2009

The influence of rewording and gesture scaffolds on the ability of first graders with low language skill to solve arithmetic word problems

Vicki Marie Samelson
University of Iowa

THE INFLUENCE OF REWORDING AND GESTURE SCAFFOLDS ON THE ABILITY OF FIRST GRADERS WITH LOW LANGUAGE SKILL TO SOLVE ARITHMETIC WORD PROBLEMS

by

Vicki Marie Samelson

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Speech and Hearing Science
in the Graduate College of
The University of Iowa

May 2009

Thesis Supervisor: Professor J. Bruce Tomblin

ABSTRACT

Purpose: This study examined the relationship between arithmetic word problem solving skill in first graders and 1) their oral language skill, 2) their nonverbal understanding of mathematical sets, and 3) rewording and gesture scaffolds designed to help the children access both the linguistic and the nonverbal content of *Compare 6* word problems.

Method: Two groups of first graders (15 with good oral language skill and 15 with low oral language skill) solved a matched set of verbal and nonverbal arithmetic problems, followed by three types of *Compare* word problems. Twenty first graders with low oral language skill (9 with low normal language (LN) and 11 with a diagnosis of language impairment (LI)) then solved orally-presented *Compare 6* word problems under 4 scaffold conditions: 1) traditional wording, 2) traditional wording + gesture, 3) rewording, and 4) rewording + gesture.

Results: Children with low oral language skill had greater difficulty solving orally-presented arithmetic word problems than their peers with good language skill, but performed comparably on a nonverbal arithmetic task. Using proportion of problems solved correctly, rewording *Compare 6* word problems was facilitative for the LN group but not for the LI group. Changing the problem wording from a *Compare 6* to a *Compare 3*, by using ‘more than’ instead of ‘fewer than’ and by eliminating pronoun anaphora, resulted in comparable performance to rewording that also included a rationale, optional verbs and placing the question first. The gesture scaffold was marginally significant for both groups.

Conclusions: The LI group did not benefit from implicitly-presented rewording or gesture scaffolds; the LN group did benefit from the rewording scaffold. The gesture scaffold was marginally facilitative despite the finding that children with low oral language skill were able to access nonverbal information in a nonverbal arithmetic task.

Empirical and anecdotal evidence suggested that, for a number of these children, rewording and gesture scaffolds altered their mental model of the word problem structure. This altered representation resulted in the use of different solution strategies. The new strategies, however, were not always correct. Implications for classroom intervention and suggestions for future research are discussed.

Abstract Approved: ______________________________
Thesis Supervisor

____
Title and Department

____
Date

THE INFLUENCE OF REWORDING AND GESTURE SCAFFOLDS ON THE ABILITY OF FIRST GRADERS WITH LOW LANGUAGE SKILL TO SOLVE ARITHMETIC WORD PROBLEMS

by

Vicki Marie Samelson

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Speech and Hearing Science
in the Graduate College of
The University of Iowa

May 2009

Thesis Supervisor: Professor J. Bruce Tomblin

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Vicki Marie Samelson

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Speech and Hearing Science at the May 2009 graduation.

Thesis Committee: ___________________________________
J. Bruce Tomblin, Thesis Supervisor

____
Martha W. Alibali

____
Susan W. Cook

____
Richard R. Hurtig

____
Karla K. McGregor

____
Amanda J. Owen

"The difficulty with defining the term *problem* is that problem solving is relative. The same tasks that call for significant efforts from some students may well be routine exercise for others, and answering them may just be a matter of recall for a given mathematician. Thus being a "problem" is not a property inherent in a mathematical task. Rather, it is a particular relationship between the individual and the task that makes the task a problem for that person."

A. H. Schoenfeld
Mathematical Problem Solving

ACKNOWLEDGMENTS

My doctoral thesis was the product of course work, research experience, and conversations with many individuals. First, to the members of my committee:

Bruce, your enthusiasm for learning and following new threads immensely enriched my doctoral experience and I extend a sincere thank you to you for mentoring my foray into the early stages of intervention research. The combination of your broad knowledge base and your ability to view the big picture without losing track of the details helped me process and organize the diverse literatures I chose to explore. As a result of our many conversations, I experienced countless synapse-strengthening ‘aha’ moments that have shaped and will continue to shape my research and my teaching.

Martha, hearing your students present their work at SRCD – Boston in 2007 influenced my thesis topic and opened the world of gesture to me. Your insightful comments guided my hypotheses and predictions, and will influence my future research.

Thank you, Susan, for joining my committee on short notice. Your gesture research influenced my hypotheses and your suggestions will shape my future research.

Richard, thank you for introducing us to the intersection of psycholinguistics and language intervention. I can’t tell you how many times I have referred back to the fantastic literature base that you provided during our Psycholinguistics seminar.

Karla, I appreciated being able to explore scaffolding and the work of Vygotsky in your seminar, where our readings and conversations influenced my dissertation hypotheses. My Scientific Writing notes are also never far from my writing desk.

Amanda, my experiences in your lab helped me learn how to develop my ideas, do disciplined research, recruit and test participants, and mentor undergraduate RAs. I enjoyed our conversations about theories, study designs, and research results. Thank you also for reading and critiquing my writing and challenging me to express my ideas and results more clearly and succinctly.

Two other professors at the University of Iowa also had a significant impact on my thinking. Nancy Jackson, you introduced me to the world of arithmetic word problems during our readings on reading comprehension. Through your seminars I was able to explore the ideas of Robbie Case, Piaget, Robert Siegler, Michael Pressley, and Walter Kintsch. Their work has shaped my research and my teaching. Thank you also for reading our weekly reaction papers and guiding us in developing the skills of writing exactly what we are thinking, and representing others' ideas accurately.

And finally, Bob McMurray, thank you for patiently guiding me into the world of online spoken word recognition. Our results have influenced the way I conceive of language processing and language disorders.

What an incredible group of researchers and professionals. Each of you mentored me in a way that allowed me to become who I wanted to be, while you guided me to become who I needed to be. That is the mark of a positive doctoral and dissertation experience! And last, but not least:

Mom and Dad, by your example and guidance you instilled a strong work ethic, a love of learning, and a trust in a higher power in your children.

Bret, it has been wonderful sharing the world of academia and research with 'our favorite son', and I look forward to many more conversations with you and Dad as you progress along your career path.

And finally, Scott, I simply couldn't have undertaken the past five years without your loving support and encouragement. I'm where I am today because of you!

ABSTRACT

Purpose: This study examined the relationship between arithmetic word problem solving skill in first graders and 1) their oral language skill, 2) their nonverbal understanding of mathematical sets, and 3) rewording and gesture scaffolds designed to help the children access both the linguistic and the nonverbal content of *Compare 6* word problems.

Method: Two groups of first graders (15 with good oral language skill and 15 with low oral language skill) solved a matched set of verbal and nonverbal arithmetic problems, followed by three types of *Compare* word problems. Twenty first graders with low oral language skill (9 with low normal language (LN) and 11 with a diagnosis of language impairment (LI)) then solved orally-presented *Compare 6* word problems under 4 scaffold conditions: 1) traditional wording, 2) traditional wording + gesture, 3) rewording, and 4) rewording + gesture.

Results: Children with low oral language skill had greater difficulty solving orally-presented arithmetic word problems than their peers with good language skill, but performed comparably on a nonverbal arithmetic task. Using proportion of problems solved correctly, rewording *Compare 6* word problems was facilitative for the LN group but not for the LI group. Changing the problem wording from a *Compare 6* to a *Compare 3*, by using 'more than' instead of 'fewer than' and by eliminating pronoun anaphora, resulted in comparable performance to rewording that also included a rationale, optional verbs and placing the question first. The gesture scaffold was marginally significant for both groups.

Conclusions: The LI group did not benefit from implicitly-presented rewording or gesture scaffolds; the LN group did benefit from the rewording scaffold. The gesture scaffold was marginally facilitative despite the finding that children with low oral language skill were able to access nonverbal information in a nonverbal arithmetic task.

Empirical and anecdotal evidence suggested that, for a number of these children, rewording and gesture scaffolds altered their mental model of the word problem structure. This altered representation resulted in the use of different solution strategies. The new strategies, however, were not always correct. Implications for classroom intervention and suggestions for future research are discussed.

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xi

CHAPTER I INTRODUCTION AND RATIONALE 1

Comprehension in the Classroom Setting 1
The Study of Arithmetic Word Problem Skill 3
The Development of Arithmetic Word Problem Skill 3
Scaffolding Arithmetic Word Problem Solving 4
Rewording the Problem 5
Providing a Gesture Scaffold 6
Summary 7

CHAPTER II THE CURRENT STUDY 9

Phase One - Goals and Hypotheses 9
Phase Two - Goals and Hypotheses 9
Phase Three - Goals and Predictions 10

CHAPTER III REVIEW OF THE LITERATURE 13

Introduction 13
Development of Numerical Cognition 13
The Study of Arithmetic Word Problem Solving 14
Schema Representation 14
Structure and Relative Difficulty of Arithmetic Word Problem Types 16
Linguistic Factors 18
Contextualized Knowledge 21
Working Memory 23
Summary 25
Individual Differences in Numerical Cognition: Language Impairment 26
Summary of Mathematical Skill in Children with Language Impairment 33
Gesture Scaffolds 34
Gesture Use 34
Access to Teachers' Gestures 35
Gesture and Language Impairment 38
Gesture Use 38
Access to Teachers' Gestures 39
Summary and Questions 40

CHAPTER IV PHASE ONE 42

Method 42
Participants 42
Language Skill 43
Age and Nonverbal Matching 44
Additional Participant Criteria 45
Procedure and Tasks 45

Visual Memory Task 46
Auditory Memory Task 46
Verbal and Nonverbal Arithmetic Problems 46
Nonverbal Arithmetic Task 47
Verbal Word Problem Task 49
Coding and Analysis of Verbal and Nonverbal Problem Accuracy 49
Reliability 50
Analyses 50
Results 50
Verbal and Nonverbal Arithmetic Problem Tasks 50
Counterbalancing 51
Potential Influencing Factors 53
Discussion 53

CHAPTER V PHASE TWO 56

Method 56
Participants 56
Task and Procedure 57
Basic Arithmetic Word Problem Task 57
Procedure 59
Coding and Analysis of Problem Accuracy 60
Reliability 61
Analyses 61
Results 61
Solution Accuracy for Three Word Problem Types: All Participants 61
Poor Language User Performance: Three Problem Types 63
Good Language User Performance: Three Problem Types 64
Filler Word Problems: All Participants 64
Poor Language User Performance: Filler Word Problems 65
Two Sets of Number Facts 66
Potential Influencing Factors 66
Discussion 67

CHAPTER VI PHASE THREE 69

Method 69
Participants 70
Selection of Problem Type 70
Stimulus Creation 73
Rewording Scaffold 73
Gesture Scaffold 73
Compare 6 Word Problem Task 75
Procedure 76
Coding and Analyses of Problem Accuracy 77
Reliability 77
Analyses 78
Results 78
Solution Accuracy 78
Post hoc Blocking Factor 79
Group by Rewording Interaction 82
Two Sets of Number Facts 84
Gesture Order Effect 85
Practice and Learning Effects 85

‘Question First’ Rewording versus Traditional Wording 86
Potential Influencing Factors 87
Discussion 88
Response to Scaffolds 88
Group Differences in Response to Scaffolds 89
Low Normal Language Skill Group 89
Language-Impaired Group 90
Performance Above Chance 90
Other Implications of the Rewording Scaffold Results 91
Marginal Effect of the Gesture Scaffold 92
Other Potential Influencing Factors 95
Zone of Proximal Development 96
Clinical Implications 97

CHAPTER VII SUMMARY AND FUTURE DIRECTIONS 99

Summary 99
Future Directions 100

APPENDIX A ARITHMETIC WORD PROBLEM TYPES 103

APPENDIX B LANGUAGE MEASURES INCLUDED IN THE LANGUAGE COMPOSITE SCORE (LCOMP) 104

APPENDIX C PHASE ONE: VERBAL AND NONVERBAL ARITHMETIC PROBLEM SETS 105

APPENDIX D PHASE TWO: WORD PROBLEM SETS 110

APPENDIX E PHASE THREE: WORD PROBLEM SETS 118

REFERENCES 126

LIST OF TABLES

Table

1. Background measures for participants in Phase One....43

2. Means and standard deviations for the verbal and nonverbal arithmetic problem solving tasks in Phase One. 51

3. Background measures for participants in Phase Two. 57

4. Means and standard deviations for the four problem types in Phase Two. 62

5. Background measures for participants in Phase Three. 71

6. Means and standard deviations for the four scaffold conditions for all 20 participants in Phase Three 79

7. Means and standard deviations for the four scaffold conditions with a post hoc blocking factor 82

LIST OF FIGURES

Figure

1. Four Scaffolding Conditions in Phase Three. 11

2. Nonverbal Arithmetic Task (Problem: 5–3=2). 47

3. Graph of Verbal and Nonverbal Arithmetic Problem Tasks. 52

4. Word Problem Task (Phases Two and Three). 60

5. Proportion Correct for the Four Problem Types in Phase Two. 63

6. Example of the Rewording and Gesture Scaffolds for a Compare 6 Problem. 74

7. Overall Proportion Correct for the Four Scaffold Conditions in Phase Three. 80

8. Proportion Correct for Rewording and Gesture Scaffolds. 83

9. Proportion Correct for the Four Scaffolding Conditions in Phase Three. 86

CHAPTER I

INTRODUCTION AND RATIONALE

The uniquely human use of language provides a significant advantage in facilitating conspecific communication, cultural interaction, and transmission of knowledge. As human cultures have advanced, awareness of the importance of knowledge transmission has resulted in the development of formal systems for recording the content and form of human interactions. More recently, structured academic settings with their formalized math, reading, and subject matter textbooks serve as vehicles for human learning. One of the challenges faced by a young child as he enters school is to integrate his existing knowledge base and current levels of cognitive and language functioning with the structured language format of the formal educational system (Justice, 2008). Although we have evidence that children with language deficits struggle with reading, writing and math in school (Cowan, Donlan, Newton & Lloyd, 2005; Donlan, 1998b; Fazio, 1999; Fey, Catts, Proctor-Williams, Tomblin & Zhang, 2004) and have poorer academic outcomes than typically-developing children (Catts, Fey, Tomblin & Zhang, 2002; Tomblin, Zhang, Buckwalter & O'Brien, 2003; Tomblin, 2005), we know little about the influence of oral language skill on children's ability to comprehend the language that they hear in primary grade classroom settings.

Comprehension in the Classroom Setting

Language demands change as a child moves from preschool into elementary school age. During the toddler and preschool years conversational language serves as the vehicle for social interaction, informal knowledge acquisition, and informal problem solving. Upon entering the formal education system, however, information is often presented in a decontextualized, expository language format. Expository language, which differs in structure from the conversational and narrative formats that preschool children are familiar with, becomes the classroom medium by which children are required to comprehend auditory and written information and to engage in problem-solving activities (Culatta & Wiig, 2006; Paul, 2006).

In order to function optimally in a classroom a child must be able to make sense of the language surrounding him, be it written or oral. This comprehension process involves a litany of skills, including attending selectively to a linguistic signal, creating a mental representation of that input, and integrating this representation with existing knowledge in order to engage in problem solving activities. Thus, in a formal education environment, a child is faced with challenging language comprehension tasks and novel problem solving tasks that are likely to require considerable language support.

The most prominent place where research and practice has acknowledged the role of language in classroom learning has been in the area of reading. Initially, researchers focused on the relationship between phonological skills and word decoding abilities (e.g. Perfetti, 1999; Share & Stanovich, 1995). More recently, as scholars have begun to differentiate the processes involved in word decoding from those used for text comprehension, the importance of higher level language skills for reading comprehension has surfaced (Catts, Hogan & Fey, 2003; Hoover & Gough, 1990; Nation & Norbury, 2005). Within much of the research on reading, however, the emphasis has been on the comprehension of text material for the construction of meaning that is contained in the text itself or the implicature allowed by the text (Kintsch & Greeno, 1985; Riley, Greeno & Heller, 1983).

There are, however, other important classroom language comprehension tasks that confront the child with an explicit problem that requires a solution. These tasks may be viewed as a special kind of discourse that requires that the child bring to the text or the listening comprehension process some particular world knowledge having to do with the problem itself, in conjunction with particular problem solving schemas and strategies. This type of classroom learning task is well exemplified in activities involving basic arithmetic word problems.

