SPECIAL EDUCATION

DOES UNDERSTANDING RELATIONAL TERMINOLOGY MEDIATE EFFECTS OF INTERVENTION ON DIFFERENCE WORD PROBLEMS FOR SECOND-GRADE STUDENTS?

ROBIN FINELLI SCHUMACHER

Dissertation under the direction of Professor Lynn S. Fuchs

The purpose of this study was to assess whether understanding relational terminology (i.e., *more*, *less*, and *fewer*) mediates the effects of intervention on difference word problems. Second-grade teachers who volunteered to participate were assigned to 1 of 3 conditions: schema-broadening word-problem intervention, calculation intervention, or business-as-usual control. Students within the word-problem intervention condition received explicit instruction on the difference problem type, which included a focus on understanding relational terminology within word problems. Analyses, which accounted for the nested structure of the data, indicated that, compared to the active and inactive contrast conditions, word-problem intervention significantly increased performance on difference problems and on understanding relational terminology and that those intervention effects on difference problems were partially mediated by students' understanding of relational terminology.

Approved ____ Date ___________________

DOES UNDERSTANDING RELATIONAL TERMINOLOGY MEDIATE EFFECTS OF INTERVENTION ON DIFFERENCE WORD PROBLEMS FOR SECOND-GRADE STUDENTS?

By

Robin Finelli Schumacher

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Special Education

August, 2010

Nashville, Tennessee

Approved:

Professor Lynn S. Fuchs, Chair

Professor Donald L. Compton

Professor Douglas Fuchs

Professor Kim Paulsen

Professor Bethany Rittle-Johnson

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First, I would like to thank Lynn Fuchs, my advisor and mentor. She has provided me with unwavering support throughout my academic experience at Vanderbilt. Her input and collaboration over the past three years more than prepared me for the culmination of my graduate degree. I would also like to thank my committee: Don Compton, Doug Fuchs, Kim Paulsen, and Bethany Rittle-Johnson. They provided new and interesting perspectives throughout my dissertation process and my final product is undoubtedly improved by their insight.

I would also like to thank my family. Stephen, my adoring husband, who put up with the long hours I devoted to my studies and provided constant support along the way. My parents, Carol and Bob, have always been supportive of my dreams and were tremendous emotional supports from afar. And finally, Millie and Melvin, who were calming forces in my life amidst my crazy work schedule. I couldn't have done it without you.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

According to recent data from The National Assessment of Educational Progress (NAEP, 2009), 69% of American fourth graders perform below the proficient level in mathematics and 18% perform below the basic level, as indexed by high stakes testing. Approximately half of the problems on these tests are word problems (NAEP). Simple word problems are part of the mathematics curriculum beginning in the primary grades; yet, mathematics instruction generally focuses on performing procedural algorithms more than solving word problems. To bolster mathematics learning in the primary grades, the National Council of Teachers of Mathematics (2006) encourages teachers to devote more attention to word-problem instruction. For these reasons, the efficacy of word-problem instruction is important to a comprehensive mathematics curriculum for primary-grade students.

Students struggle to solve word problems even when they perform well on corresponding computation problems, suggesting they fail to understand the language of word problems (Briars & Larkin, 1984; Cummins, Kintsch, Reusser, & Weimer, 1988; Hegarty, Mayer, & Green, 1992; Lewis & Mayer, 1987; Riley & Greeno 1988). In fact, incorrect answers on word problems are often the result of correct computation performed on incorrect problem representations (Lewis & Mayer, 1987). Mathematics word-problem solving is distinctly different from calculations because problems are presented linguistically, requiring students to read and interpret the problem, choose a solution strategy, set up the problem, and calculate to find the answer. Enhancing understanding of the language of word problems (the first step in the process) may increase primary-grade students' ability to understand the word-problem structure and therefore set up and solve the problem correctly (Stern, 1993). The purpose of the present study was to assess whether the effects of intervention on difference problems are mediated by second-grade students' understanding of relational terminology (e.g., which in this paper refers to the terms *more*, *less*, and *fewer*).

In this introduction, we explain three types of simple word problems, the defining features of the most difficult of these problem types (i.e., difference problems), and we discuss how the relational terminology and linguistic features of difference problems potentially make them more challenging. Then, we review effective word-problem interventions that include a focus on difference problems. Finally, we explain the purpose of the present study and identify our hypotheses.

Difference Problems and Relational Terminology

Simple word problems, which are solved using one-step addition or subtraction, are the major types of word problems incorporated into the primary-grade mathematics curriculum. These problems are classified into three types by their semantic structures and the situation described in the story as *total*, *difference,* or *change* problems (Riley, Greeno, & Heller, 1983). Several researchers have used this classification structure (e.g., Cummins et al., 1988; Morales, Shute, & Pellegrino, 1985; Powell, Fuchs, Fuchs, Cirino, & Fletcher, 2009; Riley et al., 1983), although these researchers sometimes use different problem-type labels. The problem-type structures reflect a combination of sets (i.e., the total problem type), a change in one set over time (i.e., the change problem type), or a comparison of sets (i.e., the difference problem type). Total and change problems prompt an action through the combination of two sets (total) and one set changing over time (change). Difference problems do not prompt an action; instead, they make a static comparison and incorporate relational terminology. Perhaps for this reason, difference problems are more difficult than total or change problems for primary-grade students, even though the computation required for all three problems types is similar (e.g., Cummins et al., 1988; Garcia, Jimenez, & Hess, 2006; Morales et al., 1985; Powell et al., 2009; Riley & Greeno, 1988).

In the present study, we focused on difference problems because of the challenge they present to primary-grade students. In difference problems, two sets or quantities are compared and through this comparison, the difference between them emerges as a third set (i.e., the difference set). In difference problems, any of these three sets can be the unknown quantity students are asked to find. Three subtypes of difference problems are formed based on which quantity is unknown. The most common subtype is the difference set unknown in which both static sets are given and the difference set is found (e.g., Jill has 5 marbles. Tom has 8 marbles. How many more marbles does Tom have than Jill?). When the difference set is given, either the compared set is unknown (e.g., Jill has 3 marbles. Tom has 5 more marbles than Jill. How many marbles does Tom have?), or the referent set is unknown (e.g., Jill has 8 marbles. She has 5 more marbles than Tom. How many marbles does Tom have?). As the unknown quantity changes, the language and story structure change. When the compared set is unknown, that set is the subject of the relational statement; when the referent set is unknown, that set is the object of the relational statement, with a pronoun used in the subject. See Table 1 for examples of difference problem subtypes and their corresponding algebraic representation.

Table 1

Subtypes of Difference Problems

Unknown (Subtype)	Difference Problems: Sample Problem	Difference Problems: Algebraic Representation	Alternative Wording: Sample Problem	Alternative Wording: Algebraic Representation
	Using Additive Relational Terminology		Equalize	
Difference	Jill has 5 marbles. Tom has 8 marbles. How many marbles does Tom have more than Jill?	8 – 5 = X	Jill has 3 marbles. Tom has 8 marbles. How many marbles does Jill need to have as many as Tom?	8 – 6 = X
Compared	Jill has 3 marbles. Tom has 5 more marbles than Jill. How many marbles does Tom have?	X – 3 = 5	Jill has 8 marbles. Tom needs 5 marbles to have as many as Jill. How many marbles does Tom have?	8 – X = 2
Referent	Jill has 8 marbles. She has 5 more marbles than Tom. How many marbles does Tom have?	8 – X = 5	Jill has 3 marbles. She needs 5 marbles to have as many as Tom. How many marbles does Tom have?	X – 6 = 2
	Using Subtractive Relational Terminology		Won't Get	
Difference	Jill has 5 marbles. Tom has 8 marbles. How many marbles does Jill have less than Tom?	8 – 5 = X	There are 10 kids at the birthday party. There are 8 cupcakes. How many kids won't get a cupcake?	10 – 8 = X
Compared	Tom has 8 marbles. Jill has 3 fewer marbles than Tom. How many marbles does Jill have?	8 – X = 3	There are 8 cupcakes at the birthday party. 2 kids won't get a cupcake. How many kids are at the party?	X – 8 = 2
Referent	Tom has 5 marbles. He has 3 less than Jill. How many marbles does Jill have?	X – 5 = 3	There are 10 kids at the birthday party. 2 kids won't get a cupcake. How many cupcakes are there?	10 – X = 2

Although these three subtypes of difference problems describe a comparative relationship, problems with unknown referent sets are the most difficult, followed by problems with unknown compared sets and then unknown difference sets (Riley & Greeno, 1988; Morales et al., 1985). One potential reason for the differential difficulty among the subtypes is the way in which the relational terminology is presented. For problems with the difference set unknown, the relational terminology is used in the question. For problems with the compared or referent sets unknown, the relational terminology is incorporated in a relational statement, which could be more difficult for students to determine the comparative relationship.

