shows the means and standard

TABLE 1

Comparison of Children With and Without Math Difficulties in Problem Solving at Pretest

	MD		*NMD*			
Measure	*Mean*	SD	*Mean*	SD	ES	F *ratio*
Age (months)	103.53	5.98	104.41	5.48	–0.18	.88
Classification						
Raven[a]	99.65	8.46	105.58	9.81	–0.67	18.32**
CMAT[b]	7.27	1.30	12.06	1.43	–3.43	535.32***
WRAT-3[a]	99.09	7.63	102.65	9.02	–0.32	7.93*
TORC[b]	9.38	2.07	10.29	1.89	–0.36	9.74**
WMC						
Executive WMC[c]	–0.28	0.44	0.15	0.74	–0.61	20.81**
Visual WMC[c]	–0.07	0.70	0.007	0.72	–0.16	.53
Pretest criterion measures						
TOMA-2[c]	–0.43	0.83	0.34	0.98	–0.85	42.12***
STAR[c]	–0.30	0.98	0.14	0.93	–0.42	9.78**
WRAT-3[c]	–0.59	0.91	0.09	0.91	–0.75	42.01***

Note. MD = students with math difficulties; NMD = students without math difficulties; ES = effect size between ability groups; Raven = Raven Colored Matrices Test (Raven, 1976); CMAT = story subtest from the Comprehensive Mathematical Abilities Test (Hresko et al., 2003); WRAT-3 = arithmetic subtest from Wide Range Achievement Test (Wilkinson, 1993); TORC = Test of Reading Comprehension (Brown, Hammill, & Weiderholt, 1995); WMC = working memory capacity; TOMA-2 = Test of Mathematical Abilities (Brown, Cronin, & McIntire, 1994); STAR = California Department of Education Standardized Testing and Reporting sample questions (2009).
[a]Standard score. [b]Scale score. [c]z-score.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

deviations for children with and without MD. Children with MD in problem-solving normed scores were in the average range for computation and fluid intelligence. Children with MD were inferior to children without MD on the executive WMC measure, but not visual-spatial WMC.

Obviously, the 25th percentile as a cut-off for identifying problem-solving deficits is arbitrary and there is no reason to assume that children in the 26th percentile and above would perform differently. Thus, our procedure to identify children at risk for MD needs justification. We compared two cut-off points (standard score of 90 and 85, or scale score of 8 or 7) from a composite score of two norm-referenced story problem subtests. We used a measure referred to as "affected-status agreement" (Waesche, Schatschneider, Maner, Ahmed, & Wagner, 2011), which in this case is the proportion of children classified as at risk by either a cut-off score of the 25th percentile (scale score of 8) or 16th percentile (scale score of 7) or both. Using both cut-off scores, the number of children identified as at risk (from a total $N = 193$ that included both risk and nonrisk samples) was 56. Using a cut-off score of 8, the number of additional children identified as at risk was 17. The affected status agreement was .77 (56/56 + 17 + 0). We computed the standard error (.06; see Waesche et al., 2011, p. 300, for formula) and determined the 95% confidence interval (.06 × 1.96). This yielded an affected status that ranged from .65 to .89. Because our status score was greater than chance (confidence intervals did not contain zero), we assumed a standard score of 90 (scale score of 8) was an appropriate cut-off score to infer that children were at risk.

In summary, our criteria for defining children at risk for MD in problem solving was a score between the 35th and 90th percentile on measures of fluid intelligence (Raven Colored Progressive Matrices Test; Raven, 1976), reading comprehension (Test of Reading Comprehension, TORC; Brown, Hammill, & Weiderholt, 1995) and calculation (subtest from the Wide Range Achievement Test-3, WRAT-3; Wilkinson, 1993), and a score at or below the 25th percentile (below a standard score of 90 or scale score of 8) on the Story Problem Solving subtest from the Comprehensive Mathematical Abilities Test (CMAT; Hresko, Schlieve, Herron, Swain, & Sherbenou, 2003). The technical manual for the CMAT subtest reported adequate reliabilities (> .86) and moderate correlations (*r*s > .50) with other normed math tests (e.g., the Stanford Diagnostic Mathematics Test). We also tested children not at risk for MD, and compared this average achieving contrast group to the MD treatment groups to determine if residual problem-solving accuracy differences at pretest were significantly reduced at posttest.

Performance on standardized measures of word problem-solving accuracy for the MD sample was at or below the 25th percentile (scale score at or below 8, standard score below 90; see Table 1), whereas their norm-referenced scores on calculation, reading comprehension, and fluid intelligence were above the 35th percentile. No significant differences emerged between children with and without MD as a function of ethnicity, χ^2 (df = 4, N = 193) = 2.65, p > .05; gender, χ^2 (df = 1, N = 193) = 1.66, p > .05; or chronological age, F (1,193) = . 95, p > .05.

Tasks and Materials

The battery of group and individually administered tasks is described below. Experimental tasks are described in more detail than published and standardized tasks. Tasks were divided into classification, pretest-only (moderator measures), and pretest/posttest measures. The sample reliabilities for each measure varied from .80 to .98.

Classification Measures. Prior to intervention, we administered norm-referenced measures to assess fluid intelligence (Raven Colored Progressive Matrices; Raven, 1976), calculation (WRAT-3; Wilkinson, 1993), reading comprehension (TORC; Brown et al., 1995), and story problems (CMAT; Hresko et al., 2003). We also administered three working memory measures from a normative measure (S-Cognitive Processing Test, S-CPT; Swanson, 1995) that captured executive processing (i.e., Conceptual Span, Updating, Digit/Sentence) and two that captured visual-spatial working memory (i.e., Mapping & Directions, Visual Matrix). The intercorrelations among the processing measures exceeded .45 and the reported Cronbach alpha reliabilities are greater than .88 (Swanson & Beebe-Frankenberger, 2004).

Pretest and Posttest Measures. We administered alternate forms of story problems from the Test of Mathematical Abilities (TOMA-2; Brown, Cronin, & McIntire, 1994) and WRAT-3 (Wilkinson, 1993) at pretest and posttest. The alternate forms were counterbalanced for presentation order. We investigated calculation accuracy (WRAT-3) because it was a required activity for all sessions and therefore a logical measure to assess possible transfer effects.

Experimental Measure: STAR Test of Word Problem-Solving Accuracy. We administered 16 sample story problems from the California Department of Education's Standardized Testing and Reporting Program (2009) for Grade 3; the problems were identical for both pretest and posttest with changes only in the numbers and proper names. The two forms were counterbalanced for presentation order. Pretest and posttests were scored as correct with a total 16 points possible.