The Study of Arithmetic Word Problem Skill

Over the past 20 years, researchers have focused on how children develop formal mathematical skill, especially in light of declining performance by students in the United States and Britain on math achievement tests (Cummins, Kintsch, Reusser & Weimer, 1988, p. 405; Klein, 2003; Verschaffel et al., 1999; Woodward, 2004). In an attempt to explain the developmental progression of arithmetic word problem skill in children, a number of studies have explored the role of linguistic factors in children's ability to solve basic word problems (Cummins et al., 1988; Cummins, 1991; Fuson, Carroll & Landis, 1996). Despite the emphasis on linguistic factors, however, none of these studies specifically considered the oral language skill levels of their young participants.

In addition to studying typical development of cognitive and procedural mathematical knowledge, researchers currently are emphasizing individual and intra-individual differences in the component skills of arithmetical cognition (e.g. Donlan, 1998a, pp. 197-199; Dowker, 2005; Siegler & Araya, 2005). Despite this focus on individual differences, we know little about the ability of children with low oral language skill to solve basic arithmetic word problems in a formal listening environment.

The Development of Arithmetic Word Problem Skill

The ability to solve an arithmetic word problem consists of a language-mediated calculation component, a strategy-based procedural component, and knowledge of the 'relationships between sets' (possibly based in nonverbal cognition), all of which are couched in a language comprehension process. For the purposes of this study, the term 'sets' refers to the real quantities that a child must conceptualize and manipulate in order to solve a word problem (e.g. in the problem "Garfield has 3 candy bars. He gave 2 of his candy bars to Goofy. How many candy bars does Garfield have left?" a child must conceptualize the sets '3' and '2' and then manipulate the relationship between these sets in order to solve the problem correctly).

Griffin (2003) advocates that in order for children to succeed in *formal* math education they must come to the table with a basic conceptual structure in place that includes the realms of 1) real quantities, that is sets of objects and the relations between those sets, and 2) verbal number labels for these real-world quantities. Formal education must provide opportunities for children "to connect their understanding of quantity to their understanding of number…" (p. 22). Griffin suggests that a breakdown in this connecting process could happen if the curriculum focuses too much on mathematical equations and the verbal 'number' counting system, and not enough on 'quantity' and set relations. Although Griffin emphasized the role of external factors (i.e. classroom language and the curriculum) in a child's ability to succeed in formal math education, an individual child could also experience difficulty comprehending the formal linguistic components of the mathematical problem, as a separate process from knowledge of quantity and knowledge of a verbal counting system. In arithmetic word problems the linguistic component (comprised of the verbal count system and the wording of the problem) interacts with the underlying problem structure (i.e. the child's conceptual knowledge of numerical sets and how those quantities are related) as well as with an individual child's oral language skill. Thus, a breakdown in the development of arithmetic word problem solving skill could arise from two sources: 1) How well the curriculum and teacher can help a child connect the linguistic component of the problem with the underlying problem structure, and 2) how well an individual child can process both the verbal component and the underlying problem structure.

Scaffolding Arithmetic Word Problem Solving

When children with low language skills have difficulty with word problem solving tasks, speech-language pathologists and educators face the challenge of designing interventions that scaffold the linguistic component so that the child can access the meaning of the underlying problem structure. According to Mayer (2003), when solving arithmetic word problems a young child must process the linguistic message, determine the underlying problem structure, select a solution strategy, and execute the solution. If the linguistic message functions as a bottleneck to word problem representation, it should be possible to reduce this bottleneck by scaffolding the linguistic and the conceptual problem structure components of the problem. Two potential scaffolds are: 1) rewording the problem to make the linguistic message more explicit and available to the child, and 2) providing a gesture scaffold to help the child access the underlying problem structure (the relationships between sets of numbers).

Rewording the Problem

A number of studies have shown that rewording a traditional arithmetic word problem in order to mitigate the effects of specific lexical bottlenecks (e.g. '*some*' or '*altogether*') (Cummins, 1991) or to make the problem text more explicit (Hudson, 1983; Decorte, Verschaffel & Dewin, 1985) facilitates problem solution in typically-developing children. In other work, Thevenot, Devidal, Barrouillet, and Fayol (2007) demonstrated that placing the question first in a word problem resulted in better performance in children with poor mathematical ability.

There is little evidence, however, as to whether explicit rewording or placing the question first will facilitate word problem representation and solution in children with low oral language skill. It is possible that other factors could mitigate the positive scaffolding effect of rewording the word problems for children with lower language skills. Prior research has demonstrated that these children may have concomitant deficits in attention and working memory processes that could in turn influence their ability to focus on and retain information presented orally (Ellis Weismer, Evans & Hesketh, 1999; Montgomery, 2003; Spaulding, Plante & Vance, 2008).

Providing a Gesture Scaffold

A gesture scaffold could also facilitate comprehension of an arithmetic word problem. A growing body of literature suggests that typically-developing children are better able to understand a spoken message when that message is accompanied by meaningful gestures (Cook & Goldin-Meadow, 2006; Goldin-Meadow, Kim & Singer, 1999). Using mathematical equivalence problems, Church, Ayman-Nolley, and Vasich (2007) discovered that children who were exposed to a conceptual gesture (for the concept of 'equals') exhibited greater learning than children who were exposed to a simple 'beat' (emphasis) gesture. They concluded that the conceptual gesture provided a concrete image that helped bridge the gap between the verbal explanation and the concept of equality. McNeil, Alibali, and Evans (2000) explored the effects of conceptual gestures (that either reinforced or conflicted with the oral message) in an oral language comprehension task, and found that reinforcing gestures were most facilitative when the spoken message was more complex. In these studies, representation of information via multiple modalities seemed to improve typically-developing children's depth of conceptual knowledge; the conceptual gestures functioned as a mechanism of cognitive support and change.

Ellis-Weismer and Hesketh (1993) found that both typically-developing kindergarteners and kindergarteners with language impairment benefited from a conceptual gesture when asked to learn nonsense words for the location of a spaceman. It is not yet known, however, whether conceptual or procedural gestures are facilitative for children with low language skills when they are asked to solve an arithmetic word problem containing formal mathematical lexical items such as "more than" and "less than".

Because gestures are hypothesized to help bridge the gap between the linguistic message and underlying (possibly nonverbal) conceptual and/or procedural knowledge, it is also important to consider the influence of nonverbal skill on ability to solve arithmetic word problems. One body of research provides evidence that children with language impairment also exhibit nonverbal processing deficits (Botting, 2005; Johnston & Smith, 1989; Miller, Kail, Leonard & Tomblin, 2001). Other evidence suggests that nonverbal cognitive skill is a relative strength in children with language deficits (Donlan & Gourlay, 1999; Donlan & Newton, 2007; Jordan, Levine & Huttenlocher, 1995). Therefore, it remains to be seen whether a nonverbal conceptual or procedural gesture will facilitate comprehension of arithmetic word problems in children with low oral language skill.

Another potential barrier to a child's ability to make use of rewording and gesture scaffolds could be the child's ability to integrate the information received from the linguistic and nonverbal (i.e. gesture) channels. Because poor reading comprehenders have difficulty constructing complete and accurate mental models from the text that they are reading (Cain & Oakhill, 1999), children with low language skill might also construct less accurate mental representations of word problems that they hear. We would then expect that these children would benefit from a gesture scaffold that facilitates a more accurate mental model. However Cain and Oakhill also concluded that poor reading comprehenders have difficulty integrating new information within an existing representation, so it is possible that additional gestural information might overwhelm children with low oral language skill. The extent to which both gesture and problem rewording facilitate arithmetic word problem representation and solution has not yet been investigated in children with low oral language skills.

Summary

I contend that the task of arithmetic word problem solving draws substantially on language abilities. Children with low oral language skills therefore will be challenged by arithmetic word problems. However by restructuring the language provided in the word problem, and by tapping into nonverbal knowledge of set relations via conceptual and procedural gesture scaffolding, children with low language skill will be more successful in this task.

In this study I ask the following questions:

1) Do children with low oral language skill solve orally-presented arithmetic word problems as accurately as children with good oral language skill?

2) Do children with low oral language skill have access to an underlying nonverbal knowledge of mathematical set relations?

3) Will scaffolds designed to address a linguistic bottleneck allow children to access this nonverbal knowledge?

3) Specifically, will rewording arithmetic word problems make the linguistic message more explicit and available to children with low oral language skill, and increase solution accuracy?

4) Specifically, will providing a gesture scaffold help children with low oral language skill access the underlying problem structure (the relationships between sets of numbers), and increase solution accuracy?

CHAPTER II

THE CURRENT STUDY

This research project consisted of three phases, in which I examined the relationships between oral language ability, nonverbal understanding of mathematical sets, basic arithmetic word problem solving skill, and two types of intervention scaffolds (one targeting linguistic factors and the other targeting the nonverbal component).

Phase One – Goals and Hypotheses

In Phase One I compared good and poor language users' performance on basic addition and subtraction problems presented in both oral and nonverbal formats. A key assumption in this work is that children with low oral language skill have the capacity to solve basic arithmetic problems when there is reduced verbal demand, i.e. when the problems are presented in nonverbal format. Under the hypothesis that underlying nonverbal skill provides support for the development of mathematical cognition, I predicted that the two groups would have similar performance on a nonverbal arithmetic problem task, if the groups were matched on a nonverbal intelligence measure. Under the hypothesis that oral language deficits function as the primary bottleneck to accurate problem representation and solution when arithmetic problems are presented verbally, I predicted that, in Phase One, children with poor oral language skill would have greater difficulty solving orally-presented arithmetic word problems than children with good oral language skill.

Phase Two – Goals and Hypotheses

The goals of Phase Two were 1) to identify those first graders who had difficulty solving basic verbal arithmetic word problems (presented orally), and 2) to determine the specific problem types that were difficult for them to solve. The results of this phase determined which children would continue to Phase Three, and which of the three types of word problems would be presented with rewording and gesture scaffolds.

For all three problem types I predicted that children with low oral language skill would have greater difficulty than children with good oral language skill. No predictions were made regarding differences in difficulty of the three types of word problems within each of the groups, because of the exploratory nature of this phase.

Phase Three – Goals and Hypotheses

Whereas Phases One and Two examined the relationship between oral language skill and arithmetic problem solving within a quasi-experimental between subjects design, Phase Three explored the influence of linguistic and gesture scaffolds on the ability of first graders with low oral language skill to comprehend and solve oral word problems, using a within subjects repeated measures design. The goal of Phase Three was to evaluate potential classroom adaptations by investigating the effects of two types of scaffolds on arithmetic word problem solving ability in those children with low oral language skill who had difficulty solving the orally-presented problems in Phase Two: 1) a rewording scaffold, that was hypothesized to mitigate the influence of a linguistic bottleneck on children's ability to solve verbal arithmetic word problems, and 2) a gesture scaffold that was hypothesized to access and support the underlying knowledge of mathematical set relations of children with low language skills.

Assuming the theoretical perspective that ability to solve a given verbal problem accurately arises out of the process of comprehending a linguistic message and integrating the representation of this message with a mental model of the problem situation (in the case of arithmetic word problems, the nonverbal knowledge of set relations, e.g. Kintsch, 1998; Mayer, 2003), we have evidence that the resulting representation influences a child's choice of solution strategy. Choice of strategy can, in turn, affect the accuracy of the solution (for a review see Alibali, 2005). Inaccurate problem representation arising from a child's inability to comprehend the linguistic component of the problem could lead that child to choose less mature solution strategies and potentially inaccurate solutions. Because children tend to resort to preferred default strategies when not sure how to interpret a word problem (Cummins, 1991), providing facilitative scaffolds should improve problem comprehension and result in more effective solution strategies with fewer solution errors.

So in Phase Three, if problem solving in children with low oral language skill is constrained by their ability to understand the problem (for linguistic reasons, conceptual reasons, or both), I predicted that linguistic and gesture scaffolds would facilitate a more accurate representation of set relations in the underlying conception of the problem, and would result in increased solution accuracy. If, as hypothesized, low oral language skill functions as a bottleneck to accurate arithmetic word problem solution, I predicted that, in Phase Three, children with poor oral language skill would solve reworded arithmetic word problems (See Figure 1, Scaffold **B**) more accurately than problems presented with traditional, formal wording (Scaffold **A**).

	Wording	
Gesture	**A.** Traditional Wording with no Gesture (Baseline)	**B.** Reworded with no Gesture
	C. Traditional Wording with Gesture	**D.** Reworded with Gesture

Figure 1. Four Scaffolding Conditions in Phase Three.

If children with low oral language skill are able to access a nonverbal representation of the relations between sets (as hypothesized in Phase One by the similar performance of children with good and poor language skill on the nonverbal arithmetic problem set), I predicted that Scaffold **C** would also result in greater solution accuracy than Scaffold **A**. It is also possible that a combination of problem rewording and a gesture scaffold (Scaffold Condition **D**) would provide additional support for the underlying representation of set relations, if rewording alone (Scaffold **B)** is not sufficient. In this case, I predicted that Scaffold **D** would result in greater solution accuracy than Scaffold **A**, Scaffold **B**, and Scaffold **C**. Here, use of a gesture scaffold in combination with rewording of the problem would provide a redundant mechanism for actual change in the child's representation of the problem.

On the other hand, if children with low language skill have difficulty integrating multiple modalities (i.e. verbal and gesture), I predicted that the gesture scaffolds (**C** and **D**) would not be more facilitative than the problem rewording condition (Scaffold **B**). In this case, the addition of a gesture scaffold would not provide increased benefit (see McNeil et al., 2000).

In summary, I predicted that evidence of facilitative scaffolding would be seen in increased solution accuracy in the reworded and/or the gesture conditions in comparison to the baseline traditional wording condition. Any difference in the relative influence of the gesture and rewording scaffolds would be evident in a comparison of the reworded-alone condition (Scaffold **B**) to the reworded with gesture condition (Scaffold **D**). If the gesture and rewording scaffolds were not facilitative, I predicted that solution accuracy would not change across the four conditions. On the other hand, facilitative scaffolds were hypothesized to result in an increase in the number of problems solved accurately.

CHAPTER III

REVIEW OF THE LITERATURE

Introduction

Dissecting the act of solving an arithmetic word problem requires identification of those component processes that are essential to accurate problem representation and solution. A review of the math word problem literature reveals a complex set of factors that influence a child's ability to listen to or read, and solve, an arithmetic word problem. The following is an overview of this literature, in conjunction with a description of the component processes inherent in an arithmetic word problem solving task. This review begins with a brief discussion of the development of early mathematical cognition in young children, and then proceeds from early studies that view schematic representation as the locus of problem solving skill to more recent studies that emphasize the influence of linguistic factors and working memory. Studies of mathematical cognition and working memory in children with language impairment are also reviewed, in addition to gesture studies that scaffold language comprehension.

Development of Numerical Cognition

Specific to arithmetic word problem solving are the processes of mathematical conceptualization and computation. The development of numerical reasoning and cognition from infancy on has been well documented (e.g. see Fazio, 1996; Fuson, 1988; Geary, 2000; Gelman & Gallistel, 1978; Ginsburg, 1983; Ginsburg, 1989). Starting in infancy and early childhood, children display evidence of quantitative concepts (Hauser & Spelke, 2004; Gelman & Gallistel, 1978), begin to conceptualize part-whole relationships and number-object correspondences (Levine, Jordan & Huttenlocher, 1992), and develop initial representations of a mental number line (Case, Okamoto, Griffin, Siegler & Kaeting, 1996; Okamoto & Case, 1996).

Even before being exposed to the academic process of solving word problems, children are able to use their developing world knowledge and knowledge of basic quantitative concepts to solve real world math problems (Ginsburg, 1989; Hughes, 1986; Nunes, Schliemann & Carraher, 1993; see Riley, Greeno, & Heller, 1983 for a review), as evidenced in their ability to trade, compare, request, and place monetary values on specific quantities of toy soldiers, baseball cards, pieces of candy, and other real world entities. For example, Hughes developed his 'Box task' to demonstrate that when preschool children are given real world manipulatives and a problem such as figuring out how many objects have been added to or removed from a box, they eagerly use their sense of numerosity to solve simple addition and subtraction problems. Gelman and Gallistel (1978) developed a successful research program around the hypothesis that 2- to 5-year-old children know more about number concepts than previous studies had demonstrated, even though these children might not yet have the conventional verbal count sequence totally memorized yet.