One way to evaluate the connection between relational terminology and performance on difference problems is to remove these terms and replace them with alternative wording. Two variations of difference problems, *equalize* and *won't get*, are common ways to rephrase difference problems without changing the problem structure. (See Table 1 for examples of alternative wording for difference problems and their corresponding algebraic representation.) Fuson, McCarroll, and Landis (1996) assessed first and second graders on difference versus equalize problems for all three subtypes and found that students consistently scored higher on equalize than difference problems. Because the phrasing of the comparative relationship is the defining feature between difference and equalize problems, these findings suggest relational terminology may explain lower performance. Similar results were documented by Fan, Mueller, and Marini (1994) when they assessed performance differences among difference, equalize, and won't get problems. They concluded performance increased for these alternative wordings because they prompt an action (like total and change problems), which difference problems with relational terminology fail to do.

A related body of work looks at the relational statements specifically within problems with unknown compared or referent sets to examine the viability of Lewis and Mayer's (1987) consistency hypothesis as an explanation for the increased difficulty of unknown referent set problems. In unknown compared set problems, the relational term (*more* or *less/fewer*) aligns with the computation required for solution (when *more* is used, addition is required; when *less/fewer* is used, subtraction is required); in unknown referent set problems this is not the case. Additionally, the relational statement in problems with the compared set unknown defines the relationship in terms of the newly introduced set; by contrast, problems with an unknown referent set define the relationship in terms of the already given set and incorporate a pronoun. For these reasons, Lewis and Mayer proposed that problems with unknown referent sets require problem solvers to rearrange the relational statement. For example, for "Jon has 4 apples. <u>He has 3 fewer apples than Eric</u>. How many apples does Eric have?," the second sentence would be rearranged by the problem solver to "Eric had 3 more apples than Jon" before determining the solution procedures. To rearrange the relational statement, the consistency hypothesis assumes the problem solver understands the symmetry of relational terminology (to change *fewer* to *more*) and can effectively reverse the position of the object and subject of the sentence to generate a relational statement that describes an equivalent relationship. Verschaffel, DeCorte, and Pauwels (1992) documented that students were more successful when solving unknown compared set problems. Yet, even though students spent more time solving problems with unknown referent sets, which suggests rearrangement, this extra time did not lead to greater accuracy, providing only mixed support for Lewis and Mayer's consistency hypothesis.

To further explore the consistency hypothesis, Stern (1993) conducted two studies assessing students' understanding of the symmetrical relationship of *more* to *less/fewer* in relation to solving difference problems with unknown referent sets. In line with the consistency hypothesis, Stern hypothesized understanding this symmetrical relationship was most pertinent to solving difference problems with unknown referent sets. First graders were presented with pictures of two quantities and asked to match relational statements with each picture. For example, students had to decide whether one, both, or neither statement (e.g., “There are 2 more cows than pigs” and “There are 2 fewer pigs than cows”) matched a picture. Although students understood the meanings of the sentences, some students failed to understand that *more* and *fewer* could be interchanged to describe the same relationship, reflecting confusion about the symmetrical relationship between *more* and *less/fewer*. Low performance on this task was significantly related to students’ ability to solve problems with unknown referent sets, providing some evidence for the consistency hypothesis. As revealed in these studies, interpreting relational terminology is one plausible explanation for poor performance on difference problems. Problems with unknown referent sets stand out as most difficult among the difference word-problem subtypes, potentially requiring an understanding of the symmetrical relationship between *more* and *less/fewer*.

Word-Problem Interventions

Intervention work on difference problems is most often situated within instructional programs on all three problem types. Much of the recent literature on word-problem instruction has relied on schema theory (Cooper & Sweller, 1987; Gick & Holyoke, 1883) to design intervention. By developing schemas for problem types (i.e., total, change, and difference), students learn to recognize defining features of each problem type, categorize a problem as belonging to a problem type, and thereby apply solution procedures efficiently to novel problems.

Jitendra and colleagues, for example, have enjoyed success explicitly teaching students to recognize distinctions among total, change, and difference word problems while representing these problem types with conceptual diagrams (e.g., Jitendra, Griffin, Deatline-Buchman, & Sczesniak, 2007; Jitendra, Griffin, Haria, Leh, Adams & Kaduvettoor, 2007; Jitendra, Griffin, McGoey, Gardill, Bhat, & Riley, 1998). Each diagram is unique to the underlying structure of the problem type. Students are encouraged to represent problem structures with the diagrams before solving the problem. For difference problems, the mathematical structure of the comparison between a bigger and small quantity the conceptual diagram depicts. Students first learn to use the diagram when all numerical information (i.e., compared, referent, and difference sets) is provided; then, word problems are presented with unknown quantities, mirroring each of the three difference word-problem subtypes. Students put the two given quantities into the diagram and put a question mark (?) in place of the unknown. To help students compute the answer, instruction on part-whole relationships (i.e., the bigger number is the whole and the smaller and difference numbers are the parts) is provided. In these studies, instruction has resulted in growth in overall word-problem performance, but effects specifically for difference problems have not been reported.

Like Jitendra, in our word-problem intervention research, we explicitly teach defining features of each problem type to scaffold problem representation and solution procedures. When working with the same three problem types (i.e., total, change, and difference problems), however, our approach to schema theory differs from that of Jitendra and colleagues in two major ways. First and most central to the present study, we teach students to represent problem structures algebraically (i.e., putting X in the place of the unknown), rather than with conceptual diagrams. For difference problems, we teach students to generate an algebraic representation based on the problem structure "B – s = D" (i.e., the bigger number is B, the smaller number is s, the difference number is D). Students are taught to identify important information in the problem, cross out irrelevant information, and put an X in the number sentence in place of the unknown. (Table 1 shows difference problem examples and their corresponding algebraic representation.) The second way in which our approach differs from that of Jitendra is that we explicitly focus students on novel problem features (such as graphs or pictures, irrelevant information, money, or combinations of problem types), which increase the challenge of recognizing the problem type. The goal is to help students become more flexible problem solvers. We call this instructional approach *schema-broadening instruction* or SBI. In a series of randomized control trials, SBI increased word-problem performance across total, change, and difference problems (Fuchs, Powell, Seethaler, Cirino, Fletcher, Fuchs, et al., 2009; Fuchs, Seethaler, Powell, Fuchs, Hamlett, & Fletcher, 2008; Fuchs, Zumeta, Schumacher, Powell, Seethaler, Hamlet et al., 2010); however, growth on difference problems was reported separately in only Fuchs, et al. (2010). Performance favored SBI for each of the three problem types; however, learning appeared less robust for difference than for total or change problems. These findings should be interpreted cautiously, however, because those analyses, which were included only for exploratory purposes, were underpowered.

In both of these research programs, an essential feature of intervention to solve difference problems is representing the problem structure as the bigger number minus the smaller number equals the difference number. Understanding relational terminology is necessary to do this successfully. However, explicit instruction on understanding relational terminology was not reported in Jitendra's work nor was it incorporated in our previous work. In Fuchs, et al. (2010), we evaluated student performance by problem type which revealed students represent the underlying structure of difference problems inaccurately more often than for total or change problems and as a result solved problems incorrectly. For this reason, we incorporated an additional focus on understanding relational terminology in the difference unit for our most recent iteration of our word-problem intervention, as described in the present study. The hope was that this additional instruction would contribute, in combination with the larger SBI intervention, to improved student ability to solve difference problems.

Purpose of the Present Study

The purpose of the present study was to gain insight into whether this newly introduced focus on relational terminology within the difference unit contributes to increased student performance on difference problems and to thereby assess whether difficulty associated with relational terminology accounts for the differential difficulty of the difference problem type. We specifically examined whether student understanding of relational terminology mediates the effects of the intervention during the unit on difference problems.