Design

Within each participating classroom, children were randomly assigned to either a control group (N = 29; n = 14 MD) or to one of four treatment conditions (i.e., verbal strategies, N = 36, n = 15 MD; verbal + visual strategies visual-only, N = 48, n = 18 MD; visual strategies-only diagramming, N = 37, n = 15 MD; materials-only no strategies, N = 43, n = 11 MD). The uneven sample size across conditions reflected some attrition as well as a further randomization within the verbal strategy-only and verbal-strategy + visual conditions. Initially, children were randomly assigned to either Presentation Order A or B. Presentation Order A included directing children to follow a six-step ordered procedure to problem solution: identify and underline the question, circle the relevant numbers, place a square around the key words (graduate tutors prompted participant by indicating key words), cross out irrelevant information, make a decision (add, subtract, or both), and solve it. Another strategy sequence tested whether eliminating irrelevant information (crossing out all distracter sentences) early in the strategy instruction sequence was related to treatment outcome. Thus, Presentation Order B rearranged the strategy sequence by first entering the task of identifying and underlining the question, crossing out irrelevant information then placing a square around the key words, circling the relevant numbers, making a decision (add, subtract, or both), and solving the problem. Our rationale for this manipulation was to test whether working memory load demands from identifying relevant information were reduced by eliminating distracting information at the beginning rather than towards the end of the strategy instruction sequence. However, presentation order was not significantly related to treatment outcomes and therefore the conditions within the verbal-only and verbal + visual were collapsed to simplify the subsequent analysis.

We computed a 2 × 5 (ability group: children with and without MD (condition) on the pretest (TOMA-2), transfer (STAR), and calculation (WRAT-3) measures. As expected, a significant effect was found for in favor of children without MD, but not for the condition or the ability group × condition interaction. No significant differences were found to be related to the random assignment to condition as a function of risk status, χ^2 ($df = 4$, $N = 193$) = 4.43, $p > .05$; or gender, χ^2 ($df = 4$, $N = 193$) = 5.5, $p > .10$.

Common Instructional Conditions. All children in the study participated with their peers in their homerooms on tasks and activities related to the districtwide math curriculum. The schoolwide instruction across conditions was the enVisionMATH learning curriculum (Pearson, 2009). General problem-solving steps included the following: (a) understand, (b) plan, (c) solve, and (d) look back. An independent evaluation of this curriculum (Resendez & Azin, 2008) indicated in random trials (teachers assigned randomly to treatment or control condition) that gains emerged in Grades 2 through 4, following guidelines outlined by the U.S. Department of Education's What Works Clearinghouse Standards (2008), with effect sizes relative to the control condition in the .20 range. A number of the elements within the curriculum were also utilized in our treatments (e.g., find the pattern). However, in contrast to district instruction, our treatment conditions focused on specific components of problem solving over consecutive sessions presented in a predetermined order. In addition, the lesson plans for the experimental condition focused directly on the propositional structure of word problems.

Experimental Conditions. Each experimental treatment condition included 20 scripted lessons administered over 8 weeks. Each lesson was 30 min in duration and was administered three times a week in small groups of two to four children. Lesson administration was done by one of six tutors (graduate students or paraprofessionals). Children were presented with individual booklets at the beginning of the lesson, and all responses were recorded in the booklet. Each lesson within the booklet consisted of four phases: warm-up, instruction, guided practice, and independent practice.

The warm-up phase included two parts: calculation of problems that required children to provide the missing numbers ($9 + 2 = x$, $x + 1 = 6$, $x - 5 = 1$) and a set of puzzles based on problems using geometric shapes. This activity took approximately 3 to 5 min to complete.

The direct instruction phase lasted approximately 5 min. At the beginning of each lesson, the strategies or rule cards were either read to the children (e.g., to find the whole, you need to add the parts) or reviewed. Across the 20 lessons, seven rules were presented. Depending on the treatment condition, children were taught the instructional intervention: verbal strategy, visual strategy-only (diagramming), or verbal + visual strategies.

Steps for the verbal strategy-only approach included directing the child to find within the word problem (a) the question sentence, and underline it; (b) the sentences with the number, and circle the number; (c) the key word, and place a square around it; and (d) the irrelevant sentences, and cross them out. The child was directed to decide what needed to be done (add, subtract, or both) and then solve the problem. For the visual-spatial only (diagramming) strategy condition students were taught how to use two types of diagrams; the first one represented how parts comprise a whole, and the second represented how quantities are compared. The diagram consisted of two empty boxes, one bigger and the other smaller, in which students were to fill in the correct numbers representing the quantities from the word problem. For the second diagram, students were presented with an equation with a question mark, which acted as a placeholder for the missing number provided in the box. Students then identified which number was needed to replace the question mark and solve the problem. For the combined verbal + visual-spatial (diagramming) strategy condition, an additional step (diagramming) was added to the six verbal-only strategy steps. This step included directing students to fill in the diagram with given numbers and identify the missing numbers (question) in the corresponding slots in the boxes (i.e., part, part, whole).

We also randomly assigned children to the materials-only condition, which served as a more rigorous control condition because it removed the overt activities related to strategy instruction (e.g., underlining, diagramming). Administration of this condition also allowed us to assess whether practice in the incremental presentation of irrelevant propositions (as found in the strategy conditions) within the curriculum materials played a significant role in posttest outcomes.

The third phase, guided practice, lasted 10 min and involved children working on three practice problems. Tutor feedback was provided on the application of steps and strategies to each of these three problems. In this phase, children also reviewed problems from the examples from the instructional phase. The tutor assisted children with finding the correct operation, identifying the key words, and providing corrective feedback on the solution. A sample problem from Lesson 12 was:

> Jackie went to the store with some friends. Jackie bought 2 jars of shiny marbles. There are 32 marbles in one jar. There are 62 marbles in the second jar. The marbles are green, red, blue, and silver. How many marbles are there in all?

The fourth phase, independent practice, last 10 min and required students to independently answer another set of three word problems without feedback. If the student finished the independent practice tasks in less than 10 min, they were presented with a puzzle to complete. Student responses were recorded for each session to assess the application of the intervention strategies and problem-solving accuracy. For the verbal + visual-strategy condition, points were recorded for correctly choosing the diagram, inserting correct numbers, applying strategies, identifying the correct operations, and correctly solving the problem (36 points possible). The mean points scores averaged across all sessions were 26.01 (SD = 6.02) and 29.39 (SD = 6.20) for children with and without MD, respectively. For the visual-strategy-only condition, points were recorded for correctly choosing the correct diagram, correctly filling in the numbers for the diagram, identifying the correct operations, and correctly solving the problem (36 points possible). The mean points scores averaged across all sessions were 23.49 (SD = 4.36) and 28.34 (SD = 4.59) for children with and without MD, respectively. For the verbal-strategy-only condition, points were recorded for identifying the correct numbers, applying strategies (e.g., underlining), identifying the correct operations, and solution accuracy (30 points possible). The mean points scores averaged across all sessions were 20.98 (SD = 2.93) and 23.25 (SD = 3.87) for children with and without MD, respectively.