Numerical skill continues to develop as children progress through elementary school (e.g. Geary, 2000; Greer, 1990), and depending on the particular school curriculum, students are introduced to formal word problems relatively early in the primary grades. Thus, an early conception of numerical quantities begins developing before the verbal counting system is fully acquired, and lays the foundation for later comprehension and use of verbal number strings, quantifiers, and other math-related lexical and syntactic constructions.

The Study of Arithmetic Word Problem Solving

Schema Representation

Much of the early research on word problem solving focused on the child's development of a formal mathematical representation of the problem (for a review, see Kintsch, 1998, pp. 332-334). Early theoretical and computational models emphasized the representation of the word problem that a child constructed, along with the inferences and computations required for accurate solution. Despite the fact that arithmetic word problem tasks are defined by their inherent linguistic content, these early models did not incorporate a specific language processing component.

For example, in a review of the existing literature, Riley, Greeno and Heller (1983) developed the argument that children's ability to solve word problems "primarily involves an increase in the complexity of conceptual knowledge required to understand the situations described in those problems" (p. 153). Although they superficially acknowledged the potential contribution of factors such as the use (or not) of concrete manipulatives, the grammatical complexity of the problem, the length and sequence of the problem statement, and a child's reading ability, Riley and colleagues focused on what they referred to as the semantic/conceptual relationships between the quantities presented in the problem statement (i.e. the problem wording). These relationships between quantities in the word problem statement, along with the associated computational procedures, are often referred to as the problem schema. In the case of arithmetic word problems, a schema can be described as a top-down control structure, in the form of an assimilated set of representations of experience and information that can be recalled as an aggregate problem-specific entity, to trigger accurate solution of the problem (Marshall, 1995; Riley, et al, 1983; Schoenfeld, 1985, pp. 50-51).

Under Riley and colleagues' (1983) model, arithmetic word problem solving skill is driven by the child's understanding of how sets of objects can be manipulated. With repeated practice, children develop part-whole schemata that they can then apply to particular problem solving situations. These schemata contain stereotypical information about the type of problem the child is being asked to solve, along with the procedures required for successful solution of that particular problem type (Blessing & Ross, 1996). Once the child has recalled the appropriate schema for a given problem type, he fills in problem-specific information (such as the quantities for each set) and executes the solution. Riley and colleagues' model of schema activation, however, did not contain an explicit language processing component despite the inherent linguistic content of word problems.

Building on their earlier work, as well as the work of Carpenter and Moser (1982), Riley, Greeno and Heller (1983) described how children activate these schemata and apply this conceptual knowledge of increasing, decreasing, combining, comparing and equalizing sets of objects to specific types of word problems (See Appendix A for a description and examples of word problem types). Riley and colleagues described three kinds of knowledge necessary for solving all types of basic arithmetic word problems: 1) a problem schema: the child's representation of the problem's semantic relations, 2) an action schema: the child's knowledge of the mathematical actions (procedures) needed to solve the problem (e.g. make a set, count all, take out part), and 3) the top-down strategic knowledge required to plan and execute the solution. Deficits in any one or more of these three knowledge bases were predicted to explain children's difficulty in solving word problems. In addition, differences in problem schemata across the various types of word problems were thought to reflect the relative difficulty levels of these different problem types.

Structure and Relative Difficulty of Arithmetic Word Problem Types

For many years preceding and following Riley and colleagues' work, researchers described the structure of basic word problems, and also explored the relative difficulty levels of the different word problem schemata across the range of conceptual and computational skill development (see Okamoto, 1996, p. 410; Riley, Greeno, & Heller, 1983, pp. 159-163; Riley & Greeno, 1988 for summaries).

All types of addition and subtraction word problems provide three pieces of information, and in any type of word problem one of these entities will be an unknown

(See Appendix A). For example, in *Change* schemata, the three bits of information are the starting amount, the amount of change, and the resulting quantity. The level of difficulty will vary depending on which unknown entity the child must solve. A problem with specified starting and change quantities is easiest to solve, but if either the starting amount or the changed amount are unspecified, the solution is more difficult for children, presumably because of the need to retain an unspecified entity until more information is supplied. Also, depending on the particular *Change* schema, the direction of the change could result in an increase (join) or a decrease (separate) in the starting quantity. Word problems are more difficult when this computational direction of change is not consistent with the wording used in that schema type. For example, a *Compare 6* schema was more difficult for college students than a *Compare 3* schema because although both require an addition computation, the relational term "less" is used in the Compare 6 schema (Lewis & Mayer, 1987). And finally, problems that involve action (e.g. *Change* schemata) are generally more readily and accurately solved by younger children than static schemata (e.g. *Compare*) (Fuson et al., 1996).

Riley, Greeno, and Heller (1983) and Riley and Greeno (1988) demonstrated that word problems requiring the same computational operation (e.g. addition), but having different semantic structures (i.e. schemata, e.g. *Compare 3* and *Change 1*, See Appendix A) are not equally difficult for first graders. Riley and colleagues attributed these differences in problem difficulty to differences in the part-whole relations used to describe the problem schemata. For example, *Combine* and *Compare* problem types involve static quantities: in *Combine* problems the child simply combines two distinct quantities to obtain a sum, and in *Compare* problems the child is required to evaluate the difference between two distinct quantities. In *Change* problems, however, the initial quantity increases or decreases, resulting in a new final quantity. Within the realm of addition computational processes, *Compare* problems are more difficult than both

Change and *Combine* schemata. For subtraction processes, *Combine* and *Compare* problems are more difficult than *Change* schemata.

In summary, the research studies I have described to this point attribute the relative difficulty of word problem types to differences in their respective schemata. Although several of these studies described the potential contribution of specific semantic terms to problem difficulty (e.g. Riley & Greeno, 1988), none explicitly tested the specific contribution of linguistic factors to a child's ability to comprehend and solve an arithmetic word problem. This body of research contends that children comprehend and solve arithmetic word problems by a top-down process of calling on specific formal schemata, but gives little indication of factors that might influence the development of these schemata.

Linguistic Factors

Some researchers who advocated for the importance of mathematical schemata in word problem comprehension also allowed that linguistic factors within a given schema-type could influence problem difficulty, and therefore problem comprehension and solution accuracy (e.g. Cummins et al., 1988; Cummins, 1991; Fuson et al., 1996). Kintsch and Greeno (1985) developed a computational model that combined Riley, Greeno and Heller's (1983) work on the semantic knowledge required for processing word problems with van Dijk and Kintsch's model of text processing (van Dijk & Kintsch, 1983). According to Kintsch's Construction-Integration Theory and Model, text is comprehended by constructing a mental model that merges specific ideas expressed in the text (the textbase) with relevant prior knowledge (the situation model: Kintsch, 1998, pp. 92-120). The rationale for studying word problems within this framework was expressed by Kintsch 1988, p. 174):

> …word problems, like all other texts, share the ambiguity and fuzziness of all natural language. Not only formal, arithmetic knowledge is involved in understanding these problems, but all kinds of linguistic and situational knowledge. What makes word problems hard—and interesting—are often

> not their formal properties, but the way a problem is expressed linguistically and the way formal arithmetic relations map into the situations being described. Thus, word problems are ideal from the standpoint of knowledge integration because it is precisely the integration of formal arithmetic knowledge and linguistic and situational understanding that is at issue here.

Kintsch and Greeno (1985) adapted Van Dijk and Kintsch's (1983) discourse comprehension model to accommodate word problem processing, by incorporating a math-specific set of strategies into the original discourse comprehension strategy set. In addition to its text processing capabilities, the revised computational model was able to form numerical sets and assign them ownership, as well as represent the cardinality of a set and the relationship and actions that might occur between sets in a word problem. Hence the model was able to classify errors as being linguistically or arithmetically-based. Kintsch and Greeno (1985) proposed that children also must develop this set of math-specific strategies that differ from other types of text comprehension strategies. Analogous to Kintsch's Construction-Integration text comprehension theory, children build a textbase representation for the input text, as well as a problem model that extracts the abstract problem-relevant information from the textbase in order to mathematically compute a solution (Kintsch & Greeno, 1985, p. 111). During the Construction phase, numerous strategies, such as 'combine two quantities' or 'compare two quantities and compute the difference', are activated while processing the text, and then during the Integration phase inappropriate representations and strategies are inhibited. Successful problem solution occurs when the correct schema is most strongly activated, wins the competition, and then triggers the appropriate computational processes (Kintsch, 1988).

To investigate the relative influence of mathematical and linguistic factors on word problem comprehension, Cummins, Kintsch, Reusser, and Weimer (1988) tested whether children's performance on word problems would vary systematically with their ability to recall the problem (i.e. a measure of comprehension) rather than with their ability to solve a numeric equivalent (e.g. 5 + 3 = ?). Using a simple regression model, only the structural recall variable met the significance criteria for entry into the model, and accounted for 72% of the variance in word problem solution accuracy. Cummins and colleagues (1988) also found support for their prediction that when difficult problems are harder to solve because of a more-complex text base, children were more likely to simplify the recall wording of these difficult problems than problems with easier schemata. Interestingly, providing a linguistically-rich, but not necessarily problem-relevant context, as opposed to the 'sparse' language of traditional word problems did not facilitate problem comprehension when the word problems were presented in a read-along format.

In their error analyses, Cummins and colleagues (1988) found that children's interpretation of individual lexical items influenced solution accuracy. For example, some children tended to interpret the word 'some' as an adjective rather than an unknown quantity to be identified and computed. As a result, these children erred by providing one of the quantities stated in the problem (a given number error), instead of computing the correct solution. Likewise, the phrase 'has x more than' was sometimes misinterpreted to simply mean 'has x', resulting in miscomprehension and solution of the word problem. Also, some children interpreted the word 'altogether' to mean that *each* child possessed that quantity of an object, rather than in sum. These findings were replicated in large part in simulation modeling, and indicated the importance of children's interpretation of individual lexical items as a factor in successful comprehension of word problems.

In a follow up study, Cummins (1991) asked first graders to solve *Combine 5* (to test 'altogether'), *Compare 4* and *6* (to test 'more/less than'), and *Change 6* (to test 'some'). One day later she asked the children to listen to similar problems and then draw their representation of each problem. Finally the children listened to the original problems from Day 1 and were asked to select the picture that best depicted the problem. She found that misrepresentations of the problems were related to specific lexical forms. For example, the most frequent error for the *Combine 5* problems was a wrong operation error, where children added instead of applying the required subtraction operation.

Cummins concluded that this addition operation was the default strategy that children used most frequently when having difficulty interpreting a word problem.

In a second experiment, she selected those children who demonstrated representation and solution errors, and asked them to listen to, solve, and draw depictions of *Combine 5* problems presented in traditional and reworded (to avoid the use of 'altogether') formats. Cummins found that for the traditionally worded *Combine* problems, inaccurate solutions were associated with inaccurate representations of the word problems. The majority of the errors were wrong operation errors (e.g. adding instead of subtracting) or given number errors (i.e. providing one of the numerals from the problem as the answer). Both of these error types were thought to reflect misinterpretation of the word problem.

Rewording the *Combine 5* problems, to facilitate conceptual representation and choice of solution strategy, resulted in a lower proportion of wrong operation and given number errors. Cummins concluded that a significant reduction in errors for the reworded problems indicated that children were able to use conceptual part-whole set knowledge appropriately if the linguistic complexity of the problem was within their comprehension ability. These findings suggest that rewording a more difficult arithmetic word problem can serve as a scaffold to facilitate problem comprehension in typically-developing first graders.

Contextualized Knowledge

A limitation of the Kintsch and Greeno model is the inability of the problem model component to also incorporate world knowledge, i.e. a situation model (Kintsch, 1988, pp. 174-175). Therefore, although the model can explain both the linguistic/semantic and math strategy errors commonly seen in error analyses of children's word problem solutions and retellings, it cannot account for the striking results reported by Hudson (1983) that the rewording of a linguistically difficult word problem to place it in a more realistic situation model resulted in increased problem comprehension for primary grade students.

Similar evidence of world knowledge is found for non-math domains. For example, Kintsch (1994b) presented evidence regarding the influence of background knowledge on a reader's ability to process expository text. In his model of comprehension processes, the situation model takes information from the text (i.e. the textbase) and incorporates this textbase into the reader's existing knowledge base. When an expository text was rewritten using more explicit language, low-knowledge adult readers performed much better on problem solving questions than when they read a less explicit version. Interestingly, high-knowledge readers performed better with the less explicit text. Kintsch concluded that readers must be able to access a deeper situation model in order to comprehend a text well enough to problem solve. It might also be possible to scaffold the formation of a situation model in an arithmetic word problem solving task.

Influenced by Kintsch's concept of a situation model as well as prior research on the effect of linguistic factors on word problem difficulty (e.g. De Corte & Verschaffel, 1987; Hudson, 1983), Riley and Greeno (1988) created a revised set of cognitive models incorporating semantic networks (analogous to van Dijk and Kintsch's textbase, 1983) and semantic models (analogous to the situation model). These models were designed to simulate a number of word problem types examined in previous studies, in order to test the relative difficulty of these problem types based on their linguistic content, problem structure, and situation model. Riley and Greeno then compared their simulations to behavioral data from typically-developing children in kindergarten through third grade. This research project yielded valuable data describing the proportions of correct solutions provided at each grade level for the 18 problem types. By comparing their behavioral data with their simulation data, Riley and Greeno also provided more a more-detailed description of the reasons for differences in relative difficulty of the *Combine, Change,* and *Compare* problem types. With regard to the influence of linguistic factors on children's ability to solve arithmetic word problems accurately, they concluded that increases in both conceptual knowledge (i.e. the relations between sets) and linguistic knowledge (e.g. specific lexical items) were necessary in order for children to experience success with all types of word problems. They also suggested, however, that the ability to understand a linguistic phrase such as 'x more than' might depend on an additional type of knowledge: the understanding that a comparison of two sets yields a new numerical value. In that case, scaffolding linguistic factors in a word problem might not facilitate an increase in solution accuracy, if a child does not have access to this additional numerical knowledge.

More recently, Thevenot, and colleagues (2007) found that French-speaking children with lower mathematical ability benefited most from moving the question to the beginning of a word problem, especially for the more difficult problems, and concluded that this placement facilitated a more accurate mental model of the problem. Interestingly, the question-first placement also virtually eliminated the differences in difficulty between *Combine 1*, *Compare 1*, and *Compare 2* problems. Their results indicate that, in addition to rewording difficult word problems, scaffolding a more accurate situation model (i.e. presenting the cognitive rationale and set relations up front) can also increase solution accuracy.

Working Memory

Working memory has also been identified as a factor contributing to children's comprehension of word problems. Functioning from the perspective that solving a word problem requires a complex interaction of text comprehension, mathematical problem solving, background knowledge, and cognitive processing demands, LeBlanc and Weber-Russell (1996) explored how working memory constrains the comprehension of word problems. Using a computer simulation and the results from previous studies of kindergarten through third grade children, they found that the combination of total number of concepts that needed to be remembered plus the number of inferences required to process the word problem (based on text integration demands) consistently predicted the simulation model's success in solving the problem. Their computer model was unique, in that it processed the word problem sentences word by word, via independent lexical entries, in an attempt to simulate human language processing. This bottom up model was in contrast to prior arithmetic word problem models, where solution strategies, relationships between numerical sets, and/or propositions representing entire sentences had been coded up front and exerted top down control (e.g. Kintsch & Greeno, 1985; Riley & Greeno, 1988; Riley, Greeno & Heller, 1983).

LeBlanc and Weber-Russell (1996) tested their computational model on *Combine 5* problems, in both a traditional wording format and two reworded formats. For each wording format, they calculated the number of concepts that the model had to remember, as well as the number of inferences that the model had to make. LeBlanc and Weber-Russell found that not only did rewording increase the solution accuracy in their model, but it also reduced the number of inferences, and therefore presumably decreased the demand on working memory.