The present study occurred within the context of a larger investigation in which classrooms were randomly assigned to word-problem (WP) intervention, calculations (CAL) intervention, or business-as-usual control (see Fuchs et al., 2010). With intervention in the larger investigation, all students received whole-class instruction, and at-risk students also received tutoring; intervention ran 17 weeks and supplemented the standard mathematics curriculum. WP intervention addressed total, difference, and change problem types, with the difference unit running weeks 8 through 13. In the larger investigation, measures of word-problem performance mixed the three problem types and were administered at the end of the 17 weeks, five weeks after the difference unit was implemented. In addition, students' understanding of relational terminology was not assessed in the larger investigation. To accomplish the present study's purpose, we added assessments specifically of performance on difference problems and of students' understanding of relational terminology. We administered these measures before and after the unit on difference problems (i.e., at weeks 7 and 14 of the larger investigation) which did not overlap the assessment period in the larger investigation. Readers should note the present study's measures were not part of the larger investigation, the purpose of which is not to isolate the effects of the relational terminology intervention features or to assess whether students' understanding of relational terminology mediates the effects of the difference word-problem type intervention.

The hypotheses in the present study are based on previous SBI intervention research showing efficacy without the relational terminology features (Fuchs et al., 2008; Fuchs et al., 2009; Fuchs et al., 2010), prior investigations of student performance on difference problems (Cummins et al., 1988; Garcia et al., 2006; Morales et al., 1985; Powell et al., 2009; Riley & Greeno, 1988), as well as earlier work suggesting a connection between understanding relational terminology and solving difference problems (Lewis & Mayer; 1987; Stern, 1993; Vershaffel et al., 1992). First, we posited that students receiving WP intervention would significantly outperform those in CAL and control conditions on difference problems, who would perform comparably to each other. Then, we conducted mediation analyses to address the following hypotheses: (a) WP intervention would significantly affect students' understanding of relational terminology; (b) students' understanding of relational terminology would in turn affect student performance on difference problems; and (c) students' understanding of relational terminology would mediate (at least partially) the effects of WP intervention on difference problems.

In these ways, we extend knowledge about whether understanding relational terminology is a generative mechanism within WP intervention for enhancing student performance on difference problems. We suggest the following proposed causal mechanism: Students provided explicit instruction on the meanings of relational terminology and the symmetrical relationship between *more* and *less/fewer* within a word-problem context will apply this understanding when interpreting difference problems, which will enhance understanding the relationships in difference problems. Understanding these relationships in difference problems will increase students' ability to accurately solve difference problems.

CHAPTER II

METHOD

Participants

In the larger investigation, 32 second-grade teachers (all female) from a large metropolitan school district volunteered to participate. Blocking within school, we randomly assigned their classrooms to one of three treatments: WP (*n* = 12), CAL (*n* = 12), or control (*n* = 8). Soon after random assignment, one CAL teacher's classroom was dissolved, leaving 11 classrooms in this condition. Consented students within these classrooms were included in the larger investigation if they had at least one T-score above 35 on the Vocabulary or Matrix Reasoning subtests of the Wechsler Abbreviated Intelligence Scale (WASI; The Psychological Corporation, 1999) and were a native English speaker or had successfully completed an English Language Learner program. As part of the larger investigation, students were also screened on calculation and word-problem measures to identify risk for inadequate learning outcomes. The screening measures were *Addition Fact Fluency* (Fuchs, Hamlett, & Powell, 2003) and *Single-Digit Story Problems* (Jordan & Hanich, 2000; adapted from Carpenter & Moser, 1984; Riley et. al., 1983). Cut-off scores for risk status were empirically derived from a previous database (Fuchs et al., 2010). Students who scored below this cut-point on both measures qualified for tutoring within the WP and CAL conditions: At-risk students in WP classrooms received word-problem tutoring; at-risk students in CAL classrooms received calculation tutoring; at-risk students in control classrooms did not receive tutoring.

For the present study, we relied on the same pool of not-at-risk and at-risk students from the larger investigation, except that we excluded students who were absent for pre- or posttesting. This resulted in 169 students in WP (142 not-at-risk and 27 at-risk students), 155 students in CAL (128 not-at-risk and 27 at-risk students), and 118 students in control (98 not-at-risk and 20 at-risk students). See teacher demographic data and student demographic and screening data, as a function of condition, in Table 2. Chi-square analysis and analysis of variance (ANOVA) revealed teachers did not differ as a function of condition on race or sex but did differ on years teaching. Post hoc analysis revealed WP teachers had significantly fewer years of teaching experience than CAL teachers (p = .035) or control teachers (p = .044), who were not different from each other (p = .94). We did not find this to be problematic because teachers did not deliver whole-class or tutoring instruction and because the role of years teaching in determining student outcomes is not clear (e.g., Wolters & Daugherty, 2007). Students did not differ as a function of condition on sex, race, subsidized lunch, years retained, or either screening measure.

Classroom and Tutoring Mathematics Instruction

We have three sources of information about what occurred for mathematics instruction. First, in the two conditions, where we designed and implemented mathematics instruction, we have a description of these experimental methods: for the WP condition (at the classroom and at the tutoring levels) and for the CAL condition (at the classroom and at the tutoring levels). Second, in the WP and CAL conditions, we coded the fidelity with which the WP and CAL instructional methods were implemented at the classroom and tutoring levels. Finally, given the focus of the present study on WP, all 32 teachers completed a questionnaire on which they reported how much time they spent on word-problem instruction in their classrooms

Table 2

Demographic Information

Variable	WP (teachers, *n* = 12) (students, *n* = 169) %	(n)	M	(*SD*)	CAL (teachers, *n* = 11) (students, *n* = 155) %	(n)	M	(*SD*)	control (teachers, *n* = 8) (students, *n* = 118) %	(n)	M	(*SD*)	χ^2	*p*	*F*	*p*
Teachers																
Male	0	(0)			0.0	(0)			0.0	(0)						
Race													2.27	0.32		
African American	25.5	(3)			18.2	(2)			0.0	(0)						
Caucasian	75	(9)			81.8	(9)			100.0	(8)						
Years Teaching			11.00	(9.42)			20.27	(10.24)			20.63	(10.54)			3.26	0.05
Students																
Male	50.3	(85)			43.2	(67)			47.5	(56)			1.78	0.41		
Race													2.61	0.96		
African American	33.9	(57)			36.1	(56)			32.2	(38)						
Caucasian	37.5	(63)			34.9	(54)			38.1	(45)						
Asian	4.8	(8)			3.2	(5)			5.9	(7)						
Hispanic	17.9	(30)			19.4	(30)			15.2	(18)						
Other	5.9	(10)			6.4	(10)			7.6	(9)						
Subsidized Lunch	75.1	(127)			82.6	(128)			76.3	(90)			2.53	0.28		
Retained	6.5	(11)			7.1	(11)			10.2	(12)			1.46	0.48		
Screeners																
Addition Fact Fluency			9.54	(4.91)			9.85	(5.18)			9.31	(4.88)			0.40	0.67
Single-Digit Story Problems			8.11	(4.11)			8.03	(3.76)			7.52	(3.61)			0.91	0.40
Math Status													1.31	0.97		
Low in word problems	20.7	(35)			20.0	(31)			22.0	(26)						
Low in calculations	11.8	(20)			10.3	(16)			11.9	(14)						
Low in both	18.9	(32)			17.4	(27)			20.3	(24)						

and what general methods they used for word-problem instruction. We did not, however, pose questions about how much time they spent on mathematics instruction generally or on calculations instruction.

Teachers' reports of their word-problem instruction. CAL and control students received the word-problem instruction their classroom teachers designed; WP students also received the word-problem instruction their classroom teachers designed; in addition, however, they also received the WP intervention our research assistant (RA) teachers delivered. To describe what occurred by condition during teacher-delivered word-problem instruction, all teachers completed a questionnaire (see Table 3 for means and standard deviations) on which they reported the average min per week they devoted to word-problem instruction. The time reported for WP teachers did not include the 90 – 120 min of instruction per week we provided, and they completed these surveys based on methods they used when we were not delivering our WP intervention.

From among the word-problem instructional strategies we provided teachers to select from (see Table 3), teachers reported similar strategies with no significant differences among conditions. However, teachers had the opportunity to describe other strategies. Three WP teachers reported using strategies from the WP intervention during their own word-problem instruction. CAL and control teachers did not report using any additional strategies.