Increments in Lesson Plans. The materials for word problems for each independent practice session included three parts: question sentences, number sentences, and irrelevant sentences. The number of sentences gradually increased across the sessions: Lessons 1 through 7 focused on identifying critical information for word problems of four sentences with one irrelevant sentence, Lessons 8 and 9 focused on five-sentence-long word problems with two irrelevant sentences, Lessons 10 through 15 focused on six-sentence-long word problems with three irrelevant sentences, Lessons 16 and 17 focused on seven-sentence-long word problems with four irrelevant sentences, and Lessons 18 through 20 focused on eight-sentence-long word problems with five irrelevant sentences.

Treatment Fidelity

During the lesson sessions, tutors were randomly evaluated by independent observers (i.e., a postdoctoral student, a non-tutoring graduate student, the project director). The observers independently filled out evaluation forms covering all segments of the lesson intervention. Points were recorded on the accuracy to which the tutor implemented the instructional sequence as intended. Observations of each tutor occurred for six sessions randomly distributed across instructional sessions. Interrater agreement was calculated on all observations and exceeded 90% across all observed categories. Tutors following each step of strategy implementation (10 observable treatment specific items were coded) yielded a mean of 98% (*SD* = .41) across all observed sequences and conditions. The mean percentage of strategy implementation as intended (maximum score of 10) for each condition was 9.78 (*SD* = .41) for verbal-only, 9.93 (*SD* = .25) for verbal + visual, 9.76 (*SD* = .43) for visual-only, and 9.59 (*SD* = .97) for materials-only.

Statistical Analysis

Children were drawn from 18 classrooms. Because the data reflected treatments used with children nested within classrooms, a hierarchical linear model (HLM; Bryk & Raudenbush, 2002) was necessary to analyze treatment effects. An HLM regression model was selected over a mixed ANCOVA model because of missing data (e.g., children moving after treatment had begun) and because the method has several advantages over an ANCOVA when children are nested within classrooms (see Snijders & Bosker, 1999, p. 43, for discussion). More important, the demarcation of MD versus NMD may be viewed as an arbitrary splitting of a continuous variable, and therefore simultaneous entry of contrasts into the regression model allowed for an assessment of the independent contribution of each subgroup to posttest scores.

The fixed and random effect parameter estimates were obtained using PROC MIXED in SAS 9.3. The criterion variables in the analysis were posttest scores for accuracy for the TOMA-2, STAR, and WRAT-3. The intraclass correlation for posttest TOMA-2 problem solving accuracy, STAR test problem-solving accuracy and WRAT-3 calculation accuracy was .05, .08, and .02, respectively. When partialed for pretest performance, the intraclass correlation was zero, .009, .006 for posttest TOMA-2 problem solving, STAR test problem solving, and calculation accuracy scores, respectively. All analyses were conducted with PROC MIXED (SAS, 9.3), estimated with full-information maximum-likelihood, and utilized robust standard errors (Huber-White) to allow for the nonindependence of observations from children nested within teachers. No doubt, concerns can be raised related to the Level 2 sample size (teachers/classrooms). According to Maas and Hox (2005), at least 50 Level 2 observations (classrooms, in this case) are needed to assure that standard errors estimates for the fixed effect components are unbiased. Others have argued that grouping of data greater than 10 is adequate (Snijders & Bosker, 1999, p. 44) for determining the estimated standard errors for the fixed effects. To make adjustments in our standard errors, we initially utilized a formula to correct the standard errors (see Hox, 2010, pp. 5–6, formula for design effects) that takes into consideration the intraclass correlation. The results did not vary from our use of robust standard errors (Huber-White). In addition we calculated the design effect for Level 2 sampling using the formula N/1+ (cluster size-1) (1-ICC; Killip, Mahfoud, & Pearce, 2004), applied in this case for problem solving accuracy as 193/1+ (18-1) (1-.05), therefore 193/16.15 = 11.95. Although there are different power assumptions (increase ICC), the overall conclusion was there was a sufficient sample for this study. In general, because our focus was on fixed effects (Level 1), the intraclass correlations were low, and variance related to random effects (error) was included in the analysis, we assume the data provided an adequate test of our exploratory hypotheses.

It is important to note that the treatment conditions were entered as binary variables (e.g., verbal-only MD group + 1, other conditions zero) that, by default, allowed comparisons of treatments to the control condition for children with MD. Thus, the intercepts in Table 2 reflect the posttest performance of the MD control group. The random effect for both models was the variance related to the teacher (classroom/tutor) effects. Snijders and Bosker (1999) argued that the power to detect significant parameters in multilevel research is frequently

TABLE 2

Hierarchical Modeling on Posttest Accuracy Scores as a Function of Treatment Condition and Ability Group

	Solution Accuracy (TOMA-2)		*Solution Accuracy (STAR)*		*Calculation Accuracy (WRAT-3)*	
Treatment Condition Comparison	*Estimate*	SE	*Estimate*	SE	*Estimate*	SE
			Fixed effects			
Intercept-Control MD	-0.17	0.20	–0.16	0.27	0.024	0.29
Verbal-only vs. Control MD						
MD	0.055	0.25	0.52	0.37	0.87*	0.38
NMD	0.58*	0.25	1.29**	0.34	1.05**	0.38
Verbal + visual vs. Control MD						
MD	0.52*	0.25	0.62	0.36	0.77*	0.37
NMD	0.85**	0.23	1.35**	0.32	1.23**	0.34
Visual-only vs. Control MD						
MD	–0.01	0.25	0.40**	0.38	0.34	0.38
NMD	0.79**	0.25	1.09**	0.34	0.97**	0.37
Materials-only vs. Control MD						
MD	0.36	0.28	0.95*	0.42	–0.07	0.43
NMD	0.98**	0.24	1.003**	0.32	0.49	0.34
Control NMD vs. Control MD						
NMD	0.60*	0.27	1.04**	0.38	0.02	0.39
Pretest	0.33**	.06	0.46**	0.079	0.71**	0.09
	Variance	SE	*Variance*	SE	*Variance*	SE
			Random effects			
Classroom	0.002	0.001	0		0	
Residual	.45**	0.05	1.01**	0.1	1.07**	0.11

Note. TOMA-2 = Test of Mathematical Abilities (Brown, Cronin, & McIntire, 1994); STAR = California Department of Education Standardized Testing and Reporting sample questions (2009); WRAT-3 = arithmetic subtest from Wide Range Achievement Test (Wilkinson, 1993); MD = children with math difficulties; NMD = children without math difficulties.
*$p < .05$. **$p < .01$.

low because of reductions in parameter reliability. For this reason, we maintained all multiple comparisons at $p < .05$.