Using Baddeley's multi-component model (Baddeley, 2003), Swanson and Beebe-Frankenberger (2004) explored the cognitive correlates that mediate skill in word problem solving. They tested two models of memory in an arithmetic word problem solving task: a short term memory/phonological processing loop model and a competing model with an executive function component that assumed a unique contribution of working memory. Two groups of first, second and third graders (typically-developing children and children who scored in the lowest quartile on standardized oral word problem and digit naming fluency tasks) were asked to solve a series of *Combine* and *Compare* word problems and to complete an assessment battery of memory, reading, and math measures. The results of a series of regression models revealed a significant relationship between working memory and children's ability to solve arithmetic word problems, even after measures of phonological processing, inhibition, processing speed, math calculation, and reading skill were partialed out. When entered by itself into the regression model, working memory contributed approximately 30% of the variance in problem solving ability. Swanson and Beebe-Frankenberger concluded that the top down process of working memory ability contributed more to children's ability to solve word problems accurately than bottom up processes (i.e. phonological processing, in this study). Their conclusion and generalization about the relative contribution of top down and bottom up processes, however, did not consider the potential contribution of language processing to skill in solving word problems, beyond the contribution of phonological processing as a measure of short term memory. Although these results suggest that working memory contributes significantly to children's ability to solve arithmetic word problems, the potential mediating influence of language ability was not considered in this study.

Summary

A summary of the literature reviewed to this point indicates that, although children have the cognitive and numeracy skills to successfully solve age-appropriate real-world arithmetic problems before entering school (e.g. Fuson, 1988; Geary, 2000; Gelman & Gallistel, 1978; Ginsburg, 1983; Ginsburg, 1989; Hughes, 1986; Okamoto & Case, 1996), continued development of formal mathematical skill requires that children acquire formal schemata for solving word problems. From the initial theoretical perspective that children develop word problem schemata over time due to increased understanding of the relationships between numerical sets, i.e. the semantic structure of the problem (along with increased skill in counting, calculating and executing strategies; e.g. Riley, Greeno, and Heller, 1983), later research has shown that linguistic processing factors also contribute to arithmetic word problem skill and schema development (also see Elman, 2007; Rumelhart, Smolensky, McClelland & Hinton, 1986). This body of research has demonstrated strong links between children's ability to solve word problems and individual lexical items in the problem (e.g. Cummins, 1991), the location of the question (Thevenot, et al, 2007), the situation model (e.g. Hudson, 1983), and working memory demands (e.g. LeBlanc and Weber-Russell, 1996).

Given these links between word problem solving ability and linguistic factors, it is important to test the relationship between children's oral language skill and their ability to solve basic arithmetic word problems, and then to examine the effect of rewording scaffolds on solution accuracy in children with low oral language skill.

Individual Differences in Numerical Cognition: Language Impairment

The rich literature of addition and subtraction word problem solving in typically-developing children presents the opportunity to extend this work into the realm of individual differences in arithmetic word problem solving. Due to the importance of language processing in successful arithmetic word problem solving, we could predict that children with low oral language skill would exhibit difficulty in building mental representations of linguistically abstract, formally-worded problems, and would continue to struggle academically with arithmetic word problems. In order to plan interventions that will help these children deal with the linguistic component of word problems, we first need to consider those studies that have investigated numeracy and formal math skill in children with language impairment.

Fazio (1994) studied the ability of 4- to 5-year-old children with Specific Language Impairment (SLI) from families with low socioeconomic status to count objects accurately. SLI is defined as a discrepancy between mental age and expressive and/or receptive language ability, in the absence of any other causal factors such as hearing impairment, overt neurological disorder, mental retardation, or autism. These children also exhibit more subtle deficits in the areas of auditory processing, nonverbal ability, phonological working memory, vocabulary skill, and argument structure (Leonard, 1998), and show subtle signs of other cognitive difficulties (Miller, Kail, Leonard & Tomblin, 2001).

Compared to their cognitive age matches (CM - matched on both age and nonverbal ability; also from families with low socioeconomic status), the children with SLI in Fazio's study (1994) were not able to rote count as high or as accurately. When asked to count sets of objects ranging in size from 3 to 9, children with SLI again were significantly less accurate than their CM peers. The errors made by the children with language impairment tended to be count sequencing errors (errors in the verbal sequence of counting words) rather than tagging errors (miscounting the 1:1 correspondence between counting words and objects), suggesting that language impairment affects verbal count sequencing ability more than it affects numerical conceptual ability. No significant difference between the SLI and CM groups was found on a test of cardinality (understanding that the last number named represented the size of the set). Children with SLI did not differ from their language-matched group (LM) on measures of counting range or accuracy, but they did perform better than the LM group on the cardinality tasks, again suggesting that language ability influenced language-mediated tasks, but not conceptual tasks.

Fazio then taught the children a gesture to body part nonverbal count sequence, and repeated the assessments. In a comparison of the gestural and oral tasks, children with SLI performed significantly better on the nonverbal tasks, suggesting a relative strength in nonverbal as compared to verbal count sequencing skill. Fazio concluded that, despite their conceptual knowledge of numerical sets and the basic procedural rules of counting, children with SLI have difficulty learning and using the sequence of number words. She suggested that this difficulty with verbal count sequencing would later be a bottleneck to the acquisition of formal arithmetic skill.

The same group of children was followed up two years later when they were first and second graders (Fazio, 1996). Children with SLI still lagged behind the CM group in ability to rote count beyond 20 and to sequence verbal count numbers. Recall of math facts was also difficult for them. When allowed to count on fingers rather than retrieve facts, their performance improved. Magnitude comparison ('Do you have more or do I have more?') and conceptual and procedural knowledge of the basic principles of addition (understanding how to add sets, and using fingers or manipulatives to carry out addition tasks) were relative strengths for the children with SLI, indicating again that these children exhibited a relative strength in nonverbal conception of number as compared to rote verbal count and number fact skill. Fazio noted that the language-impaired children's accuracy increased when they were allowed to use objects to count, and when the cardinality of the counted set did not exceed the child's accurate rote counting ability.

Three years later, when the children were fourth and fifth graders, Fazio (1999) found that although both the cognitive age match group (CM) and a younger (third grade) group (YM) were now relying more on number fact retrieval to solve 2-digit addition, subtraction, multiplication and division problems, the group with SLI still relied on counting strategies rather than fact retrieval. In addition to making more calculation errors than both the CM and YM groups, and making more procedural errors (e.g. borrowing and carrying errors) than the CM group, children with SLI also required more time to complete the calculations, and performed significantly worse than both groups on timed math fact measures. There were no group differences in wrong operation errors (e.g. adding instead of subtracting). These results suggested that the children with SLI remained vulnerable to making errors when they were required to process and retrieve memorized number facts, and that this deficit influenced their ability to solve 2-digit arithmetic problems.

Donlan, Bishop and Hitch (1998) found evidence of unimpaired symbolic processing in 6- to 7-year-old children with SLI on a nonverbal magnitude comparison task. The SLI group and a typically-developing comprehension-matched control group were asked to look at pairs of arrays of dots, Arabic numerals, line drawings of houses and line drawings of animals, and to select the one that was larger, as quickly as possible. In order to test whether verbal mediation was a factor (especially for the Arabic numerals), children were asked to say the word ‘ice-cream’ repeatedly while completing the task. There was no group difference in error rates, but there was a significant group effect of response latency, with the SLI group responding faster than the control group. The concurrent verbal interference task did not disrupt either group. Donlan and colleagues concluded that symbolic processing is not globally impaired in children with language impairment, and that this magnitude comparison task does not require verbal mediation.

In a follow up study, Donlan and Gourlay (1999) tested language-impaired children and two control groups, an age and nonverbal match group (AC) and a comprehension match group (LC), on single- and double-digit Arabic numeral magnitude comparison tasks. They found no difference in response time between the SLI and AC groups, and both groups responded faster than the younger language comprehension match group. When school experience for double-digits was controlled, there was no difference between the SLI and AC groups in either accuracy or response time. Donlan and Gourlay concluded that verbal skill and spoken number knowledge might not be a necessary condition for the development of numeracy skill. Both of these studies support the existence of a separate nonverbal representation system that underlies the development of numerical skill, and that might be able to be tapped to scaffold the development of formal numerical ability.

Additional support for the possible dissociation of nonverbal and verbal skill in the acquisition of numerical knowledge was provided by Arvedson (2002). A group of preschoolers diagnosed with SLI and their age-matched (CA) and grammar-matched (LA) controls completed a series of numerical problem solving tasks that did not require verbal counting. Although the SLI group performed lower than the CA group on 28 of the 60 trials, overall the SLI group performed more like the CA group than like the LA group (based on effect sizes), indicating that nonverbal numerical processing can occur independently of language level. However, when required to count verbally, the accuracy of the SLI group dropped by 50%. Arvedson commented anecdotally that the language-impaired children as a group did not choose to use verbal or finger counting, and when required to count "they reacted as if they had been distracted from their numerical reasoning task and were asked to add another, perhaps separate, task to the one they were working on." (p. 980). Donlan & Newton (2007) suggested that this verbal counting deficit might be attributable not only to a verbal procedural sequencing deficit, but to the syntactic complexity of the English count system beyond the numeral 10. In contrast to the English system, number sequences in many of the Asian language are much more transparent, and Asian children with language impairment do not experience as much difficulty with count sequences beyond 10 (Donlan & Newton, 2007).

In two additional studies Donlan and colleagues continued to explore the relationship between number knowledge and language in children with SLI. Cowan, Donlan, Newton, and Lloyd (2005) studied the influence of nonverbal reasoning, language comprehension, working memory, and educational exposure on a set of math skills: counting, fact retrieval, calculation skill, orally-presented word problems (*Change 1, 2, 5, 6*), transcoding (converting between Arabic and verbal number systems, e.g. '15' versus 'fifteen'), and multi-digit magnitude judgment. Seven to nine-year old children with SLI performed significantly lower on all tasks than their age matches. The strongest predictors of arithmetic word problem performance were language comprehension, working memory, and nonverbal skill.

Donlan, Cowan, Newton and Lloyd (2007) looked at the relationship between language impairment and two types of math skill: cognitive (i.e. underlying conceptual knowledge) and procedural (e.g. calculation ability). They suggested that children with SLI could develop conceptual understanding, based on a relative strength in nonverbal functioning, prior to their development of procedural skills that are mediated by verbal ability (also see Donlan, 1998a). Compared to their age-matched peers, 8-year-old children with SLI exhibited significant deficits in the procedural tasks of counting and basic calculation (fact retrieval with manipulatives; single and multi-digit addition and subtraction), and in the conceptual task of multi-digit magnitude comparison. However on a difficult nonverbal conceptual task where participants were asked to judge basic arithmetic principles (e.g. commutation: $\sigma + \gamma = \gamma + \sigma$) in equations containing abstract Greek symbols, children with SLI performed as well as their age matches. Donlan and colleagues concluded that language may differentially mediate and affect the development of cognitive and procedural knowledge, and that children with SLI may have a relative strength in their conceptual knowledge of basic arithmetic principles. These results suggest that children with SLI may be able to draw on their relatively stronger nonverbal conceptual skills to scaffold their development of formal numeric ability.

Jordan, Levine, and Huttenlocher (1995) compared four groups of kindergarten and first grade children on three arithmetic problem formats: nonverbal problems, verbal arithmetic word problems and number fact problems (single digit, two- and three-term problems for all three problem formats). The four groups included children with low language but good nonverbal skill; children with nonverbal impairment only; children with both deficits; and typically-developing children (TD). The nonverbal and verbal problem tasks are described in detail in Chapter IV, Procedure and Tasks. Jordan and colleagues found that performance on the three problem formats was sensitive to both language and nonverbal ability in different ways. The low language group performed significantly lower than the TD group on the verbal word problems and verbal number facts, but performed comparably to the TD group on the nonverbal problem task. The group with nonverbal impairment did not differ significantly from the TD group on any of the three problem formats, possibly because they used verbal mediation to solve the nonverbal problems. The group with both language and nonverbal deficits performed lower than the TD group on both the nonverbal and verbal problem formats, but, surprisingly, performed as well on the number facts. However when finger use (as a calculation strategy) was entered as a covariate, the adjusted mean for the number fact problems for the TD group was higher than the adjusted mean for the group with both language and nonverbal deficits.

When performance on the three problem formats was compared within each group, the low language group showed the largest discrepancy between the nonverbal problem format and the two verbally-mediated formats (verbal word problems and verbal number facts). Jordan and colleagues' (1995) results indicate a relative strength in the development of nonverbal arithmetic skill in the low language group and suggest that a nonverbal (e.g. gesture) scaffold might facilitate word problem solution for these children.

Newton is currently studying how language impairment constrains reasoning and puzzle-solving skills by using matched spoken and picture tasks (Donlan & Newton, 2007). On a reduced array selection task with low working memory load, but high inferencing load, children with language impairment performed worse than age mates, but better than language matches, especially when a rationale for each of the problems preceded a pictured representation of the reasoning task (e.g. If you eat too many donuts, you will get fat.) Although this task did not involve mathematical skill, the results indicate that providing children with a real-world rationale for engaging in a task is facilitative.

Summary of Mathematical Skill in Children with Language Impairment

Across this set of studies we see that English-speaking, primary school-age children with language impairment have deficits in the syntactic process of rote counting, with prolonged difficulty in count sequences that extend above 10. Number fact retrieval (and mental calculation) is also difficult for these children, although use of manipulatives is facilitative in some cases. Calculation accuracy also improves for these children if their accurate rote counting ability is not exceeded (Fazio, 1996). Several studies indicated that magnitude comparison and conceptual knowledge of basic mathematical principles are relative strengths in children with language impairment (Donlan et al, 1998; Donlan & Gourlay, 1999; Donlan et al, 2007; Fazio, 1996). Jordan and colleagues (1995) provided additional evidence that successful nonverbal numerical processing can occur independently of language level.

Although arithmetic word problems are quite difficult for this population (Cowan et al, 2005), the evidence that children with language impairment have relative strengths in several areas of nonverbal numerical cognition suggests that reducing the linguistic load in word problems could potentially scaffold greater solution accuracy. Earlier in the literature review I presented evidence of how rewording a word problem, placing the question first, and enhancing the situation model facilitates comprehension of arithmetic word problems in typically-developing children. One of the aims of this study is to test whether scaffolds designed to address a linguistic bottleneck in children with low oral language skill will allow them to access their nonverbal numerical knowledge.

Another way to facilitate the comprehension of arithmetic word problems in children with low oral language skill might be to help these children access the underlying problem structure (the relationships between sets of numbers) by providing a gesture scaffold. In the following section I will review the literature on gesture scaffolding in typically-developing children and children with language impairment, and consider whether gesture scaffolds might be facilitative for children with low oral language skill.

Gesture Scaffolds

Gesture Use

Evidence for a gesture scaffold originated in research on how speakers use gesture to convey the conceptual content of their messages. Alibali, Kita, and Young (2000) asked 5-year-old children to either describe how two objects looked different (a descriptive task), or to explain why the two objects did or did not have the same quantity (a conservation judgment). Despite the fact that the children used similar types of lexical items in the two conditions, with similar corpus frequencies, they produced more substantive, non-redundant gestures in the higher cognitive load conservation task. Alibali and colleagues concluded that gestures reveal the speaker's cognitive conceptualization and are involved in the conceptual planning of messages.

Church and Goldin-Meadow (1986) asked 28 5- to -8-year-old children to make conservation judgments on a set of Piagetian tasks, and then explain their responses. All but one of the children produced gestures while providing their explanations. Those children who produced more gesture-speech mismatches (where the information conveyed in gesture was different from the information conveyed in the accompanying verbal message) expressed a more substantive level of reasoning, particularly in their gestures. Church and Goldin-Meadow argued that these children possessed a higher level of knowledge that they had not yet integrated into their verbal responses, and thus were in a transitional state. The researchers predicted that children who provided gesture-speech mismatches would benefit more from instruction in conservation equivalence than children who provided gesture-speech matches. After instruction, the gesture-mismatch group was significantly more likely to add a new equivalence explanation to their verbal conservation explanations than the gesture-match group. In addition, for 61% of those children who added new verbal explanations, evidence of their new explanation was found in the gestural component of their pretest responses. These results indicate that 1) the gesture channel provides a window into a child's cognitive representations, 2) children's gestures often provide more information than their verbal explanations, especially for children who are in a transitional state of knowledge acquisition, and 3) instruction facilitates the integration of gesture and verbal explanations.

Beyond establishing that children know more than what they say, Garber, Alibali, and Goldin-Meadow (1998) demonstrated that this gestural knowledge is often represented uniquely in a nonverbal format. When fourth grade children produced gesture-speech mismatches while explaining their solutions to mathematical equivalence problems (e.g. 5 + 2 + 3 = __ + 3), the knowledge expressed in gesture was often not expressed in their verbal explanations. In addition, when asked to rate the correctness of a set of solutions to similar problems, these children rated the solutions that they produced uniquely in gesture significantly higher than solutions they did not produce during the explanation task. These results suggest that the children had implicit access to the knowledge that they had produced in their own gestures. Under the theory that this nonverbal gesture information represents cognitive information that the child has at least partial access to, but has not yet integrated into his speech, we can then ask whether children would be able to access and benefit from information produced in gestures used by their teachers.