Among the strategies teachers selected from, the average teacher rating for reliance on the basal program, *Houghton Mifflin Math* (Greenes et al., 2005), was 4.23 (*SD* = 0.72) across conditions, indicating heavy reliance. The basal program guides teachers to help students (a) understand, plan, solve, and reflect on the content of word problems, (b) apply problem-solution rules, and (c) perform calculations. Word problems from the basal text require simple arithmetic

Table 3

Teacher Survey Information (n = 31 teachers)

Variable	WP (*n* = 12) % (n)	M	(*SD*)	CAL (*n* = 11) % (n)	M	(*SD*)	Control (*n* = 8) % (n)	M	(*SD*)	*F*	*p*
Minutes of Word-Problem Instruction per Week		30.00	(14.92)		55.00	(37.88)		53.75	(33.96)	2.05	0.15
Problems from Basal Text		3.91	(0.83)		4.50	(0.67)		4.25	(0.46)	2.10	0.14
Strategies from the Basal Text		3.92	(0.67)		4.18	(0.60)		4.38	(0.74)	1.19	0.32
Stategies for Interpreting and Deriving Solution Procedures											
Keyword Strategy		4.83	(0.39)		4.82	(0.41)		5.00	(0.00)	0.77	0.47
Number Family Approach		2.83	(1.27)		2.64	(1.21)		3.25	(1.17)	0.60	0.56
Meta-Cognitive Strategies		3.00	(1.35)		2.73	(1.62)		2.50	(1.2)	0.31	0.74
Solution Procedures/ Work Shown											
Graphic Organizers		3.17	(0.84)		3.09	(1.38)		3.38	(1.19)	0.15	0.86
Number Sentences		4.75	(0.45)		4.55	(0.69)		4.63	(0.52)	0.39	0.68
Unknown after equal sign only	4 (33)			8 (72)			5 (63)				
Unknown in any position	8 (67)			3 (28)			3 (37)				
Draw Pictures		4.58	(0.52)		4.55	(0.69)		4.75	(0.46)	0.32	0.72
Provide Word Label for Answer		4.25	(1.06)		3.82	(1.25)		3.75	(0.71)	0.71	0.50

Note. Teachers responded to each question, except, minutes of word-problem instruction, using the following Likert Scale. 1 = never; 2 = rarely; 3 = every once in a while; 4 = sometimes; 5 = almost always.

for solution and are the same problem types included in the WP intervention (i.e., total, change, difference). Even so, problem types within the basal rarely present problems for which the unknown is in the first or second position of the algebraic representation of the problem (i.e., the compared or referent set within difference problems), which is an important feature of the WP intervention and a major focus of the present study. Across conditions, teachers also reported an average rating of 4.87 (*SD* = 0.34) for keyword instruction in which students are taught to identify specific key words and decide whether to add or subtract based on these terms alone, also indicating heavy reliance. Keyword instruction eliminates the need to understand the underlying word-problem structure (a pertinent feature of the WP intervention) and in the case of difference problems, where the term *more* indicates addition and the term *less/fewer* indicates subtraction. Therefore, keyword instruction renders the meanings of relational terminology irrelevant. In terms of expectations for student work, teachers reported heavy reliance on students showing how they derived solutions. Teachers reported an average rating of 4.65 (*SD* = 0.55) for requiring students to solve problems with a number sentence and an average rating of 3.97 (*SD* = 1.05) for requiring a word label (both of which occurred in WP intervention). For drawing pictures, teachers reported an average rating of 4.61 (*SD* = 0.56) (which did not occur in WP intervention).

Control condition. As reported in Table 3, therefore, essential features of word-problem instruction for the control condition included (a) using the basal text for strategies and sample problems, (b) keyword instruction, (c) writing number sentences, and (d) drawing pictures. Due to reliance on the basal text, difference problems with unknown compared or referent sets were rarely used. Three control teachers did report teaching students to represent and solve word problems with number sentences with the unknown information occurring before the equal sign (the first or second position of the number sentence), which aligns with WP instruction. In terms of the amount of time devoted to word-problem instruction, teachers in the control condition reported less than what was delivered in the WP condition.

CAL condition. CAL teachers reported a pattern of word-problem strategies that resembled those of control teachers. They relied on the basal text and keyword instruction and taught students to write number sentences or draw pictures to show their work. These students also received CAL instruction delivered by our RA teachers (referred to as MathWise). We delivered CAL instruction in addition to what teachers provided in their mathematics curriculum.

Our instruction focused on the concepts and procedures for addition and subtraction basic facts and double-digit addition and subtraction with and without regrouping. CAL instruction included the same number of whole-class and tutoring lessons as the WP intervention across the 17 weeks, and each classroom and tutoring lesson provided the same amount of researcher-delivered instructional time as the WP intervention.

WP condition. As reported on the survey, WP teachers relied on several methods for their own word-problem instruction, some of which aligned with the WP intervention (labeling answers and solving equations which vary the position of the unknown) and some of which did not align with our methods (keyword instruction and drawing pictures). Teachers reported spending, on average, 30 additional min per week of instruction beyond what our WP intervention provided.

To help readers connect the results of this study to the literature, we note that the efficacy of this WP intervention, which is known as Pirate Math, has been demonstrated previously (see Fuchs et al., 2008; Fuchs et al., 2009; Fuchs et al., 2010). As part of the larger investigation, however, we strengthened several instructional components of Pirate Math, including the incorporation of explicit instruction on relational terminology in the difference unit.

Pirate Math whole-class instruction, which occurs two times per week (45-60 min per session), was incorporated into the teachers' standard mathematics block. Pirate Math begins with an introductory unit that addresses foundational problem-solving skills important to the subsequent three units, each of which focuses on one of the three problem types, with cumulative review. The difference unit, relevant to the present study, is the third unit of the larger program. It runs six weeks from weeks 8 through 13 of the larger program. At-risk students also receive individual tutoring three times per week for 13 weeks; five weeks are devoted to the difference unit (occurring between weeks 8 and 13). Each tutoring session lasts 30 min. The instructional sequence in tutoring follows the whole-class instruction so that students work on the same skills in both settings with students being introduced to new skills during whole-class instruction. Tutoring augments whole-class instruction rather than replicating it by teaching additional strategies and providing further explanation for the hardest concepts addressed in whole-class instruction.

Pirate Math, which is based on SBI (Fuchs, et al., 2008; Fuchs, et al., 2009; Fuchs, et al., 2010), explicitly teaches students to conceptualize word problems in terms of problem types, to recognize defining features of each problem type, and to represent the structure of each problem type with an overarching equation (i.e., $a + b = c$; $d – e= f$). For each problem type, students are taught to RUN (i.e., Read the problem, Underline the question, and Name the problem type) through the problem before solving it. Students name the problem type by thinking of their defining features. In the present study, we focus on difference problems, for which students identify whether two things are being compared. Then students use the following procedure to solve the difference problem. First, they identify the overarching equation that represents the underlying structure of difference problems: the bigger amount minus the smaller amount equals the difference ($B – s = D$). Second, students identify the unknown in the difference problem they are solving and place X under that position of the equation. Third, students identify, check off, and write important numbers under the difference equation, $B – s = D$. Fourth, students write the mathematical signs (- and =). Finally, students solve the problem by finding X and labeling the numerical answer.

X can occur in any of the three positions of the difference equation (i.e., the missing number might be B, s, or D). (In the introductory unit, students are taught a simple procedure to solve for X.) When the compared or referent set is unknown, the X is positioned underneath B or s in the difference equation. For the present study, for which the difference unit is the focus, students are first taught to solve problems with the difference set unknown (as in "Jill has 8 marbles. Tom has 5 marbles. How many more marbles does Jill have than Tom?"), because these problems are the easiest subtype of difference problems (Riley & Greeno, 1988; Riley et al., 1983). In this problem, determining which numbers represent the bigger and smaller quantities requires an understanding of whether 8 is bigger or smaller than 5, thus representing B – s = D as 8 – 5 = X. When the difference set is given, and either the compared set (e.g., Jill has 8 marbles. Tom has 5 fewer marbles than Jill. How many marbles does Tom have?) or the referent set (e.g., Jill has 8 marbles. She has 5 more marbles than Tom. How many marbles does Tom have?) is unknown, problems are more difficult to represent because determining whether the unknown is the bigger or smaller amount requires understanding of the relational statement in the problem. To represent the problem structure when the difference set is given (as in the two examples just provided), students must translate the relational statement to understand that 5 is the difference, Tom has the smaller amount, and Jill has the bigger amount. This will lead them to represent B – s = D as 8 – X = 5. These harder difference problem subtypes are introduced once students have practiced representing and solving the easier difference problem subtype in which the difference set is unknown (and the missing information is in the third position of the algebraic representation).