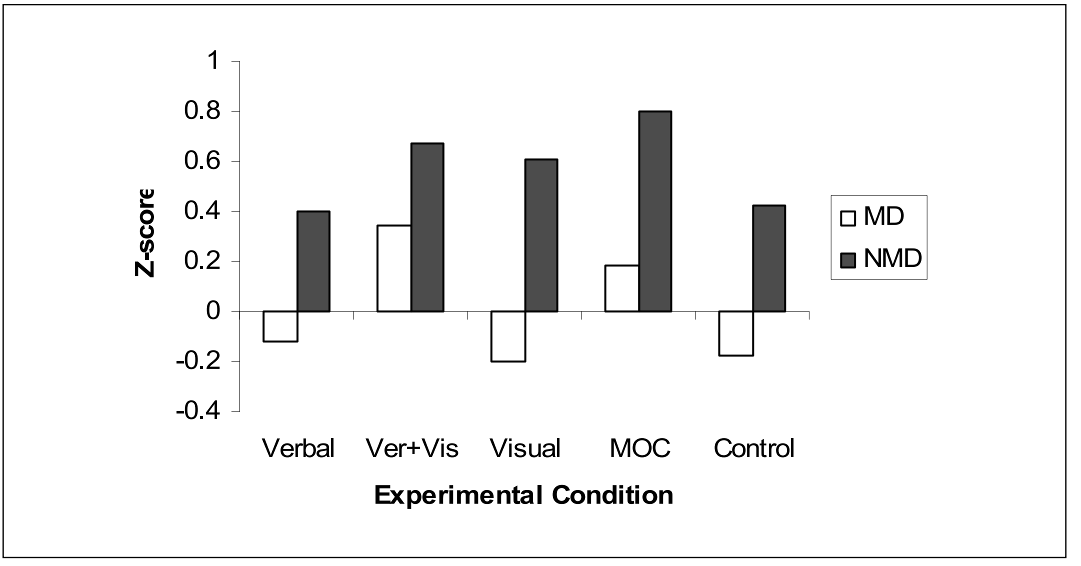
RESULTS

Pretest

Prior to the analysis of posttest scores, we computed a 2 (MD vs. NMD) × 5 (verbal, verbal + visual, visual, materials-only, control) mixed ANOVA on pretest scores for problem solving (TOMA-2, STAR) and calculation (WRAT-3). As expected, a significant main effect occurred for ability group for the TOMA-2, STAR and WRAT-3, $F(1, 181) = 39.48$, 11.26, and 43.28, respectively. No significant main effects occurred for treatment condition or ability group × condition interaction (all $p > .10$). We also determined if our random treatment assignments were potentially biased in the distribution of WMC, fluid intelligence, and reading measures. As expected, all *F*s were significant for ability group, but nonsignificant for the treatment or the ability group × treatment interaction emerged (all $p > .10$). However, as a precaution, in subsequent analysis we entered reading and fluid intelligence as potential moderators in the

FIGURE 1

Adjusted Posttest Solution Accuracy z-Scores (TOMA-2)



Note. TOMA-2 = Test of Mathematical Abilities (Brown, Cronin, & McIntire, 1994); Verbal = materials + verbal-only strategies; Ver+Vis = materials + verbal + visual strategies; Visual = materials + visual or diagram only; MOC = materials-only condition without overt strategy instruction; Control = classroom instruction as usual.

subsequent regression analysis. Entry of these variables did not improve the model fit, nor change the direction of the results, and therefore the results are not reported.

Posttest Accuracy

A shown in Table 2, the HLM analysis performed on posttest scores adjusted for pretest effects. To facilitate comparisons across dependent measures, the posttest scores were converted to *z*-scores based upon the mean and standard deviations of the pretest scores.

Problem-Solving Accuracy: TOMA-2. As shown on the left side of Table 2, the random effect for the intercept between classrooms was nonsignificant. Overall, the results indicated that posttest problem solving *z*-scores for the children in the control condition who were at risk for MD were nonsignificant and no better than chance. The adjusted posttest means for all conditions are shown in Figure 1. As expected, significant parameter estimates occurred for all children without MD relative to the MD control group. However, the important finding was that adjusted posttest scores for children with MD in the verbal + visual condition were significantly better than those scores of children with MD in the control condition. These results were followed up with a mixed ANCOVA comparing the 10 groups. As expected, the main effect for each condition was significant, $F(9,169) = 4.26$, $p < .001$. To simplify the Tukey analysis, post-hoc comparisons ($\alpha = .05$) were computed within the ability groups. For children with MD, a significant posttest advantage ($p < .05$) was found for verbal + visual when compared to other conditions (verbal + visual > materials-only > verbal-only = visual-only = control). No reliable effects (all $p > .05$) emerged between conditions for children without MD.

Clearly, these statistical outcomes were related to the power in our analysis. Thus, to address this issue, effect sizes are reported in Table 3. We calculated Hedge's $g = \gamma/[(SD_1^2)(N_1) + (SD_2^2)(N_2)/2]^{1/2}$, where γ is the HLM coefficient for the intervention effect, which represents the mean difference between treatment adjusted for both Level 1 and Level 2 covariates, N_1 and N_2 are the sample sizes, and SD_1 and SD_2 are the unadjusted standard deviations for the comparison condition, respectively. The

TABLE 3

Effect Sizes for Adjusted Posttest Scores as a Function of Treatment Conditions within Groups

	Students With MD			Average-Achieving Peers		
Treatment Condition	*TOMA-2*	*STAR*	*WRAT-3*	*TOMA-2*	*STAR*	*WRAT-3*
Verbal-only[a]						
vs. Verbal + visual	–0.62	–0.09	0.07	–0.3	–0.04	–0.2
vs. Visual	0.06	0.1	**0.47**	–0.27	0.22	0.05
vs. MOC	–0.39	–0.39	**0.90**	–0.48	0.36	**0.53**
vs. Control	0.05	0.31	**0.83**	0.04	0.1	**0.96**
Verbal + visual[a]						
vs. Visual	**0.67**	0.2	**0.40**	0.04	0.26	0.25
vs. MOC	0.23	–0.29	**0.83**	–0.18	0.4	**0.73**
vs. Control	**0.61**	**0.40**	**0.76**	0.32	0.13	**1.16**
Visual-only[a]						
vs. MOC	–0.45	–0.49	**0.43**	–0.22	0.14	**0.48**
vs. Control	–0.01	0.21	0.37	0.29	–0.11	**0.91**
MOC[a]						
vs. Control	**0.41**	**0.67**	–0.05	**0.49**	–0.25	**0.45**

Note. Effect sizes at or > .40 are in boldface. MD = math difficulties; TOMA-2 = Test of Mathematical Abilities (Brown, Cronin, & McIntire, 1994); STAR = California Department of Education Standardized Testing and Reporting sample questions (2009); WRAT-3 = arithmetic subtest from Wide Range Achievement Test (Wilkinson, 1993); MOC = materials-only condition without overt strategy instruction.
[a]Positive effect size in favor of verbal + visual.