Access to Teachers' Gestures

Goldin-Meadow, Kim, & Singer (1999) explored gesture and cognition in this reverse direction, by looking at how teacher's gestures influenced third and fourth grade students' ability to solve and explain mathematical equivalence problems. Both verbal and gestural problem solving strategies used by the teachers were coded and compared to the children's responses. The children were significantly more likely to reiterate the teacher's solution strategies verbally when the teacher had presented the strategy in speech accompanied by a gesture, as opposed to speech without a gesture. It was also of interest that in 9 of 12 instances where a teacher presented an incorrect strategy in gesture, accompanied by the correct strategy in her speech, children reiterated the incorrect strategy in their responses. In addition, when the children reiterated a teacher's strategy that had been presented only as a gesture, 91% of these reiterations were in the verbal modality. These results provided evidence that children can take advantage of information conveyed via gesture, and that children were able to transform information into speech that was presented uniquely in gesture. In this typically-developing population, however, we could assume that the children's oral language skills were within normal limits. It remains to be seen whether children with low oral language skill will also be able to make use of gesture information that accompanies arithmetic word problems whose language they don't understand. It also remains to be seen whether children with low oral language skill will be able to integrate gesture with the speech it accompanies.

Also using mathematical equivalence problems, Church, Ayman-Nolley, and Vasich (2007) found that typically-developing children who were exposed to a representational gesture (for the concept of 'equals') exhibited greater learning than children who were exposed to a simple 'beat' gesture, and concluded that the representational gesture helped bridge the gap between the teacher's verbal explanation and the underlying concept of equivalence. Cook and Goldin-Meadow (2006) looked at the relationships between gesture use by the teacher, subsequent gesture use by third and fourth graders, and those children's success on equivalence problems. They found that a representational gesture by the teacher helped typically-developing third and fourth graders understand the teacher's verbal message, and resulted in increased expression of the equalizer strategy in gesture by the children when they explained their own solutions. During the instruction session, there was a significant difference in solution accuracy between those children who did not express a correct strategy in either speech or gesture and those who either expressed the correct strategy in speech or in speech plus gesture. Children who verbalized the correct strategy (but did not gesture it), however, solved almost as many problems correctly as children who expressed the correct strategy in both speech and gesture. It is of interest to note that, on a posttest that required both retention of similar material and extension of equivalence knowledge to a new problem type, those children who had expressed their solutions via both speech and gesture solved significantly more problems correctly than those children who expressed equivalence only in speech, or those who did not express equivalence in either speech or gesture. These results provided evidence that typically-developing children were able to make use of the information conveyed in the teacher's gesture to increase their problem solving ability, especially over time, as opposed to immediately during instruction. As Cook and Goldin-Meadow cautioned, however, not all of the children exposed to teacher gestures actually used an equivalence gesture themselves, raising the possibility that not all children were cognitively ready to process the concept of mathematical equivalence.

McNeil, Alibali, and Evans (2000) explored the effects of reinforcing versus conflicting gestures in an oral language comprehension task with two groups of children: preschool children, who were predicted to rely more on contextually-based comprehension strategies, and kindergarten children, who would be expected to rely more on language-based strategies. They found that reinforcing gestures were most facilitative when the spoken message was more complex. Under conflicting gesture conditions, however, kindergartener comprehension was more adversely affected than preschool comprehension and preschoolers were not affected by conflicting gestures under either easy or complex message conditions. McNeil and colleagues concluded that reinforcing gestures accompanying complex messages provide a redundant signal in another modality, but conflicting gestures might not have an effect on younger children who might only be able to focus on one modality, even under easy message conditions.

It remains to be seen whether this potential developmental progression of attention allocation will have an effect on children with low oral language skill as they solve gesture- and language-scaffolded word problems.

Across all of these studies, representation of information via multiple modalities seemed to improve typically-developing children's depth of conceptual knowledge, and reinforcing gestures functioned as a mechanism of cognitive support and change. These children did attend to gesture, and those gestures influenced their performance. The current study will examine whether children with low oral language skill also benefit from information presented in the gesture modality and whether reducing the verbal load of a word problem will make the gesture more accessible.

Gesture and Language Impairment

Gesture Use

To date, we have limited information on gesture use in children with language impairment (also see Capone & McGregor, 2004, for a review of gesture development in typical children and children with language impairment). Evans, Alibali and McNeil (2001) found that 7- to 9-year-old children with SLI and phonological working memory deficits gestured at a similar rate to a group of younger, typically-developing children matched on conservation judgment ability. In addition, when providing explanations for their conservation task judgments, the children with SLI expressed more sophisticated information via gesture than they were able to express verbally. Finally, the children with SLI produced more information uniquely in gesture than the judgment match group. These results indicated that, like typically-developing children, these children with language and phonological deficits possessed an implicit knowledge of some aspects of conservation, even though they were not always able to express this information verbally.

Mainela-Arnold, Evans and Alibali (2006) also studied gesture use during Piagetian conservation task explanations in three groups of children: a group of 7- to 10-

year-old children with SLI, a chronological age match group (CA), and a conservation judgment match group (CM). Mainela-Arnold and colleagues coded both speech and gesture for ‘external’ explanations (i.e. perceptually-based) versus ‘internal’ explanations (i.e. cognitively-based explanations expressing the initial equality and/or the transformation of the quantities). They found that, compared to the CA group, the children with SLI expressed fewer, higher-level internal explanations in the verbal modality, but they produced a similar proportion of internal explanations as the CM group. In the gesture modality, however, all three groups produced similar proportions of internal explanations, again indicating that children with SLI are able to access and express higher-level explanations in gesture than in speech. In discussing their results, Mainela-Arnold and colleagues cautioned that, because the children in the SLI group did not express more internal explanations in the gesture modality than the younger CM group, it was more likely that it was simply easier for children to express a perceptually-based external gesture than it was to express an internal explanation in gesture. In this case, the more-easily produced perceptual gesture might be encouraging the children with SLI to focus more on perceptual than on higher-level cognitively-based explanations.

Access to Teachers’ Gestures

Only a couple of studies have investigated the effects of gesture on language-impaired children’s comprehension. Ellis Weismer and Hesketh (1993) studied the effects on novel word acquisition of varying three input conditions: rate of speech, prosodic stress, and use of gesture cues. With regard to the gesture input condition, a group of kindergarteners with SLI and a group of typically-developing age mates (CA) were asked to listen to and learn some nonsense words for the locations of a spaceman. Iconic gestures were paired with each of the location words. Following training, both groups comprehended the novel words significantly better in the gesture condition than in the no gesture condition. There was not an effect of gesture for the production condition, however. It is of interest that 3 of the 4 children who showed the largest gains in comprehension with the gesture cues were also those children who were identified with comprehension deficits.

In an eyetracking study that looked at individual differences in the use of word learning cues, McGregor and Rost (2007) found that an eye gaze gesture cue was facilitative for both typically-developing children and children with language impairment in a spoken word recognition task. A facilitative eye gaze, as compared to a contradictory or neutral eye gaze, resulted in faster, more accurate word comprehension and novel word mapping in both groups.

In these two studies word learning in children with language impairment was facilitated by using locative and eye gaze gestures. It remains to be seen whether children with low oral language skill will benefit from a gesture designed to scaffold an arithmetic word problem task.

Summary and Questions

Although arithmetic word problems have been studied in typically-developing children and children with math disability (e.g. Jordan, Hanich & Kaplan, 2003; Thevenot, et al, 2007), few researchers have specifically explored the influence of language deficits on ability to represent and solve basic addition and subtraction word problems. Children in American school systems are exposed to arithmetic word problems as early as first grade so the question remains how we can facilitate the acquisition of word problem skill in a population of children who demonstrate a relative weakness in the linguistic, memory, and count sequence components of mathematical knowledge, and a potential relative strength in nonverbal numeracy and basic numerical concepts. Given the confined problem space of basic addition and subtraction word problems with their explicit correct answers, two questions will be addressed in this study: 1) how do the child-internal factors of oral language skill and nonverbal understanding of mathematical sets interact with the external problem factor of linguistic complexity, and 2) will a gesture scaffold and/or a rewording scaffold allow children with low oral language skill to access their conceptual knowledge of part-whole set relations more readily, and thereby solve the word problems more accurately.

CHAPTER IV

PHASE ONE

Method

In Phase One, two groups of children, one group with good oral language skills and the other group with poor oral language skills, completed a set of basic addition and subtraction problems presented in both oral and nonverbal formats. The oral format tested the hypothesis that linguistic factors are a bottleneck to accurate arithmetic word problem solving in children with poor oral language skill, and the nonverbal problem solving task tested whether nonverbal knowledge of mathematical set relations would be similar in the two groups.

Under the hypothesis that oral language deficits function as the primary bottleneck to accurate problem representation and solution when arithmetic word problems are presented orally, I predicted that, in Phase One, children with poor oral language skill would have greater difficulty solving the orally-presented arithmetic word problems than children with good oral language skill. Under the hypothesis that nonverbal factors also support the development of mathematical cognition, I predicted that the two groups would have similar performance on the nonverbal task, if they are matched on a measure of nonverbal intelligence.

Participants

Thirty first graders participated in Phase One, including 15 children classified as poor language users (PLU) and 15 age mates classified as good language users (GLU). All participants were part of a larger group of children recruited by the University of Iowa Child Language Research Center (CLRC) from school districts that had administered the Iowa Tests of Basic Skills (ITBS) to their first graders. The children in this phase were recruited from two public school districts in the Midwest. One of the participants was from a bilingual home where Spanish was spoken; however the mother was a proficient speaker of English. All families signed consent forms permitting their children to participate in an arithmetic word problem study, as well as release forms for their children's ITBS scores. Participants received $10.00 in compensation per visit, as well as small prizes.

The 15 first graders in the GLU group (10 boys, 5 girls) ranged in age from 82 months to 97 months (M = 89.4, SD = 4.3). The 15 first graders in the PLU group (10 boys, 5 girls) ranged in age from 81 months to 99 months (M = 88.3, SD = 5.7). Table 1 summarizes the background measures for the two groups of participants in Phase One.

Table 1. Background measures for participants in Phase One.

	GLU (n = 15)		PLU (n = 15)	
Measure	M	(SD)	M	(SD)
Age (months)	89.4	(4.3)	88.3	(5.7)
LCOMP(*z*-score)	-0.02	(0.59)	-1.22	(0.50)
NVIQ (*z*-score)	-0.32	(0.74)	-0.52	(0.84)
Simon (high score)	3.93	(1.03)	3.87	(0.92)
Digit Span (Total)	9.80	(1.97)	8.33	(1.92)

Note: See Appendix B for a description of the LCOMP measures. The NVIQ composite was based on two subtests of the WISC-III: Picture Completion and Block Design. The Simon task is a measure of visual memory. Digit Span Total is the sum of Digits Forward and Digits Backward scores from the WISC-III Digit Span Subtest.

Language Skill

Determination of language skill was based on a composite measure of oral language skill (LCOMP) employed by the CLRC (Tomblin, Records, & Zhang, 1996).

(See Appendix B for a list of language measures included in the LCOMP) The GLU group included children who achieved an LCOMP score at or above the 25th percentile (a *z*-score >= -.67), while the PLU group included children who achieved an LCOMP score below the 25th percentile (a *z*-score < -.67). The 25th percentile was selected as the language skill cut off because the purpose of this study was to investigate the ability of children with lower language skill to solve basic word problems in a class-room-like listening environment. Children in the lowest quartile of oral language skill were thought to be most at risk for having difficulty solving word problems in a listening environment.

Age and Nonverbal Matching

The two groups were matched on age (+/- 3 months), and were matched as closely as possible on nonverbal cognitive ability, as measured by a composite standard score (M = 20; SD = 5) from the Picture Completion and Block Design subtests of the Wechsler (NVIQ) Intelligence Scale for Children, 3rd edition (WISC-III; Wechsler, 1991). For two pairs, the age difference exceeded 3 months; however in one pair the child with good language skill was 9 months older, and in the other pair the child with poor language skill was 5 months older. Age mates were recruited from the same school district to control for educational experience.

Based on their LCOMP and NVIQ scores, the two groups contrasted on oral language skill, but were similar in age and on nonverbal cognitive ability. To confirm that the groups were appropriately matched, a multivariate analysis of variance (MANOVA) was conducted, with group (GLU, PLU) as the between subjects factor, and age, LCOMP and NVIQ as the dependent variables. There was a main effect of group ($F(3, 26) = 13.59$, $p < .001$, $\eta^2 = .61$). Planned comparisons indicated that, as expected, the GLU and PLU groups did not differ on age ($F(1, 28) = .38$, $p = .54$, $\eta^2 = .01$) or on NVIQ ($F(1, 28) = .50$, $p = .49$, $\eta^2 = .02$). The two groups did differ, however, on LCOMP ($F(1, 28) = 36.51$, $p < .001$, $\eta^2 = .57$).

It is important to note here that, because the GLU and PLU groups were matched on NVIQ, the GLU group tended to be comprised of average and above average first graders, but not 'superstar' first graders. In other words, despite the significant group difference in language ability, the group of participants as a whole did not span the entire range of language abilities.

Additional Participant Criteria

All participants demonstrated nonverbal cognitive scores within the normal range (WISC-III; Picture Completion and Block Design subtests: composite >= 10) and passed a hearing screening. Thresholds greater than 30 dB in both ears in two frequencies (500, 1000, 2000, and 4000 Hz) resulted in failure. By parent report, no child had a history of sensory impairment, neurological impairment, or a diagnosis of autism or pervasive developmental disorder. The LCOMP language measures, as well as a modification of the Simon visual memory game, the WISC-III digit span subtest, and a hearing screening, were completed near the end of the first grade year. The experimental tasks were completed between May of the first grade year and the end of September of the second grade year. In order to ensure that all children had verbal count sequence skills, children were also asked to count forward from 1 to 10 and backward from 10 to 1. All participants were able to complete these verbal counting tasks in two attempts or less, without using manipulatives.

Procedure and Tasks

The five tasks for Phase One were administered during one session lasting approximately one hour: the Simon visual memory task, the WISC-III digit span subtest, the hearing screening, and the basic addition and subtraction problem tasks (presented in verbal and nonverbal formats). Participants were tested individually in a specially-equipped van, parked in a quiet location at the child's home or school. All LCOMP language measures, NVIQ subtests, and the five tasks for Phase One were administered by an experienced CLRC examiner, who was trained on all tasks by the primary investigator.

Visual Memory Task

A computerized modification of the Simon Task was administered to obtain a measure of visual memory (H. J. Hsu, personal communication, November 10, 2008; Pisoni & Cleary, 2004). After watching a ring of yellow buttons light up in random order, the children were asked to replicate the sequence by clicking on the buttons in the same order. Rather than using the multi-colored ring of buttons from the original Simon game, mono-colored yellow buttons were used, to mitigate the influence of verbal (color name) mediation in a visual memory task. Visual memory was tested as the children attempted to complete increasingly longer sequences, and capacity was measured as the highest level at which a child accurately replicated the sequence in two attempts or less.

Auditory Memory Task

The WISC-III Digit Span subtest was administered to obtain a measure of auditory verbal memory for numbers. For the Digits Forward section, children heard a series of single digit numbers, and were asked to repeat the series in the same order as presented. For the Digits Backward section, children heard a series of single digit numbers, and were asked to repeat the series in reverse order. Digit span was computed as the sum of the number of correct forward and backward trials.

Verbal and Nonverbal Arithmetic Problems

The verbal and nonverbal arithmetic problem tasks were patterned after Jordan, Levine, and Huttenlocher (1995). Each problem set (verbal and nonverbal) contained 5 two-term addition problems, 5 two-term subtraction problems, and 4 three-term problems that combined addition and subtraction procedures (see Appendix C for the problem sets and protocols). The matched sets of nonverbal and verbal arithmetic problems contained augends (the first term in an addition problem), addends (the amount added to the augend), minuends (the first term in a subtraction problem) and subtrahends (the amount subtracted from the minuend) ranging from 1 to 9, and sums or differences less than 10. Problem order was randomized, and all participants received the same problems in the same order for both the nonverbal and verbal tasks. Half of the children received the verbal problem set first, and the remaining children were given the matching nonverbal problems first, to control for any influence of completing one problem set before the other.