During whole-class instruction on the difference unit, RA teachers introduce a new instructional component that focuses on helping students interpret relational statements is central to the present study. First, the meanings of *more*, *less*, and *fewer* are defined and reviewed. Second, students are taught to cover the difference number to decide what is the bigger or smaller amount using the relational term (e.g., in "Tom has 5 fewer marbles than Jill," the student covers 5 and re-reads; if Tom has less than Jill, he has the smaller amount and Jill therefore must have the bigger amount). Third, students practice writing an alternative relational statement that preserves the structure of the original sentence by switching the subject and object of the sentence and switching the relational term to its opposite (e.g., *more* instead of *less/fewer*). In the case of "Tom has 5 less marbles than Jill," the new sentence would read, "Jill has 5 more marbles than Tom." In the new sentence, Jill still has the bigger amount and Tom the smaller amount and the difference amount is 5. These activities have two purposes: (a) to teach the meaning of the relational terms for determining the bigger and smaller amounts in problems with unknown compared or referent sets and (b) to teach the symmetrical relationship of *more* to *less/fewer*. This activity occurs in 9 of the 11 difference unit lessons.

In tutoring, for which 15 lessons teach difference problems, the following activities address relational meaning. First, the "difference game" teaches students an additional strategy for determining the bigger and smaller amounts. In the first tutoring lesson of the difference unit, students learn the foundational elements for the game including (a) a review of the meanings of *more*, *less*, and *fewer* (as taught in whole-class instruction); (b) a review of which items are being compared in the relational sentence (as taught in whole-class instruction); (c) instruction on the meanings of the greater than (>) and less than (<) symbols (not addressed in whole-class instruction); and (d) application of the (>) and (<) to the relational term (i.e., *more, less* ,or *fewer*) to assist determining which amount is bigger or smaller (not addressed in whole-class instruction). After this initial instruction, students play the "difference game." The tutor provides a relational sentence. Students underline the two things being compared, writing the (>) or (<) symbols over the relational term in the sentence, and write B and S on the two things being compared. For example, consider this relational statement, “Jess has $5 more than Kesha.” First, students underline Jess and Kesha; second, they find the relational term, *more*, and write a greater than symbol (>) above; and finally, students write B over Jess’s name and S over Kesha’s name. This shows Jess has the bigger amount and Kesha has the smaller amount.

In the second through fifth difference unit lessons, students play the difference game with three relational sentences. In the sixth lesson, the stimulus changes from a relational statement to a whole difference word problem, as students apply the same three steps just described. After completing those steps, students identify whether the unknown is the bigger or smaller amount. In subsequent lessons, students play the difference game with the RA tutor mixing the stimulus between relational statements and difference word problems. In lesson 11 students learn to determine the bigger and smaller quantities when presented with new relational terminology (e.g., *older/younger* and *taller/shorter*). For the remainder of the difference unit, the stimuli in the difference game mix relational statements and difference word problems with all the relational terminology taught.

When solving difference problems throughout tutoring, RA tutors remind students to use methods from the difference game to interpret the relationship and accurately represent the problem structure in their equations. When students struggle during problem-solving activities, the tutor encourages them to use the (<) and (>) symbols and write “B” or “S” above items to facilitate correct problem representations.

In addition, as part of Pirate Math, students are taught to broaden students’ schemas for the problem type (regardless of the position of the unknown quantity in the problem). In line with SBI (Fuchs et al., 2008; Fuchs et al., 2009; Fuchs et al., 2010), students are therefore taught about novel problem features (e.g., graphs or pictures, irrelevant information, money). By sensitizing students to novel problem features that preserve the underlying structure of problem types, we hope students will become more flexible and capable problem solvers.

Each whole-class lesson incorporates five specific activities: (a) a review of previously taught concepts and solution strategies; (b) the daily lesson, which introduces new material; (c) teacher-led seatwork, which involves students working along with the teacher on one word problem pertinent to day daily lesson; (d) partner work, in which students solve two to three word problems in dyads; and (e) individual work, in which students are accountable for completing one word problem independently for which they earn points and monitor their daily progress. Each tutoring lesson comprises four activities: (a) a two-minute drill and practice activity for reviewing previously taught foundational skills; (b) the lesson, in which introduces new strategies and provides and guides students through two to three word problems while gradually decreasing support; (c) a two-minute sorting game in which students receive practice in recognizing the defining features of and naming problem types; and (d) completion of an independent word problem for which students earn points and monitor daily progress.

Delivery of Instruction

In the larger investigation, six RA teachers taught whole-class WP and CAL programs. Across the larger 17-week study, RA teachers delivered 34 whole-class lessons in each of the two RTI conditions: WP and CAL. For WP, 11 lessons were part of the difference unit (lessons 15 – 25). Each RA teacher was responsible for teaching one to three WP or CAL classrooms for the duration of the investigation. Twelve RA tutors delivered individual tutoring three times a week for 13 weeks, with a total of 39 tutoring lessons, 15 of which were part of the difference unit (lessons 13 – 27). Tutoring occurred outside the classroom at times when students did not miss important instruction, as determined by the classroom teacher. Each RA tutor delivered tutoring to three to six at-risk students, approximately half of whom received WP tutoring; the other half CAL tutoring.

Training occurred for the larger investigation prior to the 17-week intervention. To facilitate the provision of a standard protocol and fidelity of implementation, we used scripts to help RAs understand the nature of instruction. Scripts were studied prior to instruction; RAs were not permitted to memorize or read scripts. RA teachers and tutors attended separate one-day trainings on intervention procedures and salient features of SBI. Three RA teachers had prior experience with Pirate Math and demonstrated their initial lessons for the novice RA teachers so they could observe lessons in action. Six RA tutors were returning from the previous year and seven tutors were new. Returning RA tutors were paired with new tutors to practice tutoring procedures prior to working with students. Subsequent trainings occurred for both RA teachers and tutors before each unit. All RAs used audio-digital recording devices to record 100% of implemented lessons. To assess whether additional training was necessary, experienced RAs conducted live observations and/or listened to audio recordings.

Fidelity of Implementation

To assess treatment fidelity, we relied on the audiotapes of the whole-class and tutoring sessions (see above). At the end of the study, we randomly sampled 20% (RAs were not aware of which lessons would be coded) to represent the WP and CAL conditions, RAs, and lessons comparably. Prior to the study, we had prepared a checklist for every lesson, which listed the essential components of that lesson. As coders listened to tapes, they checked the essential components to which the RA adhered, with a percentage of essential components then derived.

In WP, fidelity for the difference unit was 93.59% for whole-class instruction and 94.23% for tutoring; reliability for relational instruction (i.e., the difference game) within tutoring lessons was 99.26%. In CAL, fidelity for lessons taught during the same timeframe as the WP difference unit (lessons 15 – 25) was 95.88% for whole-class instruction and 92.03% for tutoring.

Measures

For the present study, we assessed performance on difference problems and on students' understanding of relational terminology. (Neither of these measures was part of the larger investigation.) *Difference Problems*, which assesses performance on representing and solving difference problems, comprises 20 problems: six with the referent set unknown, six with the compared set unknown, four with the difference set unknown, and four with the difference set unknown with alternative wording (i.e., two were equalize problems and two were won't get problems). Additionally, two problems from each subtype include irrelevant information. Four problems (two referent set unknown and two compared set unknown) used an alternative presentation by placing the relational statement as the first sentence instead of the second sentence; this alters the expected word-problem format without changing the structure of the difference problem. The score is the number of correct answers with correct word labels, with a maximum score of 40. Alpha for this sample was .88. See Table 4 for sample problems.

Relational Tasks assesses understanding of relational terminology in the context of word problems. It comprises two activities. The first presents eight difference problems with the compared or referent sets unknown. Rather than solving problems, students determine which quantity is bigger and smaller and which quantity is unknown (i.e., the bigger, smaller, or difference amount). See Table 4 for a sample problem. The second activity comprises eight relational statements that are similar to those phrased in problems in which the compared set is missing. For each relational statement, students are instructed to determine which sentence preserves the described relationship. Students are given three sentences for which to choose along with a “none of the above” option. See Table 4 for a sample problem. The first activity assesses understanding of relational terminology as it directly relates to representing difference problems algebraically. The second activity assesses understanding of the symmetrical relationship of the terms *more* to *less/fewer*. Alpha for this sample was .91.