Level 2 coefficients were adjusted for the Level 1 covariates under the condition that the Level 1 covariate (pretest) was grand mean centered. For the interpretation of the magnitude of the effect sizes, Cohen's (1992) distinction was used; an *ES* of 0.20 is considered small, and *ES* of 0.50 and 0.80 are considered moderate and large, respectively.

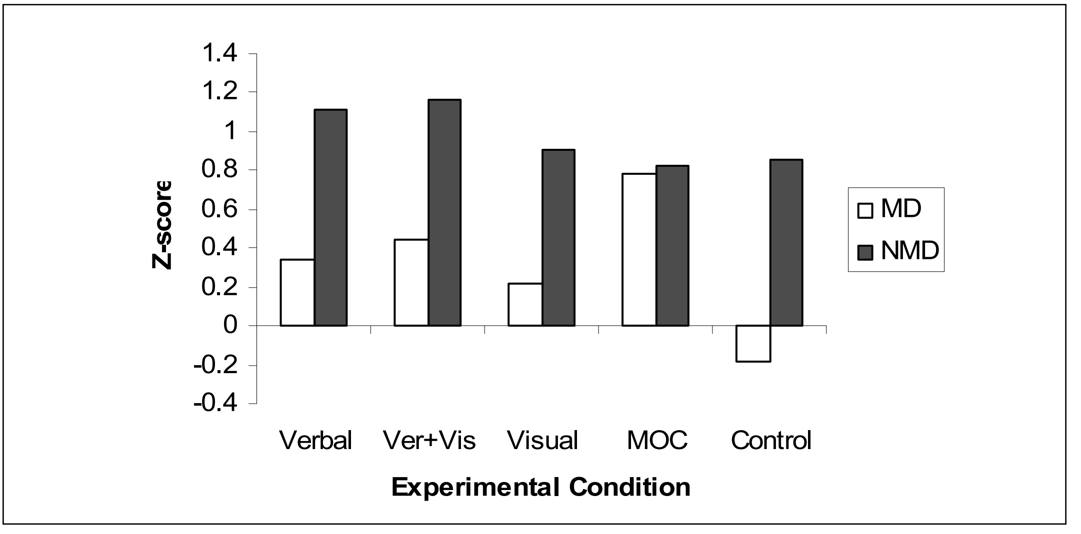
The results shown in Table 3 indicate that effect sizes were in the moderate range for the verbal + visual condition when compared to the control condition for children with MD. None of the effect sizes for treatment comparisons exceeded .50 for children without MD. We also computed effect sizes on adjusted posttest scores between ability groups within each condition. The negative effect sizes (in favor of the NMD group) between ability groups within conditions were –.72, –.41, –1.09, –.83, and –.69 for the verbal-only, verbal + visual, visual-only, materials-only, and control conditions, respectively. The results revealed that differences between ability groups were smaller in the verbal + visual condition and largest for the visual-only condition.

Problem Accuracy: STAR. As shown in the middle of Table 2, children with MD in the materials-only condition yielded significantly higher adjusted posttest scores than children with MD in the control condition. As expected, all children without MD outperformed children with MD in the control condition. Figure 2 shows adjusted posttest scores for all conditions.

These results were again followed up with a mixed ANCOVA comparing the 10 groups. As expected, the main effect for group was significant, $F(9,168) = 2.79$, $p < .01$. Tukey analyses were computed within ability groups. As shown in Figure 2, a significant advantage ($p < .05$) was found for materials-only condition relative to the control condition (materials-only > control) but not ($p > .05$) for the other treatment conditions (materials-only = verbal + visual = verbal-only) for children with MD. No reliable effects (all $p > .05$) emerged between conditions for children without MD. As shown in Table 3, a moderate effect size emerged between materials-only and the control conditions for children with MD, whereas none of the effect sizes between treat-

FIGURE 2

Adjusted Posttest Solution Accuracy z-Scores, Experimental Problem-Solving Measure (STAR)



Note. STAR = California Department of Education Standardized Testing and Reporting sample questions (2009); Verbal = materials + verbal-only strategies; Ver+Vis = materials + verbal + visual strategies; Visual = materials + visual or diagram only; MOC = materials-only condition without overt strategy instruction; Control = classroom instruction as usual.

ments and control exceeded .40 for children without MD. A follow-up analysis compared mean scores between groups within conditions; effect sizes were –.73, –.67, –.66, –.02, and –.85 for the verbal-only, verbal + visual, visual-only, materials-only, and control conditions, respectively. All effect sizes except for materials-only favored children without MD.

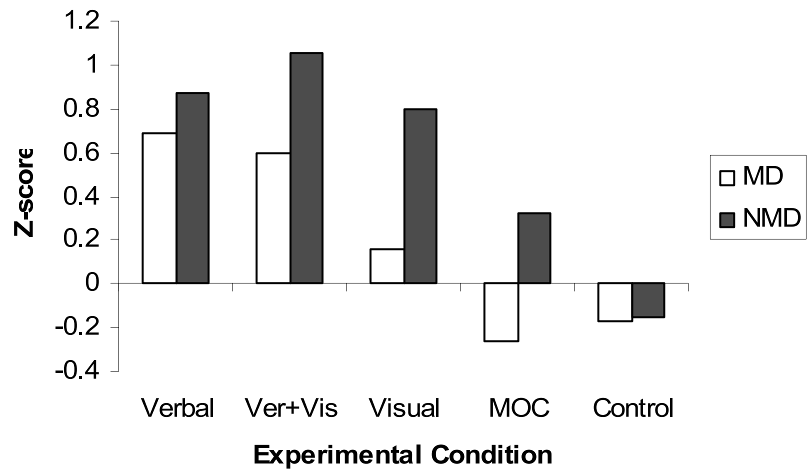
Summary

Depending on the criterion measure, children with MD's posttest problem-solving performance was higher under the verbal + visual or the materials-only conditions when compared to the control conditions. The above results do beg the question, however, as to why children in the materials-only condition shared some of the same advantages as children in the verbal + visual condition. To test whether these conditions made different demands on children's WMC, we computed the correlations between WMC in the executive system at pretest and problem-solving accuracy at posttest performance within each condition. We selected the executive system because no significant group performance differences at pretest occurred on the visual-spatial working memory measure. For children with MD, the correlation between WMC at pretest and posttest problem solving accuracy was .54, –.18, –.37, –.72 and .66, for verbal-only, verbal + visual, visual-only, materials-only, and control conditions, respectively. The positive coefficients suggested that children with higher WMC did better than those with low WMC. In contrast, the negative correlation suggested that compensatory processing occurred such that children with lower WMC benefited more from the conditions than children with higher WMC. Coupled with the previous findings, the correlation analyses suggested that verbal + visual draws fewer resources from a working memory system than the other conditions, whereas compensatory processing appears to be particularly strong for children with MD in materials-only treatment condition.