Nonverbal Arithmetic Task

For this task, the child was seated across the table from the examiner. Both child and examiner manipulated a set of yellow plastic disks on their own workspace mat. The examiner also used an opaque rectangular cover with an opening in one side, so that she could easily put in or remove disks (See Figure 2).

Figure 2. Nonverbal Arithmetic Task (Problem: 5–3=2).

To teach the procedure, the examiner placed three disks on her mat, in view of the child, and said, “See this? Now watch what I do.” The rectangular cover was then placed over the disks. Next, the examiner placed three disks on the child’s mat, lifted the cover to show the disks on her own mat, and said, “See, yours is just like mine.” The demonstration was repeated, except this time the child was asked to place the appropriate number of disks on his mat after the examiner placed and covered her quantity of disks. The examiner specifically instructed, “Make yours just like mine.” If the child placed an incorrect number of disks on his mat, his response was corrected, and the demonstration item was repeated. All participants were told that they could solve the problems in any way they wanted, including ‘using your fingers’, using the yellow plastic disks, counting, or ‘doing it in your head’. Flexibility in solving the problems via a method preferred by individual children was encouraged during the practice set.

For the addition problems, the examiner placed a horizontal line of disks corresponding to the augend on her mat, and said, “See this? Now watch what I do next.” After covering the line of disks, the experimenter placed a second line of disks corresponding to the addend in view of the child, and then slid them under the cover through the opening, one at a time. The examiner then asked the child to “Make yours just like mine.” The child was then expected to place the appropriate number of disks, representing the sum, on his mat.

For the subtraction problems, the experimenter placed a line of disks corresponding to the minuend in view of the child and said, “See this? Now watch what I do next.” After covering the line of disks, the experimenter slid the number of disks corresponding to the subtrahend out from under the cover, one by one, and the child was expected to reproduce the difference on his mat. The examiner did not supply verbal quantity labels for the minuend and subtrahend, nor did she ask the child to provide verbal responses. However a verbal response from a child, rather than a disk manipulation, was accepted. At no time were the two terms of the problems in view simultaneously, similar to the situation in verbal word problems where a child is expected to remember quantities that were presented earlier in the problem.

Verbal Word Problem Task

This task was the verbal complement to the nonverbal arithmetic task. The matching nonverbal and word problem sets contained the same 14 addition and subtraction problems. The verbal problems were presented in *Combine 1* and *Change 2* formats.

After each participant was seated at the table facing the examiner, two easy practice problems were presented (one *Combine 1* and one *Change 2*), to ensure that the child understood the task. After securing the participant's attention, each word problem was read once, with natural inflection. All problems were presented orally, using traditional formal wording, and the children were asked to provide an answer to each problem. (See Appendix C for the verbal problem set). All participants were told that they could solve the problems in any way they wanted, including 'using your fingers', manipulating the same yellow plastic disks used in the nonverbal task, counting, or 'doing it in your head'. Flexibility in solving the problems via a method preferred by individual children was encouraged during the practice set.

Coding and Analysis of Verbal and Nonverbal Problem Accuracy

All sessions were video recorded for later analysis using a Panasonic PV-L857 Palmcorder. Prior to coding and analysis, the analog videos were converted to digital format using Roxio Easy Media Creator 8 Media Import and VideoWave software. Backup audio recordings were made with a Marantz PMD101 audio cassette recorder and an Electro-Voice N/D267AS dynamic cardioid vocal microphone. The examiner also completed an online hand-coded form for each child. Data were coded based on the child's verbal or nonverbal responses, and solution accuracy (proportion correct). The accuracy data were used to compare verbal and nonverbal performance on the matched sets of problems.

Reliability

An undergraduate research assistant, trained in coding the nonverbal and verbal tasks, independently recoded a randomly selected subset of the 30 video files (6 video files: 3 GLU and 3 PLU files, representing 20 percent of the participants). Point-to-point agreement between the trained examiner and the independent coder was calculated for two measures: the child's response to each problem (nonverbal problems: 98.81%; verbal problems: 100%), and the accuracy of each response, based on the child's final response (nonverbal problems: 97.62%; verbal problems: 100%).

Analyses

For Phase One, preplanned analyses included two independent samples *t*-tests to test group differences in proportion correct on the verbal and nonverbal arithmetic problem tasks.

Results

Unless otherwise specified, an alpha level of .05 was used for all statistical tests, *p* values are two-tailed, and effect sizes (d or η^2) are reported where appropriate.

Verbal and Nonverbal Arithmetic Problem Tasks

Table 2 summarizes the means and standard deviations of the two groups for the verbal and nonverbal problem solving tasks in Phase One. Two independent samples *t*-tests were conducted to test two predictions: 1) children with good language skill would solve the verbal word problems with greater accuracy than children with low language skill, but 2) children with good and poor oral language skill would not differ in solution accuracy on the nonverbal task. Because Levene's test of equality of variances was significant for the verbal problems ($F(1, 28) = 7.40$, $p < .05$), results for equal variances not assumed are reported for the verbal task. A significant group difference was found on the verbal task ($t(21.44) = 2.88$, $p < .01$, $d = 1.10$), but no group difference was found on the nonverbal task ($t(28) = .42$, $p = .68$, $d = 0.13$; see Figure 3).

Table 2. Means and standard deviations for the verbal and nonverbal arithmetic problem solving tasks in Phase One.

	GLU (n = 15)		PLU (n = 15)	
Measure	M	(SD)	M	(SD)
Nonverbal Problems	0.73	(0.18)	0.70	(0.30)
Verbal Problems	0.80	(0.11)	0.63	(0.20)

Counterbalancing

Because half of the participants received the nonverbal task first, and the remaining half received the nonverbal task second, it was possible that an order effect influenced the results for one or both groups. Two one-way analyses of variance were conducted (verbal problem accuracy, nonverbal problem accuracy), with administration order as the between groups factor (verbal first, verbal second) and proportion correct as the dependent variable. For both analyses there was no significant difference between the two administration orders: Verbal accuracy ($F(1, 28) = .12$, $p = .73$, $\eta^2 = .00$); nonverbal accuracy ($F(1, 28) = .00$, $p = 1.00$, $\eta^2 = .00$).

To summarize, these results supported both predictions in that, although the means for the PLU group were lower than those for the GLU group on both tasks, the children with good oral language skill performed significantly better than the children with low oral language skill on the verbal arithmetic word problem task, but the two groups performed similarly on the nonverbal arithmetic problem task.

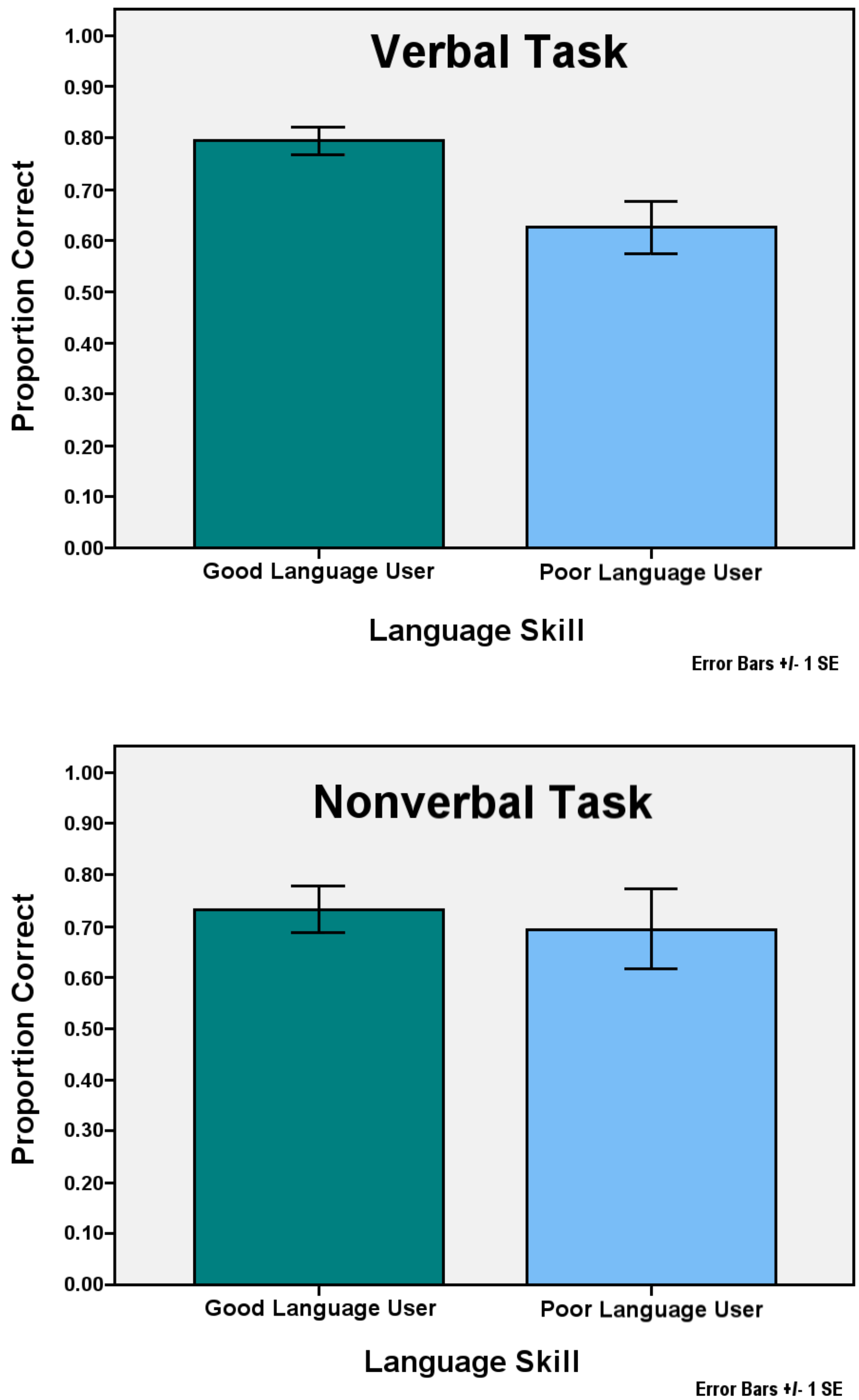

Figure 3. Graph of Verbal and Nonverbal Arithmetic Problem Tasks.

Potential Influencing Factors

It could be argued that children with poor oral language skill have an underlying deficit in working memory that would influence their performance on both the verbal and nonverbal tasks in Phase One. To explore group differences in visual memory and auditory memory skill, two independent samples *t*-tests were conducted. A significant group difference was found for the Digit Span auditory memory measure ($t(28) = 2.07$, $p < .05$, $d = 0.75$), with the GLU group performing better than the PLU group. It is therefore possible that the group difference on the verbal problem task was due to the inability of the children with poor language skill to recall the maximum of three numbers necessary to complete the three-term problems. A visual inspection of the forward digit span data, however, revealed that no child, regardless of language skill, recalled fewer than 4 digits. All 6 children whose highest forward recall was 4 digits, however, belonged to the PLU group. There was no group difference on the Simon visual memory task ($t(28) = .187$, $p =$ ..85, $d = 0.06$). These results indicate that working memory is modality specific, and suggest that we should exert caution in referring to 'working memory' as a single construct.

<u>Discussion</u>

In Phase One I compared good and poor language users' performance on basic addition and subtraction problems presented in both oral and nonverbal formats. These verbal and nonverbal arithmetic problem tasks were patterned after Jordan, Levine, and Huttenlocher (1995). It was necessary to establish that first graders with poor oral language skill have greater difficulty solving orally-presented arithmetic word problems than their peers, but can still access an underlying knowledge of mathematical set relations in a nonverbal arithmetic task. Theoretically, this ability to access nonverbal mathematical knowledge would then help children make use of a nonverbal gesture scaffold to access the underlying problem structure of an arithmetic word problem.

As predicted, and as Jordan and colleagues (1995) found, children with good language skill solved the verbal word problems with greater accuracy than children with low language skill, but the two groups did not differ in solution accuracy on the nonverbal task. So, oral language skill was related to these first graders' ability to solve orally-presented arithmetic word problems accurately, but even the children with poor oral language skill were able to access an underlying knowledge of the relations between sets in order to solve nonverbal arithmetic problems comparably to children with good oral language skill.

The finding that the GLU and PLU groups differed in the Digit Span auditory verbal memory task, but did not differ in the Simon visual memory task supports MacDonald and Christiansen's (2002) argument that verbal working memory is not a separate entity from language function and both are tapping the same underlying language skill. Under their theoretical perspective, we shouldn't be surprised if children with low language skill do poorly on a verbal working memory task. This then changes the argument from one of a causal relationship where memory capacity drives language skill to one where we consider the quality of the representation in a given realm (i.e. the quality of the knowledge base for a word problem task) rather than the capacity of that realm.

Under the same theoretical perspective, we would then expect that groups that perform comparably on the nonverbal task would also perform comparably on a test of visual memory. On the other hand, the children with poor oral language skill might have relied on a visual memory capacity to drive their performance on the nonverbal problem task. The children with good oral language skill might also have relied on visual memory, but it is also possible that these children used verbal mediation to complete the nonverbal task. Likewise, group differences on the Digit Span auditory memory measure might have contributed to the difference in group performance on the orally-presented arithmetic word problems. No matter which theoretical perspective we assume, the results of the nonverbal and verbal arithmetic word problem tasks in Phase One set the stage for considering possible gesture and rewording scaffolds that help children with low oral language skill access both the nonverbal and the verbal components of arithmetic word problems.

CHAPTER V

PHASE TWO

Method

The purpose of Phase Two was to identify those first graders who had difficulty solving basic verbal arithmetic word problems (presented orally), and to determine the specific problem types that were difficult for them to solve. I predicted that children with poor oral language skill would have greater difficulty than children with good oral language skill on all three problem types.

Participants

The same 30 children who participated in Phase One continued into Phase Two. In addition, seven more first graders who qualified as poor language users (an LCOMP *z*-score < -.67) were recruited from a third public school district in the Midwest. These seven children were added to the PLU group to ensure that we would have baseline measures and sufficient power for Phase Three. This resulted in two groups: 15 first graders in the GLU group (10 boys, 5 girls), ranging in age from 82 months to 97 months (M = 89.4, SD = 4.3) and 22 first graders in the PLU group (16 boys, 6 girls), ranging in age from 81 months to 99 months (M = 88.6, SD = 5.0). Two of the participants were from bilingual homes where Spanish was spoken; however in both cases the mother was a proficient speaker of English. All families signed consent forms permitting their children to participate in an arithmetic word problem study. Participants received $10.00 in compensation per visit, as well as small prizes. Table 3 summarizes the background measures for this larger set of 37 children.

Table 3. Background measures for participants in Phase Two.

Measure	GLU (n = 15) M	GLU (n = 15) (SD)	PLU (n = 22) M	PLU (n = 22) (SD)
Age (months)	89.4	(4.3)	88.6	(5.0)
LCOMP(*z*-score)	-0.02	(0.59)	-1.33	(0.54)
NVIQ (*z*-score)	-0.32	(0.74)	-0.55	(0.78)
Simon (high score)	3.93	(1.03)	3.81	(1.01)
Digit Span (Total)	9.80	(1.97)	7.95	(1.84)

Note: See Appendix B for a description of the LCOMP measures. The NVIQ composite was based on two subtests of the WISC-III: Picture Completion and Block Design. The Simon task is a measure of visual memory. Digit Span Total is the sum of Digits Forward and Digits Backward scores from the WISC-III Digit Span Subtest.

Task and Procedure

Basic Arithmetic Word Problem Task

This task consisted of 18 problems, including five each of the *Compare 3*, *Compare 5,* and *Compare 6* problem types that were expected to be difficult for first graders to solve (See Appendix D for the Phase Two problem sets). These three types were selected based on Riley and Greeno's (1988) finding that first grade children performed with no greater than 40% accuracy on these particular problem types, even with manipulative support. All three problem types were also conducive to the rewording and gesture scaffolds planned for Phase Three. Because both the *Compare 3* and *Compare 6* problems required an addition process for correct solution, but the *Compare 5* problems required a subtraction process, three filler problems were included that also required subtraction, in order to achieve a better balance in the problem set between the two mathematical operations. In addition, the math facts used in the filler problems were easier than those used in the *Compare* problems, to provide a measure of success for all participants.