Testing Procedure

Pretesting for the present study occurred after winter break (week 7 of larger intervention) in one 50-minute testing session on Difference Problems and Relational Tasks. Posttesting on the same measures occurred at the completion of the difference unit, approximately seven weeks after pretesting (Week 14 of larger intervention). Students in all three conditions were pre- and posttested in the same timeframe.

RA teachers administered all pre- and posttests. Scripted protocols were provided to ensure tests were administered consistently. For difference problems, students were asked to solve each word problem; sample items were not provided. Students were directed to show their work; however, students were not prompted to use strategies from the WP intervention. For Relational Tasks, testers first explained and students completed sample items to ensure understanding of the task and how to mark answers. Students were given the opportunity to ask questions before testing began. For both measures, students were instructed not to work ahead of the tester so everyone worked on the same item at the same time. Testers read each problem twice, and advanced the class to the next item when all but two students were finished.

Table 4

Sample Items from Measures (Administered at Pre- and Posttest)

Relational Tasks

Task 1	Rachel drew 8 pictures. She drew 6 more than Carl. How many pictures did Carl draw? Who has the bigger amount? Write B over the person's name Who has the smaller amount? Write S over the person's name What are you trying to find? Circle bigger, smaller or difference bigger smaller difference
Task 2	Choose the sentence that means the same thing as: Nancy has 8 more lollipops than Jen. A Jen has 8 fewer lollipops than Nancy. B Nancy has 8 less lollipops than Jen. C Jen has 8 more lollipops than Nancy. D None of the above.

Difference Problems

Problem Subtype	*Word Problem*	*Algebraic Representation*	*Answer*
Unknown Difference Set (relational terminology)	Kesha has 7 pencils. Lynn has 3 pencils. How many more pencils does Kesha have than Lynn?	7 - 3 = X	4 pencils
Unknown Difference Set (Alternative Wording)	Mrs. Smith has 7 pencils. She has 9 students in her class. How many students won't get a pencil?	9 - 7 = X	2 students
Unknown Compared Set	Larry has 9 pencils. Carol has 5 less than Larry. How many pencils does Carol have?	9 - X = 5	4 pencils
Unknown Referent Set	Jay has 3 pencils. He has 5 fewer pencils than Kate. How many pencils does Kate have?	X - 3 = 5	8 pencils
Problem with Irrelevant Information	Sandy has 10 pencils that are 4 different colors. Rick has 2 less pencils than Sandy. How many pencils does Rick have?	10 - X = 2	8 pencils
Alternative Presentation	Mark has 8 fewer pencils than Drew. Mark has 4 pencils. How many pencils does Drew have?	X - 4 = 8	12 pencils

CHAPTER III

DATA ANALYSIS AND RESULTS

Table 5 displays pretest, posttest, and improvement scores on the outcome measure (Difference Problems) and the mediator (Relational Tasks) for the three conditions and across the CAL and control groups. Performance for problem subtypes and problem variants for the Difference Problem measure are also provided.

Preliminary Analyses

To assess whether we could combine CAL and control conditions to form one comparison group, we used regression analysis. We assessed Difference Problem outcomes, controlling for the pretest Difference Problem covariate and the hierarchical structure of the data. We assessed outcomes on the Relational Tasks, this time controlling for the Relational Tasks pretest covariate and the hierarchical structure of the data. These analyses revealed no significant differences between these two groups; therefore, we combined the CAL and control conditions to form one comparison condition.

To assess pretreatment comparability of conditions (WP vs. comparison), we used regression analysis that accounted for the hierarchical structure of the data. We assessed the effects of treatment condition on pretest scores for Difference Problems and Relational Tasks, which revealed a significant difference for Difference Problems (p = .02), with the WP condition scoring higher) but not for Relational Tasks (p = .83). To account for the difference between groups on the Difference Problem measure, we included the Difference Problem pretest as a covariate in all subsequent analyses.

Table 5

Pre- and Post Test Ouctome Measures

	WP		CAL		Control		Comparison CAL+Control		Across	
classrooms	(*n* = 12)		(*n* = 11)		(*n* = 8)		(*n* = 19)		(*n* = 31)	
students	(*n* = 169)		(*n* = 155)		(*n* = 118)		(*n* = 273)		(*n* = 442)	
Measures	M	(*SD*)	M	(*SD*)	M	(*SD*)	M	(*SD*)	M	(*SD*)
Relational Tasks										
Pretest	15.46	(7.82)	15.16	(7.67)	15.13	(6.88)	15.15	(7.33)	15.26	(7.52)
Posttest	19.47	(8.53)	16.82	(7.35)	17.64	(7.18)	17.18	(7.28)	18.05	(7.85)
Improvement	4.01	(6.59)	1.66	(5.20)	2.52	(4.95)	2.03	(5.10)	2.79	(5.79)
Difference Problems										
Overall (20)										
Pretest	14.44	(9.25)	12	(7.19)	11.49	(7.81)	11.78	(7.46)	12.08	(8.28)
Posttest	16.65	(9.30)	12.63	(6.10)	12.05	(7.51)	12.38	(6.74)	14.02	(8.08)
Improvement	2.21	(9.51)	0.63	(6.15)	0.55	(5.53)	0.6	(5.88)	1.21	(7.51)
Unknown Difference Set (4)										
Pretest	2.72	(2.12)	2.31	(1.67)	1.98	(1.74)	2.17	(1.70)	2.38	(1.89)
Posttest	3.51	(2.29)	2.46	(1.61)	1.98	(1.66)	2.26	(1.65)	2.73	(2.01)
Improvement	0.78	(2.36)	0.15	(1.50)	0.00	(1.38)	0.09	(1.44)	0.35	(1.88)
Alternative Wording (4)										
Pretest	3.59	(2.02)	3.32	(1.52)	3.06	(1.56)	3.22	(1.54)	3.36	(1.75)
Posttest	3.75	(1.79)	3.38	(1.25)	3.37	(1.39)	3.38	(1.31)	3.52	(1.52)
Improvement	0.15	(2.17)	0.06	(1.55)	0.31	(1.36)	0.17	(1.47)	0.16	(1.77)
Unknown Compared Set (6)										
Pretest	4.65	(3.03)	3.83	(2.54)	3.83	(2.76)	3.84	(2.63)	4.14	(2.82)
Posttest	5.03	(3.22)	3.97	(2.10)	3.89	(2.81)	3.93	(2.43)	4.35	(2.81)
Improvement	0.38	(3.43)	0.13	(2.29)	0.06	(2.43)	0.10	(2.35)	0.21	(2.81)
Unknown Referent Set (6)										
Pretest	3.48	(3.05)	2.53	(2.46)	2.61	(2.68)	2.56	(2.55)	2.91	(2.79)
Posttest	4.37	(3.12)	2.82	(2.17)	2.81	(2.70)	2.81	(2.41)	3.41	(2.80)
Improvement	0.89	(3.35)	0.29	(2.21)	0.19	(2.22)	0.25	(2.21)	0.49	(2.71)
Irrelevant Information (6)										
Pretest	3.64	(3.13)	2.75	(2.42)	2.72	(2.65)	2.74	(2.51)	3.08	(2.80)
Posttest	4.54	(3.26)	2.97	(2.36)	2.77	(2.65)	2.88	(2.49)	3.52	(2.92)
Improvement	0.9	(3.33)	0.22	(2.15)	0.05	(2.17)	0.15	(2.16)	0.44	(2.69)
Alternative Presentation (4)										
Pretest	2.71	(1.91)	2.12	(1.64)	2.18	(1.70)	2.15	(1.66)	2.36	(1.78)
Posttest	2.98	(2.02)	2.34	(1.26)	2.11	(1.75)	2.24	(1.50)	2.52	(1.75)
Improvement	0.27	(2.16)	0.22	(1.51)	-0.07	(1.43)	0.10	(1.48)	0.16	(1.77)