Calculation. As shown on the right side of Table 2, the important finding was that children with MD in the verbal and verbal + visual conditions yielded significantly higher posttest scores than the control MD condition. These results were again followed up with a mixed ANCOVA comparing the 10 groups. As expected, the group

FIGURE 3

Adjusted Posttest Calculation Accuracy z-Scores (WRAT-3)



Note. WRAT-3 = arithmetic subtest from Wide Range Achievement Test (Wilkinson, 1993); Verbal = materials + verbal-only strategies; Ver+Vis = materials + verbal + visual strategies; Visual = materials + visual or diagram only; MOC = materials-only condition without overt strategy instruction; Control = classroom instruction as usual.

main effect was significant, $F(9,169) = 3.52$, $p < .001$. The Tukey analyses were again computed between conditions within ability groups. As shown in Figure 3, a significant advantage ($p < .05$) was found for the verbal-only and verbal + visual only conditions relative to the other conditions (verbal-only = verbal + visual > materials-only = control; visual-only = materials-only = control) for children with MD. For children without MD, a significant advantage (p < .05) in posttest performance was found for all strategy conditions relative to the control conditions. In addition, posttest accuracy scores for the materials-only condition were statistically comparable to the control condition (verbal = verbal + visual = visual > control, verbal + visual > materials-only, materials-only = control). Table 3 shows that treatment effects within both ability groups were particularly strong when verbal and verbal + spatial conditions were compared to the control conditions (all *ES* > .80). The effect sizes between ability groups within conditions were .07, –.25, –.54, –.56, and –.24 for the verbal-only, verbal + visual, visual-only, materials-only, and control conditions, respectively. Posttest performance differences between ability groups were highest for the visual and materials-only conditions when compared to the other conditions.

DISCUSSION

The purpose of this exploratory study was to identify an effective intervention to improve word problem-solving accuracy for children at risk for MD. We specifically wanted to determine whether intensive strategy instruction plus practice handling irrelevant information within word problems would improve posttest performance in children with MD relative to children without MD in the control condition. Treatment conditions included either verbal and/or visual-spatial strategies and/or materials-only with systematic increases in the number of irrelevant sentences for solving word problems across 20 instructional sessions. Three major findings occurred.

First, depending on the problem-solving measure, children with MD in the verbal + visual and materials-only conditions outperformed children with MD in the control condition. This finding occurred on a standardized measure

(TOMA-2) and an experimental transfer measure (STAR). We assume that these interventions for children with MD worked for different reasons. As suggested by the negative correlation between WMC and posttest accuracy, the materials-only condition appeared to compensate for low working memory abilities. In contrast, the verbal + visual spatial condition appeared to fit within Mayer's (2005) hypothesis suggesting that combining verbal and visual resources places fewer demands on WMC.

It is interesting that strategy conditions did not improve the problem-solving performance in children without MD. These findings fit with our hypothesis that if declarative knowledge is intact (e.g., reading comprehension, computational knowledge), strategy instruction may be redundant and fail to provide helpful information (possibly redundant or elaborative information) to solve word problems.

Second, both ability groups benefited from strategy conditions for calculation relative to the control conditions. Verbal strategy conditions were found to be particularly robust compared to the materials-only and control conditions. However, the reason that these findings differ from those found on measures of problem-solving accuracy is unclear. We assume that improvements in calculation accuracy occurred because such activities were part of each lesson plan and therefore practice and feedback played a major role in the performance of children with and without MD. As shown in Figure 2, strategy instruction clearly provided an additional boost in performance when compared to the control conditions for both groups. Thus, we assume that because all training sessions involved practice in calculation and that strategies directed children's attention to number combinations, these two activities played a major role in the outcomes. A competing interpretation is that drill and practice in calculation skills were emphasized in the general education classroom. To address this issue, we took into consideration variance between classrooms (random effects related to teacher) in our analysis. However, the variance related to intercepts between classrooms was not significant. Thus, although teachers placed a varying emphasis on calculation skills (curriculum-based measurement drills), this variation did not appear to play a major role in accounting for the results. Our findings are consistent, however, with studies that have shown that interventions providing intensive drill and practice on basic math facts as part of their lesson plans facilitate calculation accuracy (e.g., Fuchs et al., 2009). What is unique is our finding that strategy training on word problems boosted calculation accuracy relative to the materials-only condition. Because the materials-only condition also provided practice in calculation skills as part of the lesson plans, we infer that the strategy training conditions facilitated transfer of calculation skills. Whereas previous studies have focused primarily on improving calculation skills and assessing transfer to problem-solving accuracy, the current study focused on problem-solving instruction and considered calculation accuracy as a transfer measure.

Because the materials-only condition also provided practice in calculation skills as part of the lesson plans, we infer that the strategy training conditions facilitated transfer of calculation skills.

Finally, the cognitive demands related to each condition varied within ability groups. The important findings were that the verbal-only strategy and control conditions made demands on WMC with MD. In contrast, posttest performance under the materials-only condition was negatively correlated with WMC, suggesting this condition compensated for cognitive processing.

Question 1: Are Some Cognitive Strategies More Effective Than Others for Children With MD?

The adjusted posttest mean *z*-scores for children with MD trained to use verbal + visual strategies were statistically comparable to children without MD in the control condition (see Figures 1 and 2). The results were in an expected direction in that strategy training would be more beneficial for children with MD than for children without MD. The results were also consistent with several investigations showing that children with MD are more likely to experience greater processing constraints in cognition when compared to children without MD (e.g., Andersson & Lyxell, 2007; Geary, 2003; Swanson & Beebe-Frankenberger, 2004). Thus, as expected, children with MD were more likely to benefit from strategy instruction than children without MD.

A meta-analysis synthesizing research on cognitive studies of MD (Swanson & Jerman, 2006) suggested that visual working memory in children with MD is relatively intact when compared to the verbal working memory processes and, therefore, is an important route for strategy training. Based on the assumption that visual memory in children with MD is relatively more intact than verbal memory (Swanson & Beebe-Frankenberger, 2004), the results related to visual-spatial strategies when compared to verbal strategy conditions for children with MD were counterintuitive. Thus, an obvious question emerges as to why visual-spatial strategies failed to help some children with MD. Our best explanation is that not all children had adequate resources to enact this visual strategy without placing excessive demands on their verbal skills. The visual-spatial strategy along with a verbal strategy, however, may have provided a technique that allowed children to focus on the relevant verbal aspects of the task. Diagramming numbers might have activated the relevant verbal information, while preventing irrelevant information from interfering with problem solving solutions.