Each individual problem contained a unique set of numbers drawn from a pool of 24 nontrivial number facts: 12 addition facts and the 12 matching inverse addition facts. For the five *Compare 3* and the five *Compare 6* problems, number facts were quasi-randomly selected from the pool. For the five *Compare 5* problems, five additional addition facts were quasi-randomly selected, and converted to subtraction facts. To complete the set of 15 problems, care was taken that all 12 number facts were utilized, in addition to three randomly selected inverse facts. For all three problem types, the augends, addends, minuends and subtrahends ranged from 2 to 9, and all sums and differences were ≤ 10. Number facts containing the same numeral for more than one of the terms were specifically not included, to eliminate ambiguity when scoring (e.g. $6 - 3 = 3$).

Five different word problems of parallel structure were created for each of the 3 problem types. Because each word problem contained a unique pair of numbers, it was necessary to ensure that differences in solution accuracy across the three problem types did not arise from differences in difficulty of the particular number pairs used within the problem sets. Therefore, *two* different sets of number facts (e.g. augend + addend) were constructed, as described above. For a given problem (e.g. the first instance of *Compare 3*), half of the children received one number fact, and the remaining children received the other number fact (See Appendix D).

The presentation order for each set of 18 word problems was randomized, and then checked so that no more than 2 problems of any given type were presented sequentially. All participants within each problem set received the same presentation order.

Procedure

The arithmetic word problem task for Phase Two was administered by the investigator during one session lasting approximately one hour. Participants were tested individually in a specially-equipped van, parked in a quiet location at the child's home or school. After completing 11 of the 18 problems, all participants took a 5 minute break. Other short breaks were taken as needed.

After each participant was seated at the table facing the examiner, two easy practice problems were presented, to ensure that the child understood the task (one *Combine 1* problem and one *Change* problem). After securing the participant's attention, the examiner read each word problem to the child with natural inflection, and repeated it one time. All problems were presented orally, using traditional formal wording. The children were asked to provide an answer to each problem, and then to explain their solution to an animal puppet who was 'trying to learn how to do math word problems': e.g. "Tell Ricky (or explain to Ricky) how you figured out your answer." In this way the child was given the opportunity to explain his or her approach to solving the problem.

The examiner placed problem-specific agent and patient figures and objects (i.e. manipulatives) in front of the child for each exemplar, as shown in Figure 4. For example, Mario and Tweety Bird figures, along with a set of Lego blocks, were arrayed before the child for a *Compare 3* problem where the child was asked to solve 'How many Legos does Tweety Bird have?' The purpose of these figure and object manipulatives was to support potential working memory limitations in first graders who must keep track of two numerical quantities in order to solve these word problems successfully (Carr & Hettinger, 2003, p. 39).

The examiner assessed the child's familiarity with the figures and objects while each problem was being set up. All participants were told that they could solve the problems in any way they wanted, including 'using your fingers', using the

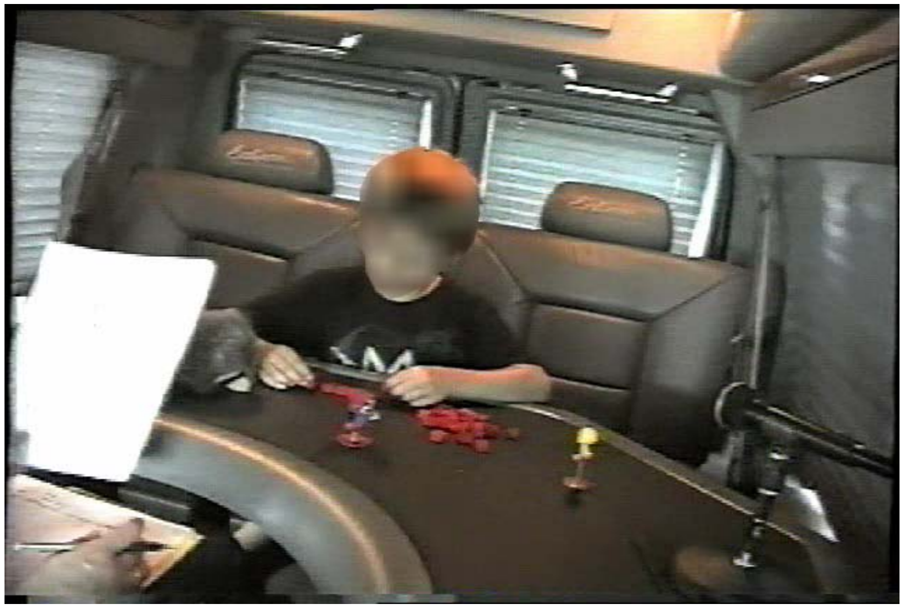

Figure 4. Word Problem Task (Phases Two and Three).

manipulatives, counting, or 'doing it in your head'. Flexibility in solving the problems via a method preferred by individual children was encouraged during the practice set.

Differential reinforcement was not provided for any of the problems, but all children were given general reinforcement by both the puppet and the examiner for their hard work and explanations.

Coding and Analysis of Problem Accuracy

All sessions were video recorded for later analysis using a Panasonic PV-L857 Palmcorder. Prior to coding and analysis, the analog videos were converted to digital format using Roxio Easy Media Creator 8 Media Import and VideoWave software. Backup audio recordings were made with a Marantz PMD101 audio cassette recorder and an Electro-Voice N/D267AS dynamic cardioid vocal microphone. The principal investigator also completed an online hand-coded form for each child. Data were coded based on verbal responses and solution accuracy (proportion correct). The accuracy data were used to compare performance on the three problem types.

Reliability

An undergraduate research assistant, trained in coding arithmetic word problem tasks, independently recoded a randomly selected subset of the 37 video files (6 video files: 3 GLU and 3 PLU files, representing 16 percent of the participants). Point-to-point agreement between the principal investigator and the independent coder was calculated for two measures: the child's response to each problem (98.15%), and the accuracy of each response, based on the child's final response (98.15%).

Analyses

For Phase Two, preplanned analyses included:

1) a mixed model (split plot univariate), repeated measures analysis of variance (ANOVA), with group (PLU, GLU) as the between subjects factor, problem type (*Compare 3*, *Compare 5*, and *Compare 6*) as the within subjects factor, and proportion correct on each problem type as the dependent variable.

2) a series of *t*-tests, to determine 1) whether the group matching criteria (LCOMP and NVIQ) were maintained when the 7 additional participants were added to Phase Two, and 2) whether there were group differences in visual memory or auditory memory skill that might have influenced group performance on word problem solution accuracy.

Results

Unless otherwise specified, an alpha level of .05 was used for all statistical tests, p values are two-tailed, and effect sizes (d or η^2) are reported where appropriate.

Solution Accuracy for Three Word Problem Types: All Participants

Table 4 summarizes the means and standard deviations for solution accuracy (proportion correct) on the *Compare 3, Compare 5, Compare 6,* and *Filler* problem types for the two groups of participants. Note that means for the *Compare* problem types are all less than 45 percent, indicating the difficulty of these three problem types for first graders. The omnibus test of the effect of the three *Compare* word problem types on

Table 4. Means and standard deviations for the four problem types in Phase Two.

	GLU (n = 15)		PLU (n = 22)	
Problem Type	M	(SD)	M	(SD)
Compare 3	0.44	(0.43)	0.28	(0.32)
Compare 5	0.23	(0.32)	0.16	(0.27)
Compare 6	0.17	(0.31)	0.13	(0.24)
Fillers	0.87	(0.25)	0.90	(0.18)

solution accuracy yielded a significant main effect of problem type ($F(2, 70) = 10.22$, $p < .001$, $\eta^2 = .23$), indicating an expected difference in the relative difficulty of the three problem types. Mauchly's Test of Sphericity was not significant ($p = .75$), indicating that the assumptions of compound symmetry were not violated. The main effect of group was not significant ($F(1, 35) = 1.03$, $p = .32$, $\eta^2 = .03$). The interaction between problem type and group was also not significant ($F(2, 70) = .76$, $p = .47$, $\eta^2 = .02$), indicating that both groups of children performed similarly across the types of arithmetic word problems.

Figure 5 illustrates the proportion correct for all problem types in Phase Two. For all participants combined, pairwise comparisons of the three *Compare* problem types yielded significant differences between *Compare 3* and *Compare 5* problems ($t(36) = 3.46$, $p < .01$, $d = 0.48$) and between *Compare 3* and *Compare 6* problems ($t(36) = 3.94$, $p < .001$, $d = 0.63$), with *Compare 3* being easier than both *Compare 5* and *Compare 6*.

There was no significant difference between *Compare 5* and *Compare 6* problems at the .05 level (p = .37).

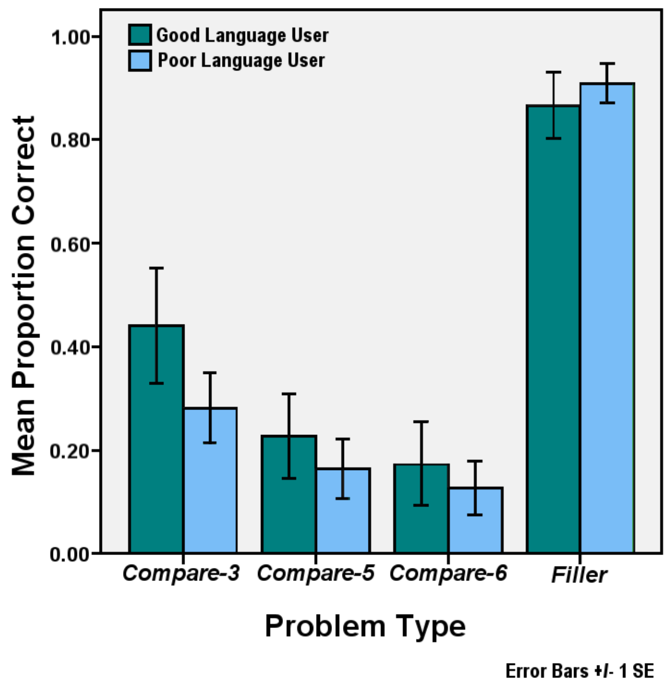

Figure 5. Proportion Correct for the Four Problem Types in Phase Two

Note: The range of chance performance is < 0.33 proportion correct for the *Compare* problem types.

Poor Language User Performance: Three Problem Types

Even though the main effect of group was not significant, one of the objectives of Phase Two was to select a problem type that was difficult for children with low oral language skill, so a one-way repeated measures analysis of variance (ANOVA) was performed on the PLU data, with one within subjects factor (problem type) and three levels (*Compare* 3, *Compare* 5, *Compare* 6). This test yielded a marginal main effect of problem type (F(2, 42) = 3.32, p = .05, η^2 = .14). Mauchly's Test of Sphericity was not significant (p = .93). Pairwise comparisons resulted in a significant difference between

Compare 3 and *Compare 6* problems ($p < .05$), a marginal difference between C*ompare 3* and *Compare 5* problems ($p = .07$) and no significant differences between *Compare 5* and *Compare 6* problems ($p = .58$). All means for proportion correct were less than 30 percent, with *Compare 3* problems being the least difficult for the PLU group, followed in difficulty by *Compare 5* and *Compare 6* problem types (See Figure 5).

Good Language User Performance: Three Problem Types

Although a group * problem type interaction was not found, performance by the GLU group was explored because the marginal effect of problem type in the PLU group would suggest that the omnibus effect of problem type was driven by the performance of the GLU group. Therefore, a second one-way repeated measures analysis of variance (ANOVA) was performed on the GLU data; with problem type as the within subjects factor, with three levels (*Compare* 3, *Compare* 5, *Compare* 6). This test yielded a significant main effect of problem type ($F(2, 28) = 7.03$, $p < .01$, $\eta^2 = .33$). Mauchly's Test of Sphericity was not significant ($p = .47$). Pairwise comparisons yielded a significant difference between *Compare 3* and *Compare 5* problems ($t(14) = 3.23$, $p < .05$, $d = .57$) and between *Compare 3* and *Compare 6* problems ($t(14) = 3.08$, $p < .05$, $d = .72$) for the GLU group, with *Compare 3* problems being significantly easier than either *Compare 5* or *Compare 6* problems. There was no significant difference between *Compare 5* and *Compare 6* problems ($p = .47$; See Figure 5).

Filler Word Problems: All Participants

The results of these ANOVAs indicated that all three problem types were difficult for both groups of children, with *Compare 3* being somewhat less difficult than *Compare 5* and *Compare 6* word problems, especially for the GLU group. This low overall performance by both groups raised the question of whether this cohort of first graders was capable of successfully solving basic arithmetic word problems. To answer this question, a second mixed model analysis of variance (ANOVA) was performed, with group (PLU, GLU) as the between subjects factor, and problem type as the within subjects factor, with four levels (*Compare 3*, *Compare 5*, *Compare 6*, and *Fillers*). The dependent variable was the proportion correct on each problem type.

The omnibus test of the effect of oral language skill on arithmetic word problem solution accuracy yielded a significant main effect of problem type ($F(3, 105) = 82.67$, $p < .001$, $\eta^2 = .70$), indicating an expected difference in the relative difficulty of the four problem types (See Figure 5). Mauchly's Test of Sphericity was not significant ($p = .242$). The main effect of group was not significant ($F(1, 35) = .60$, $p = .45$, $\eta^2 = .02$). The interaction between problem type and group was also not significant ($F(3, 105) = 1.22$, $p = .31$, $\eta^2 = .03$), indicating that both groups of children performed similarly on the arithmetic word problems.

Pairwise comparisons yielded significant differences between *Compare 3* and *Compare 5* problems ($p < .01$, $d = .49$), between *Compare 3* and *Compare 6* problems ($p < .001$, $d = .64$), and between *Compare 3* and the *Fillers* ($p < .001$, $d = -1.87$). There were also significant differences between the *Fillers* and both *Compare 5* and *Compare 6* problem types (ps < .001). There was no significant difference between *Compare 5* and *Compare 6* problems at the .05 level. The *Filler* problem type was easier than the other three problem types for both groups of participants, indicating that these first graders were capable of solving an easier *Filler* arithmetic word problem.

Poor Language User Performance: Filler Word Problems

Because the group of interest in Phase Two was the PLU group, a one-way repeated measures analysis of variance (ANOVA) was performed on the PLU data, with one within subjects factor (problem type) and four levels (C3, C5, C6, Filler). This analysis yielded a significant main effect of problem type ($F(3, 63) = 58.97$, $p < .001$, $\eta^2 = .74$). Mauchly's Test of Sphericity was not significant ($p = .23$). Pairwise comparisons resulted in significant differences between the *Fillers* and the other three problem types (*Compare 3*, *Compare 5*, and *Compare 6*; ps < .001), with *Fillers* being easier than the other three problem types. As before, there was a significant difference between *Compare 3* and *Compare 6* ($p < .05$) and a marginal difference between *Compare 3* and *Compare 5* ($p = .07$). There was no significant difference between *Compare 5* and *Compare 6* problems ($p = .58$). These results confirmed that children with low oral language skill were capable of solving an easier *Filler* arithmetic word problem.

Two Sets of Number Facts

Recall that two different sets of number facts (e.g. augend + addend) were created, in order to ensure that differences in solution accuracy across problem types did not arise from differences in difficulty of the particular number pairs used within the problem sets (e.g. for the first instance of a *Compare 3* problem, half of the children received one number fact, and the remaining children received the other number fact). Four independent samples *t*-tests were conducted (one for each problem type; equal variances not assumed for the *Compare 3* problem type) to test whether the two sets of number facts differed in difficulty. The children who received Set 1 did not perform significantly differently than the children who received Set 2 for any of the four problem types (ps > .36).

Potential Influencing Factors

To test 1) whether the group relationships in LCOMP and NVIQ were maintained when the 7 additional participants were added to Phase Two and 2) whether there were group differences in visual memory or auditory memory skill that might have influenced group performance on word problem solution accuracy, a series of independent samples *t*-tests were conducted. See Table 3 for group means and standard deviations. Although the PLU group means were lower than the GLU group means for all four measures, *t*-tests of the four dependent variables revealed significant differences only for the two language-related measures: LCOMP ($t(35) = 6.99$, $p < .001$, $d = 2.34$) and Digit Span ($t(35) = 2.91$, $p < .05$, $d = 0.98$). The GLU group performed better than the PLU group on both measures. Results for the remaining two *t*-tests were not significant: NVIQ (p = .38) and Simon-High (p = .74). These results confirm that the group relationships in LCOMP and NVIQ were maintained for the larger participant group in Phase Two, and also confirm the findings in Phase One that children with low oral language skill performed similarly to children with good oral language skill on nonverbal measures, but performed significantly lower on language-based measures.