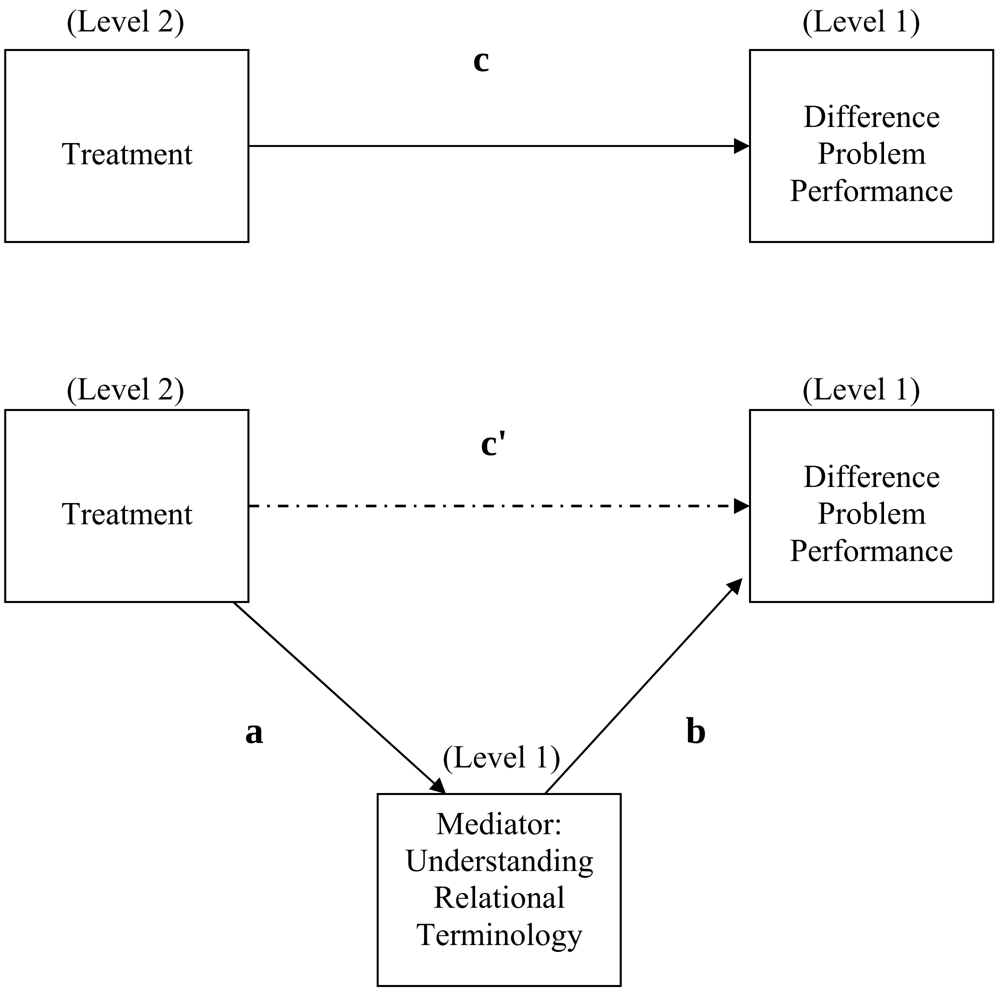
Note. The numeral next to each for Difference Problems is the number of test items for each problem category.

Assessing the Proposed Causal Mechanism

To test whether intervention effects were mediated by students' understanding of relational terminology, we conducted a mediation analysis in three steps (MacKinnon, 1994; MacKinnon, 2008; Zhang, Zypher, & Preacher, 2009) using regression analysis while accounting for the hierarchical structure of the data. Treatment was a Level 2 (classroom) variable and the mediator (Relational Tasks) along with the outcome (posttest Difference Problems) were Level 1 (student) variables. Figure 1 displays the mediation model we tested along with relevant paths we used to assess the direct and indirect effects.

In the first step of the mediation analysis, we assessed the effects of treatment condition on the outcome (posttest Difference Problem performance) controlling for pretest Difference Problem performance (path c on Figure 1). The first model on Table 6 shows treatment condition was a significant predictor of Difference Problem performance. In the second step, we assessed the effects of treatment on the mediator (posttest Relational Tasks). The second model on Table 6 shows treatment condition was a significant predictor of the mediator (path a on Figure 1). In the third and final step of the mediation analysis, we assessed the effects of treatment condition and the mediator (posttest Relational Tasks) on the outcome (posttest Difference Problems), controlling for Difference Problem prestest performance (paths b and c'). The third model shows treatment condition and the mediator each remain significant when both are included in the model (paths b and c'). Thus, understanding relational terminology did not fully mediate intervention effects. To determine whether the indirect effect was however significant, we applied the Sobel test (Baron & Kenny, 1986), which was significant (p = .048), indicating that understanding relational terminology partially mediates intervention effects on difference word problems.

Figure 1 *Mediation Model*



(*Figure 1*) Mediation Model Tested. The first panel shows the direct effect (path c) from the first step in the mediation analysis. The second panel shows the second (path a) and third steps (paths b and c') in the mediation model (based on Zhange, Zypher, & Preacher, 2009). Treatment is a Level 2 (classroom) variable and the mediator (Relational Tasks) and the outcome (Difference Problems) are Level 1 (student) variables.

Table 6

Mediation Analysis to test the Proposed Causal Mechanism

Model		*B*	*SE*	*p*	PC
Step 1	Treatment on Difference Problems posttest				
	Constant	16.42	0.78	< 0.001	
	Difference Problem (pretest)	0.58	0.05	< 0.001	
	Treatment	-3.92	0.93	< 0.001	c
Step 2	Treatment on Relational Tasks posttest				
	Constant	19.47	0.83	< 0.001	
	Treatment	-2.29	1.07	0.032	a
Step 3	Treatment and Relational Tasks on Difference Problems posttest				
	Constant	5.03	0.97	< 0.001	
	Difference Problem (pretest)	0.38	0.06	< 0.001	
	Relational Tasks posttest	0.32	0.06	< 0.001	b
	Treatment	-2.54	0.86	0.003	c'

Note . PC = the path coefficient relevant to the mediation analysis in Figure 1. To assess the significance of the indirect effect, we looked at paths a*b to assess the interaction.

CHAPTER IV

DISCUSSION

The purpose of the present study was to extend research on SBI for simple word problems by focusing on the most difficult problem type, difference problems (Riley & Greeno, 1988). We looked at the effects of WP intervention (based on SBI) against active and inactive competing conditions. The active condition, CAL, controlled for researcher-designed and delivered mathematics instruction that was structured similarly to WP by including the same amount of whole-class and tutoring instruction we added to the curriculum. The inactive condition controlled for history and maturation effects. We combined these two conditions to serve as one comparison group for three reasons. First, both conditions received similar word-problem instruction (as evidenced by teacher self reports), which did not include SBI. Second, there were no significant effects on Difference Problems or Relational Tasks between the CAL and control conditions. Third, including a group with only eight classrooms decreases power to detect a minimum effect size (*ES*) of 0.83, whereas having two groups with at least 12 classrooms (12 WP and 19 comparison), achieves the power to detect a minimum *ES* of 0.66.

Our primary goal was to assess whether students' understanding of relational terminology, which reflects an innovative instructional component of the intervention, mediates WP intervention effects. This relational terminology instruction was situated within the unit on difference problems because understanding relational terminology is most relevant for this problem type (Lewis & Mayer, 1987; Stern, 1993). We proposed a causal mechanism in which we anticipated students' understanding of relational terminology increases performance on difference problems. We posited that when students receive explicit instruction on the meanings of relational terminology, they will understand the relationships in difference problems better. This improved understanding of relational terminology will in turn increase students' ability to set up solution procedures and solve difference problems correctly. Research indicates that students who lack understanding of relational terminology have difficulty solving difference problems (Fan et al., 1994; Fuson et al., 1996; Stern, 1993); yet, to our knowledge, no prior studies have assessed how including instruction on understanding relational terminology within intervention affects performance.

In the first step of the mediation analysis, we found significant effects for WP intervention on difference problem outcomes. The WP intervention, based on SBI, includes several instructional components that likely contributed to students' increased performance. For example, students were taught to understand the underlying structure of the problem type, with the goal of fostering students' comprehension of the story structure, and to name the problem type before applying solution procedures. We also taught students to apply specific procedures for identifying the bigger, smaller, and difference quantities and to identify which of these quantities is unknown and asked for in the question. In addition, we taught students to represent the problem algebraically and to solve problems by applying solution rules. Recognizing and naming the problem type before solving it is central to applying the correct procedures for word problems in SBI. As with the present study, prior randomized control studies have found similar results for SBI when assessing performance for total, change, and difference problems together (Fuchs, et al., 2008; Fuchs, et al., 2009; Fuchs, et al., 2010; Jitendra et al., 1998; Jitendra, Griffin, Deatline-Buchman et al., 2007; Jitendra, Griffin, Haria, et al., 2007); however, this is the first study demonstration of effects specifically on difference problems, the most difficult of the three problem types.