The results related to the materials-only condition are less clear. We provide two explanations. The first is that a distinction can be made between training that reflects a process general approach (i.e., practice on problem solving tasks that increase processing demands) and a process specific approach (e.g., verbal + visual strategies). We infer that children with poor problem-solving skills may draw upon both strategies. The second explanation is that different training strategies can produce similar mean performances while also yielding different relationships with WMC. Negative correlations were found between WMC at pretest and problem-solving accuracy at posttest, but positive correlations in the other conditions. Thus, although WMC may be related to posttest outcomes for certain strategies, the overall results suggest that different resources from WMC may have been tapped.

Question 2: Do Cognitive Strategies Place Different Demands on Cognitive Processes in Children With MD?

Although Question 2 was partly answered in the above section, it must be placed in context of the models discussed in the introduction. The first model argued that if reading, computation, and general fluid intelligence are fairly intact for children with MD, reliable use of cognitive strategies supersedes the role that any individual differences in WMC might play. In contrast to this hypothesis, however, we found that WMC did correlate with posttest measures. Thus, the results do not support the notion that WMC plays a secondary role in strategy outcomes as related to problem-solving accuracy.

A second model suggested that a limited-capacity working memory system underlies word problem-solving difficulties in children with MD. This model has extensive support in the literature and is consistent with several theorists who adopt a general resource approach in which individual differences on cognitive and aptitude measures draw on a limited supply of working memory resources (e.g., Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008). This model assumes that although WMC may act in tandem with other processes, this general system may operate independent of strategy conditions. That is, WMC may be related to the level of problem-solving performance, but it does not moderate outcomes related to strategy conditions. The present study clearly shows that WMC accounted for important variance in problem solving accuracy for some strategy conditions, suggesting that certain strategies draw upon more working memory resources than others.

A third model suggested that strategy training compensates for individual differences in WMC. This model predicted that outcomes related to strategy training would significantly interact with individual differences in WMC. It was assumed that strategy training would free up resources for children with relatively weak WMC. Such was the case in this study for only one condition (i.e., materials-only). Thus, the results suggest that no other clear compensatory processing occurred for children with MD relative to the direct strategy conditions.

Implications

Our findings have three applications for current research. First, this study may account for why some children benefit from strategy instructions and others do not. We found that a key variable in accounting for the outcomes was WMC. Clearly, WMC would not be the only processing variable to account for the outcomes; however, its role in this study appeared to be fairly robust. It may be the case that when children with both computation and reading difficulties are included in the analysis, effects would be different.

Second, the results showed improvement on a norm-referenced test. The majority of intervention studies for problem solving have shown gains on experimental measures and less so on standardized measures (Powell, 2011). Thus, we were able to improve performance substantially on materials related to standardized tests. Although we used *z*-scores (based on raw scores) rather than national norms to compare treatment conditions, it is important to note that children with MD who were in the verbal + visual and materials-only conditions exceeded their peers when compared to the other conditions.

Finally, curriculum instruction that includes solving problems with gradual increases in irrelevant propositions may be an effective procedure for boosting solution accuracy in children with MD. As shown in Figures 1 and 2, the materials-only condition was just as effective as some of the overt strategies conditions in boosting solving accuracy in children with MD. This may have occurred for several reasons, but mostly likely related to the difficulty such children find in discriminating relevant from irrelevant facts in word problems (e.g., Passolunghi et al., 2001). Our previous work has shown that children with MD engage in number grabbing, selecting numbers from text without regard to the number's relationship to the problem's meaning. To avoid number grabbing, children must be directed to consider relevant information within the context of increasing irrelevant information (i.e., competing semantic and numerical information).

Summary

The adjusted posttest scores showed that general and specific strategy conditions allowed children with MD in problem solving to exceed their peers in problem-solving accuracy. The results also suggest that different processes are accessed across strategies, some drawing extensively on WMC whereas others drawing less on those processes.

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. http://dx.doi.org/10.1016/j.jecp.2006.10.001

Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal, 103*, 51–73. http://dx.doi.org/10.1086/499715

Brown, V. L., Cronin, M. E., & McIntire, E. (1994). *Test of mathematical abilities* (2nd ed.). Austin, TX: PRO-ED.

Brown, V. L., Hammill, D., & Weiderholt, L. (1995). *Test of reading comprehension*. Austin, TX: PRO-ED.

Bryk, A. S., & Raudenbush, S. W. (2002). *Hierarchical linear models: Applications and data analysis methods.* London, England: SAGE.

California Department of Education. (2009). *STAR sample test questions*. Retrieved from: http://starsamplequestions.org/

Censabella, S., & Noël, M. P. (2008). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology, 14*, 1–20. http://dx.doi.org/10.1080/09297040601052318

Cohen, J. (1992). "A power primer." *Psychological Bulletin, 112*, 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Colom, R., Abad, F. J., Quiroga, M. Á, Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence, 36*, 584–606. http://dx.doi.org/10.1016/j.intell.2008.01.002

Cook, J. L., & Rieser, J. J. (2005). Finding the critical facts: Children's visual scan patterns when solving story problems that contain irrelevant information. *Journal of Educational Psychology, 97*, 224-234. http://dx.doi.org/10.1037/0022-0663.97.2.224

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128,* 309–331. http://dx.doi.org/10.1037/0096-3445.128.3.309

Fletcher, J. M., Epsy, K. A., Francis, P. J., Davidson, K. C., Rourke, B. P., & Shaywitz, S. E. (1989). Comparison of cutoff and regression-based definitions of reading disabilities. *Journal of Learning Disabilities, 22,* 334–338. http://dx.doi.org/10.1177/002221948902200603

Fuchs, L., Fuchs, D., Prentice, K., Hamlett, C. L., Burch, M., Hamlett, . . . Janeck, D. (2003). Enhancing third-grade student's mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology, 95,* 293–304. http://dx.doi.org/10.1037/0022-0663.95.2.293

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

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating number combination and word-problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101*, 561–576. http://dx.doi.org/10.1037/a0014701

Geary, D. C. (2003). Math disabilities. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 199–212). New York, NY: Guilford.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*, 1202–1242. http://dx.doi.org/10.3102/0034654309334431

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*(4), 9–15. http://dx.doi.org/10.1111/j.1467-7687.2009.00848.x

Hox, J. J. (2010). *Multilevel analysis* (2nd ed.). New York, NY: Routledge.

Hresko, W., Schlieve, P. L., Herron, S. R., Sawain, C., & Sherbenou, R. (2003). *Comprehensive math abilities test*. Austin, TX: PRO-ED.