In the second step of the mediation analysis, we found significant effects for WP intervention on understanding relational terminology. Within the difference unit, we incorporated an instructional component that taught students the meanings of relational terms and the symmetry between *more* and *less/fewer*. Within whole-class instruction, students learned to look at relational statements and determine which amount is bigger or smaller by focusing on the relational term in the sentence. Then students learned to write equivalent relational statements by changing the relational term and switching the subject and object of the sentences. In addition, tutored students learned to find the relational statement within a difference problem, to write the greater than (>) or less than (<) symbol over the relational term, to write "B" and "s" above the bigger and smaller amounts, and to determine which quantity is unknown (i.e., to determine where to position X). Via these activities within the whole-class and tutoring settings, students received explicit instruction on understanding relational terminology within the SBI program. Prior research has assessed students' understanding of relational terminology (e.g., Stern, 1993), showing that understanding relational terminology is important when solving difference problems with unknown referent sets. Yet, instruction on relational terminology has not been the focus of previous work. Our findings add to the existing literature by suggesting that instruction on relational terminology enhances students' understanding of relational terminology within the context of difference problems.

The third step in the mediation analysis provides support, at least in part, for our proposed causal mechanism. In this step of the analyses, we simultaneously included treatment and the relational understanding mediator variable as predictors of different problem outcome performance. This showed that WP intervention and the mediator each remained significant predictors (paths b and c') when included in the same regression equation. To assess the indirect effect, we applied the Sobel test, which tests the interaction between the effect of intervention on the mediator and the mediator effect on the outcome (paths a and b on Figure 1; see Table 6 for coefficient values). The Sobel test was significant, showing that intervention effects were partially mediated by understanding relational terminology. According to MacKinnon (2008), conducting mediation analysis identifies whether specific program components are successful, and in this case, the partial mediation of understanding relational terminology suggests this instruction is one component of SBI for difference word problems that determines efficacy.

The present study's findings suggest teaching students to understand relational terminology had a positive impact on performance for difference problems, with the major focus on the meanings of *more*, *less*, and *fewer* and determining bigger and smaller quantities. Explicit instruction on other relational terms (e.g., *longer*, *shorter*, *older*, *younger*) was not included in whole-class instruction at all and was incorporated into only three days of tutoring instruction. Moreover, our outcome and mediator measures only included items focused on *more*, *less*, and *fewer* and determining bigger and smaller quantities. In future research, it would be interesting to assess whether instruction on this limited set of relational terms transfers to a broader set of vocabulary.

In terms of implications for schools, our findings suggest word-problem instruction that teaches students to comprehend word problems is important. In our WP intervention on difference problems, we relied on SBI, which teaches students to read and interpret the story structure before solving the problem, while focusing students' attention on understanding relational terminology. By contrast, word-problem instruction that removes the need to comprehend word problems, such as keyword instruction, on which CAL and control teachers reported heavy reliance, may disadvantage students, especially on more difficult problem types. For difference problems, keyword instruction would result in incorrect answers for half of the six subtypes (problems with unknown referent sets when *more* or *less* are used or unknown difference sets when the word *more* is used). Previous research shows that students relying on keywords will appear to comprehend difference problems with unknown compared sets because they applied correct solution procedures, even though they may not understand them (Stern, 1993; Verschaffel et al., 1992). Therefore, instruction should instead focus on improving students' understanding the underlying meaning of word problems.

Before concluding, we note our study's limitations. First, we relied on teacher self reports to ascertain the amount of time they allocated for word-problem instruction and the instructional methods they used. Based on these reports, students in the WP condition received more word-problem instruction than students in the CAL or control conditions. Unfortunately, we did not ask teachers to report the amount of time devoted to calculation instruction or to mathematics instruction in general. Because the CAL condition was an active control group, we know the amount of researcher-delivered mathematics instruction was equivalent between CAL and WP. Also, we assume that CAL teachers allocated additional instructional time of their own to calculations and that total mathematics instructional time between CAL and WP was therefore similar. Nevertheless, we recognize that students in the WP condition received more instruction specifically on word problems; therefore, we cannot dismiss the possibility that their superior learning is, at least in part, due to additional instructional time. At the same time, the mediation analyses, demonstrating that our instructional focus on relational terminology was a partial mediator of that learning, substantiates the effects of the WP instruction beyond a simple addition of instructional time. A second limitation is that, as part of the larger investigation, the difference unit occurred after seven WP instructional weeks had occurred. In this way, students in the WP condition had already enjoyed the benefits of SBI within an introductory unit as well as the unit focusing on the total problem type. This was reflected in the superior pretest performance for the WP students. To control for this difference, we included pretest as a covariate; however, it remains the case that the WP students began the present study performing higher on difference problems, with an ES of 0.32 compared to the combined contrasting condition. A third limitation is that we did not assess maintenance effects. Future work should include additional maintenance measures to assess whether intervention effects sustain over time. These limitations not withstanding, results suggest that SBI promotes learning of difference problems, the most difficult problem type addressed in the primary grades. Moreover, findings suggest that understanding of relational terminology partially mediates the effects of SBI on difference problem learning and that incorporating an instructional focus on relational terminology within SBI may afford added value.

REFERENCES

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173 – 1182.

Briars, D. J. & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. *Cognition and Instruction, 1*, 245-296.

Carpenter, T. P. & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal of Research in Math Education, 15*, 179 – 202.

Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem solving transfer. *Journal of Educational Psychology, 79,* 374-362.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*, 405 – 438.

Fan, N., Mueller, J. H., & Marini, A. E. (1994). Solving difference problems: Wording primes coordination. *Cognition and Instruction, 12*, 355 – 369.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., Hamlett, C. L., & Zumeta, R. O. (2009). Remediation number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. Journal of Educational Psychology, 101, 561 – 576.

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children, 74*, 155 – 173.

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. *Elementary School Journal, 110*, 440 – 463.

Fuson, K. C. & Carroll, W. M. (1996). Levels in conceptualizing and solving addition and subtraction compare word problems. *Cognition and Instruction, 14*, 345 – 371.

Garcia, A. I., Jimenez, J. E., & Hess, S. (2006). Solving arithmetic word problems: An analysis of classification as a function of difficulty in children with and without arithmetic LD. *Journal of Learning Disabilities, 39*, 270 – 281.

Gick, M.L., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: evidence from students' eye fixations. *Journal of Educational Psychology, 84*, 76 – 84.

Jitendra, A. K., Griffin, C. C., Deatline-Buchman, A., & Sczesniak, E. (2007). Mathematical word problem solving in third-grade classrooms. *Journal of Educational Research, 100*, 283 – 302.

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' mathematical problem solving. *Journal of Educational Psychology, 99*, 115 – 127.

Jitendra, A. K., Griffin, C., McGoey, K., Gardill, C, Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk or with mild disabilities. *Journal of Educational Research, 91*(6), 345-356.

Lewis, A. B. & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology, 79*, 363 – 371.

MacKinnon, D. P. (1994) Analysis of mediating variables in prevention and intervention research. Eds. Cezares, A. & Beatty, L. A. *National Institute on Drug Abuse Research Monograph Research, 139*, 132 – 159. Washington DC, Public Institutes of Health.

MakKinnon, D. P. (2008). Introduction to statistical mediation analysis. New York: Lawrence Erlbaum and Associates.

Morales, R. V., Shute, V. J., & Pellegrino J. M. (1985). Developmental differences in understanding and solving simple word problems. *Cognition and Instruction, 2*, 41 – 57.

National Council of Teachers of Mathematics (2006). www.nctm.org.

National Assessment of Educational Progress (2009) http://nces.ed.gov/nationsreportcard/

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2009). Do word-problme features differentially affect problem difficulty as a function of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities, 42*, 99 – 110.

Riley, M. S. & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*, 49 – 101.

Riley, M. S., Greeno, J. G., & Heller, J. H. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp.153 – 196). San Diego, CA: Academic Press.

Stern, E. (1993). What makes certain arithmetic word problems involving the comparison of sets so difficult for children? *Journal of Educational Psychology, 85*, 7 – 23.

The Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: Harcourt Brace & Company.

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*, 85 – 94.

Wolters, C. A. & Daugherty, S. G. (2007). Goal structures and teachers' sense of self-efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99*, 181 – 193.

Zhange, Z., Zypher, M. J., & Preacher, K. J. (2009) Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods, 12*, 695 – 719.