IDEA Regulations, 34 C.F.R. § 300 (2012).

Jitendra, A. K., Griffin, C. C., McGoey, K., Gardill, M. C., Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk for mild disabilities. *Journal of Educational Research, 91,* 345–355. http://dx.doi.org/10.1080/00220679809597564

Killip, S., Mahfoud, Z., & Pearce, K. (2004). What is an intracluster correlation coefficient? *Annals of Family Medicine, 2,* 204–208. http://dx.doi.org/10.1370/afm.141

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

Kolloffel, B., Eysink, T., de Jong, T., & Wilhelm, P. (2009). The effects of representation format on learning combinatories from an interactive computer. *Instructional Science, 37*, 503–515. http://dx.doi.org/10.1007/s11251-008-9056-7

Maas C., & Hox, J (2005). Sufficient sample sizes in multiple regression analysis. *Methodology, 1,* 86–92. http://dx.doi.org/10.1027/1614-2241.1.3.85

Marzocchi, G. M., Lucangeli, D., De Meo, T., Fini, F., & Cornoldi, C. (2002). The disturbing effect of irrelevant information on arithmetic problem solving in inattentive children. *Developmental Neuropsychology, 21*(1), 73-92. http://dx.doi.org/10.1207/S15326942DN2101_4

Mastropieri, M., Scruggs, T., & Shiah, R. (1997). Can computers teach problem-solving strategies to children with mental retardation? *Remedial and Special Education, 18,* 157–165. http://dx.doi.org/10.1177/074193259701800304

Mayer, R. E. (2005). Cognitive theory and multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816819.004

Montague, M. (1997). Cognitive strategy instruction in mathematics for students with learning disabilities. *Journal of Learning Disabilities, 30*, 164–177. http://dx.doi.org/10.1177/002221949703000204

Montague, M. (2008). Self-regulation strategies to improve mathematical problem solving for students with learning disabilities. *Learning Disability Quarterly, 31*(1), 37-44.

Montague, M., Warger, C., & Morgan, T. H. (2000). Solve it! Strategy instruction to improve mathematical problem solving. *Learning Disabilities Research & Practice, 15*, 110–116. http://dx.doi.org/10.1207/SLDRP1502_7

Ng, S. F., & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education, 40*, 282–313.

Passolunghi, M. C., Cornoldi, C., & De Liberto, S. (2001). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition, 27,* 779–790. http://dx.doi.org/10.3758/BF03198531

Passolunghi, M. C., & Siegel, L.S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80,* 44–57. http://dx.doi.org/10.1006/jecp.2000.2626

Pearson. (2009). *Scott Foresman-Addison Wesley enVisionMATH.* Lebanon, IN: Pearson Education. Available http://www.pearsonschool.com/index.cfm?locator=PSZu69

Powell, S. R. (2011). Solving word problems using schemas: A review of literature. *Learning Disabilities Research and Practice, 26,* 94–108. http://dx.doi.org/10.1111/j.1540-5826.2011.00329.x

Raven, J. C. (1976). *Colored progressive matrices test.* London, England: Lewis.

Resendez, M., & Azin, M (2008). *A study on the effects of Pearson's 2009 enVisionMath program.* Jackson, WY: PRES Associates. Retrieved from http://www.pearsoned.com/wp-content/uploads/envisionmath_efficacy_study_year1_final_report.pdf

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled. *Child Development, 60,* 973–980. http://dx.doi.org/10.1111/j.1467-8624.1989.tb03528.x

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2000 ed.). Thousand Oaks, CA: Sage.

Swanson, H. L. (1995). *S-cognitive processing test (S-CPT): A dynamic assessment measure.* Austin, TX: PRO-ED.

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. http://dx.doi.org/10.1037/0022-0663.96.3.471

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children's word problem solution. *Journal of Experimental Child Psychology, 55,* 374–395. http://dx.doi.org/10.1006/jecp.1993.1021

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76*, 249–274. http://dx.doi.org/10.3102/00346543076002249

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. http://dx.doi.org/10.1037/0022-0663.100.2.343

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816819.003

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence, 38*, 255–267. http://dx.doi.org/10.1016/j.intell.2009.12.003

Unsworth, N., & Engle, R. W. (2007). On the division of short-term memory and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133,* 1038–1066. http://dx.doi.org/10.1037/0033-2909.133.6.1038

Van Garderen, D. (2007). Teaching students with LD to use diagrams to solve mathematical word problems. *Journal of Learning Disabilities, 40*, 540–553. http://dx.doi.org/10.1177/00222194070400060501

Waesche, J. B., Schatschneider, C., Maner, J. K., Ahmed, Y., & Wagner, R. K. (2011). Examining agreement and longitudinal stability among traditional and response-to-intervention based definitions of reading disability using the affected-status agreement statistic. *Journal of Learning Disabilities, 44,* 296–307. http://dx.doi.org/10.1177/0022219410392048

What Works Clearinghouse. (2008, December). *Procedures and standards handbook (version 2.0)*. Washington, DC: U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_standards_handbook.pdf

Wilkinson, G. S. (1993). *The wide range achievement test* (3rd ed.). Wilmington, DE: Wide Range.

Xin, Y. P. (2008). The effect of schema-based instruction in solving mathematics word problems: An emphasis on prealgebraic conceptualization of multiplicative relations. *Journal for Research in Mathematics Education, 39*, 526–551.

Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *The*

Journal of Special Education, 32, 207–225. http://dx.doi.org/10.1177/002246699903200402

ABOUT THE AUTHORS

H. LEE SWANSON, Distinguished Professor & Peloy Chair; **MICHAEL J. OROSCO,** Assistant Professor; and **CATHY M. LUSSIER,** Academic Coordinator, Graduate School of Education, University of California, Riverside.

This article is based on a study funded by the U.S. Department of Education, Cognition and Student Learning in Special Education (USDE R324A090002), Institute of Education Sciences, awarded to H. Lee Swanson. The authors are indebted to Loren Alberg, Catherine Tung, Dennis Sisco-Taylor, Kenisha Williams, Garett Briney, Kristi Bryant, Orheta Rice, and Olga Jerman in the data collection and/or task development. Appreciation is given to the Academy for Academic Excellence/Lewis Center for Educational Research (Corwin and Norton Campus). Special appreciation is given to Chip Kling, Sandra Briney, and Jan Gustafson-Corea. This article does not necessarily reflect the views of the U.S. Department of Education or the school district.

Address correspondence concerning this article to H. Lee Swanson, Graduate School of Education, Educational Psychology, University of California, Riverside, CA 92521 (e-mail: Lee.Swanson@ucr.edu).

Manuscript received September 2012; accepted November 2012.