Discussion

The purpose of Phase Two was to identify those first graders with poor oral language skill who had difficulty solving basic verbal arithmetic word problems (presented orally), and to determine the specific problem type(s) that were difficult for them to solve. The selected children and problem type would then move on to Phase Three, to test the effects of rewording and gesture scaffolds. Based on the results of Phase Two, one problem type would be selected for the scaffolding phase; this problem type would need to be difficult enough so that there would be room for improvement in a test of the three scaffold conditions.

A set of *Compare 3, Compare 5,* and *Compare 6* word problems was presented to children in both the GLU and PLU groups. Based on work by Riley and Greeno (1988), it was expected that these three problem types would be difficult for first graders, and the current study confirmed the earlier findings. Contrary to my prediction, however, there was no significant difference in performance between the GLU and PLU groups, indicating that these problem types were difficult for both groups of first graders in this study. It is possible that this null finding resulted from the fact that, because the two groups were matched on NVIQ, the GLU group did not contain ‘superstar’ first graders at the higher end of the language skill continuum. As a result, the GLU group performed within the range of chance on the *Compare 5* and *Compare 6* problem sets and the PLU group performed within the range of chance on all three problem types. In particular this could be interpreted as a floor effect for the PLU group on the *Compare 5* and *Compare 6* problem types.

Because the overall goal of this study was to test rewording and gesture scaffolds on children with low oral language skill, relative performance on the three problem types for the PLU group was examined. The PLU group performed within the range of chance on *Compare 3, Compare 5,* and *Compare 6* problems, making all three problem types potential candidates for Phase Three. The finding that the PLU group achieved less than 30 percent accuracy on all three problem types, however, also raised the possibility that an arithmetic word problem task was simply too difficult for children with low oral language skill. This concern for floor effects was partially allayed, because this PLU group did perform with greater than 90% accuracy on the easier filler problems. Given these results, the challenge was to select a problem type that would not be too difficult for the PLU group, but would also allow room for improvement. An in depth discussion of the rationale for choosing a problem type for Phase Three is presented in the following chapter.

CHAPTER VI

PHASE THREE

Method

Only the participants from the PLU group in Phase Two were eligible to continue into Phase Three, because the purpose of the third phase was to investigate the influence of rewording and gesture scaffolds on the ability of first graders with low oral language skill to comprehend and solve oral word problems. The *Compare 6* problem type was selected for Phase Three, based on the rationale described later. A rewording scaffold was chosen to make the linguistic component of the word problem more accessible to children with low language skill, and a gesture scaffold was chosen to make the nonverbal conceptual and procedural relations between the sets more accessible. These two scaffolds resulted in four test conditions: 1) a baseline condition with traditional, formal wording and no scaffolds (Scaffold A), 2) a problem rewording condition with no gesture (Scaffold B), 3) a traditional wording condition with gesture (Scaffold C), and 4) a condition with both problem rewording and gesture scaffolds (Scaffold D). (See Figure 1).

Specifically, I predicted that children with low oral language skill would solve reworded *Compare 6* word problems more accurately than traditionally worded *Compare 6* problems. In addition I predicted that this group of children would also be able to solve *Compare 6* word problems presented with an accompanying procedural gesture more accurately than *Compare 6* problems presented without gestures. This effect of gesture would be most evident in the rewording condition, but might also be evident in the traditional wording condition, if the gesture is capable of changing children's representation of the problem structure. Finally, if children with low oral language skill are able to integrate the information from the rewording and gesture scaffolds, I predicted that the combination of two scaffolds would result in higher solution accuracy than any of the other 3 conditions.

Participants

Of the 22 children in the PLU group in Phase Two, only 20 children moved on to Phase Three. One of the children moved out of the study area, and was no longer available to continue. Another child was disqualified because he consistently used an 'always add' strategy for all three problem types in Phase Two, regardless of whether a problem type required an addition or subtraction process. The remaining 20 children's accuracy on the *Compare 6* problem set in Phase Two ranged from 0 out of 5 correct to 3 out of 5 correct (M = 0.10, SD = 0.20). These scores allowed all 20 children room to improve within the scaffolding conditions of Phase Three. Therefore, 20 children with poor oral language skill (14 boys, 6 girls), ranging in age from 81 months to 95 months (M = 88.1, SD = 4.7), continued into Phase Three. Participants received $10.00 in compensation per visit, as well as a larger prize at the end of the study. Table 5 summarizes the group means and standard deviations for all language and nonverbal measures for this set of 20 children.

Selection of Problem Type

In Phase Two, children solved three types of word problems, *Compare 3*, *Compare 5,* and *Compare 6*. As reported above, the PLU group achieved less than 30 percent accuracy on all three problem types, thereby making all three problem types potential candidates for the scaffolding conditions. However, on the *Compare 3* problem set, two of the children with low language skill achieved 80 percent or better accuracy, thereby reducing the number of participants that had room to improve in the scaffolding conditions. A final decision to use *Compare 6* problems for Phase Three was based on the following rationale:

1) During Phase Two a number of the participants expressed that they did not

Table 5. Background measures for participants in Phase Three.

	LN (n = 9)		LI (n = 11)	
Measure	M	(SD)	M	(SD)
Age (months)	86.2	(4.1)	89.6	(4.7)
LCOMP(*z*-score)	-0.91	(0.16)	-1.65	(0.55)
NVIQ (*z*-score)	-0.31	(0.63)	-0.65	(0.89)
Simon (high score)	3.89	(1.05)	.3.91	(0.94)
Digit Span (Total)	8.22	(1.72)	7.45	(1.64)

Note: See Appendix B for a description of the LCOMP measures. The NVIQ composite was based on two subtests of the WISC-III: Picture Completion and Block Design. The Simon task is a measure of visual memory. Digit Span Total is the sum of Digits Forward and Digits Backward scores from the WISC-III Digit Span Subtest.

understand what 'fewer than' meant, and several of the children interpreted the phrase to mean 'none'. A *Compare 6* problem contains the semantic phrase 'fewer than', but requires an addition process in order to solve the problem accurately. The linguistic phrase 'fewer than' might conflict with a procedural gesture that represents the addition of two sets. Therefore a gesture condition with the traditionally-worded *Compare 6* problem type would be a stringent test of the ability of the gesture to change a child's underlying representation of the problem structure.

A potential rewording scaffold for *Compare 6* problems would be to use the *Compare 3* problem wording, which contains the semantic phrase 'more than' and, like the *Compare 6* problem, requires an addition process in order to solve the problem accurately. By incorporating the *Compare 3* problem wording into the Phase Three scaffolding conditions, I would also be able to test the effect of gesture on this problem type.

2) By contrast, the *Compare 5* problem type contains the semantic phrase 'more than', and requires a subtraction process in order to solve the problem accurately. Based on limited pilot data of children who were good language users, I felt that it would be more difficult to reword this problem type so that an accompanying procedural 'subtract' gesture scaffold would facilitate problem comprehension.

3) Continuing the *Compare 3* problem type into Phase Three would have resulted in two fewer participants, as well as fewer rewording options (i.e. the option of substituting 'more than' for 'less than' would not have been available). In addition, traditionally worded *Compare 3* problems were particularly easy for both kindergarten and first grade pilot subjects with higher oral language skills. Only one of eight pilot first graders did not perform well on the traditional wording, the rewording, or the rewording + gesture conditions.

4) A first grade pilot child with average oral language skills had difficulty with traditional and reworded *Compare 6* problems, but responded well to a gesture and rewording scaffold.

5) And finally, under the theoretical perspective that, over time and educational experience, children build a mathematical schema for each problem type, and given prior evidence that *Compare 3* problems tend to be easier than *Compare 5* and *Compare 6* problems for first graders as a whole, selecting a relatively more difficult problem type (e.g. *Compare 5* or *Compare 6*) would ensure that few of the participants had already developed a reliable schema for the problem. If a reliable schema had not yet been constructed, then children would have to rely on comprehension of the linguistic message and/or the procedural gesture in order to access the underlying problem structure.

Given these factors, I selected the *Compare 6* problem type for the scaffolding conditions in Phase Three.

Stimulus Creation

See Appendix E for Phase Three word problem sets.

Rewording Scaffold

The goal of a rewording scaffold is to make the linguistic message more available to a child with low language skill. The following protocol was used to create the *Compare 6* rewording scaffold:

1) An introductory sentence, orienting children to the specific problem situation and characters, was followed by a statement of the question that the child was expected to solve. This 'question first' rewording provided up front information about the unknown quantity, as well as a rationale under which to view the problem structure (Thevenot, et al, 2007).

2) *Compare 3* 'more than' wording was substituted for the traditional *Compare 5* phrase 'less than'. The 'more than' semantic phrase was predicted to make the underlying problem structure (i.e. the process of adding two sets) more available to children with low oral language skill.

3) In addition, the use of *Compare 3* wording resulted in all characters being represented explicitly as nominals rather than pronouns. Because children with language impairment are known to have difficulty with anaphoric reference (e.g. Chien & Wexler, 1990; van der Lely & Stollwerck, 1997), this rewording replaced potentially confusing pronouns with noun referents.

4) Redundant information was also provided to make the overall wording more explicit (Kintsch, 1994a). For example, verbs were included in optional settings: 'Monkey ate 2 apples. Rhino ate 7 more apples than Monkey (*ate*).'

Gesture Scaffold

A procedural gesture was created, to make the nonverbal set relations of a *Compare 6* problem more transparent, and to indicate the addition procedure that is required by that problem type. Figure 6 contains an example of a *Compare 6* problem, presented in traditional wording, reworded scaffold, and traditional wording accompanied by the gesture scaffold. The left hand represents the augend, the right hand represents the addend, and both hands, first moved toward each other and then cupped together, represent the process of addition and the final solution sum. All *Compare 6* problems were presented using the same gesture scaffold. Under the rewording + gesture scaffold condition, the gestures were scripted to the same phrases as in the traditional wording + gesture condition. One additional gesture was added to the rewording condition: As I read the 'question-first' statement (e.g. 'You need to figure out how many ice cream cones *Bunny* ate.'), I pointed to the character named in the statement. No other gestures were produced while I presented the word problems.

Traditional Wording: Duck ate 7 ice cream cones. He ate 3 fewer ice cream cones than Bunny. How many ice cream cones did Bunny eat?

Reworded: Duck and Bunny were eating ice cream cones. You need to figure out how many ice cream cones Bunny ate. Duck ate 7 ice cream cones. Bunny ate 3 more ice cream cones than Duck ate. How many ice cream cones did Bunny eat?

Traditional Wording + Gesture: Present left hand cupped, palm up over Duck (scripted to 'Duck ate 7 ice cream cones). Present right hand cupped, palm up and move to meet the left hand (scripted to 'ate 3 fewer ice cream cones' or 'ate 3 more ice cream cones'). Move both hands, cupped together, over to Bunny (scripted to 'How many ice cream cones did Bunny eat?').

Figure 6. Example of the Rewording and Gesture Scaffolds for a *Compare 6* Problem.

Compare 6 Word Problem Task

This task consisted of 16 problems: ten *Compare 6* problems that required an addition process for correct solution, and six filler problems that required a subtraction process for correct solution. Five of the 10 *Compare 6* problems were presented using traditional wording, and the other 5 were presented using the rewording scaffold. Each individual problem contained a unique set of numbers drawn from the same pool of 24 nontrivial number facts used in Phase Two. The same quasi-randomization procedure described above was used to select the 10 number facts. Problems containing the same numeral for more than one of the terms were specifically not included, to eliminate ambiguity when scoring (e.g. $6 – 3 = 3$). For all *Compare 6* problems, the augends and addends ranged from 2 to 9, and all sums were ≤ 10. The math facts used in the filler problems were easier than those used in the *Compare* problems, to provide a measure of success for all participants.

In order to accommodate a within subjects design with two gesture scaffold conditions (*With Gesture* and *Without Gesture*), two parallel sets of word problems, each containing five *Compare 6* traditionally-worded problems, five *Compare 6* reworded problems, and six filler problems were created. Because each word problem contained a unique pair of numbers, it was necessary to ensure that differences in solution accuracy across the two gesture conditions did not arise from differences in difficulty of the particular number pairs used within the problem sets. Therefore, two different sets of number facts (e.g. augend + addend) were constructed, as described previously. For a given problem (e.g. the first instance of *Compare 6 - Reworded*), half of the children within each gesture condition received one number fact, and the remaining children in that gesture condition received the other number fact.

During Phase Three, each child participated in two visits, at least one week apart: one visit for the *With Gesture* conditions; the other visit for the *Without Gesture* conditions (See Figure 1). Session order was counterbalanced, with half of the children receiving the two *With Gesture* scaffold conditions in their first visit, and the remaining children receiving the *With Gesture* scaffolds during their second visit. During the first of the two visits all children used the same set of character and object manipulatives, but half of the children heard one set of number facts and the remaining children heard the second set of number facts. During the second visit a different set of character and object manipulatives was used and each child received the number fact set that he/she had not heard in the previous visit. The presentation order for each set of 16 word problems was randomized, and then checked so that no more than 2 problems of any given type were presented sequentially. All participants within each problem set received the same presentation order.

Procedure

The *Compare 6* arithmetic word problem tasks for Phase Three were administered by the investigator during two visits, described above, and lasting approximately one hour each. Participants were tested individually in a specially-equipped van, parked in a quiet location at the child's home or school. After completing 10 of the 16 problems, all participants took a 5 minute break. Other short breaks were taken as needed.

As previously described, each child was seated at the table facing the examiner and two easy *Change* problems were presented for practice. Each word problem was read to the child, and repeated once. The standardized gestures described above accompanied all problems in the *With Gesture* condition. The children were asked to provide an answer to each problem, and then explain their solution to an animal puppet who was 'trying to learn how to do math word problems': e.g. "Tell Rex how you figured out your answer." In this way the child was given the opportunity to explain his or her approach to solving the problem.

As previously described, the examiner placed problem-specific agent and patient figures and objects (i.e. manipulatives) in front of the child for each exemplar. For example, Cow and Monkey figures, along with a set of pumpkins, were arrayed before the child for a *Compare 6* problem where the child was asked to solve 'How many pumpkins did Monkey buy?' (see Figure 4). The purpose of these figure and object manipulatives was to support potential working memory limitations in first graders who must keep track of two numerical quantities in order to solve these word problems successfully (Carr & Hettinger, 2003, p. 39).

The examiner assessed the child's familiarity with the figures and objects while each problem was being set up. All participants were told that they could solve the problems in any way they wanted, including 'using your fingers', using the manipulatives, counting, or 'doing it in your head'. Flexibility in solving the problems via a method preferred by individual children was encouraged during the practice set.

Differential reinforcement was not provided for any of the problems, but all children were given general reinforcement by both the puppet and the examiner for their hard work and explanations.

Coding and Analysis of Problem Accuracy

All sessions were video recorded for later analysis using a Panasonic PV-L857 Palmcorder. Prior to coding and analysis, the analog videos were converted to digital format using Roxio Easy Media Creator 8 Media Import and VideoWave software. Backup audio recordings were made with a Marantz PMD101 audio cassette recorder and an Electro-Voice N/D267AS dynamic cardioid vocal microphone. The principal investigator also completed an online hand-coded form for each child. The recorded data were coded based on verbal responses and solution accuracy (proportion correct). The accuracy data were used to compare performance on the three problem types.

<u>Reliability</u>

An undergraduate research assistant, trained in coding arithmetic word problem tasks, independently recoded a randomly selected subset of the 40 video files (4 from the

With Gesture visits and 4 from the *Without Gesture* visits, representing 20 percent of the video files from Phase Three). Point-to-point agreement between the principal investigator and the independent coder was calculated for two measures: the child's response to each problem (99.22%), and the accuracy of each response (100 %), based on the child's final response.

Analyses

For Phase Three, preplanned analyses included a two-way within subjects analysis of variance (ANOVA). The within subjects factors were rewording with two levels (traditional and reworded) and gesture with two levels (with gesture and without gesture). The dependent variable was proportion correct on each of the scaffolding conditions.

Results

Unless otherwise specified, an alpha level of .05 was used for all statistical tests, p values are two-tailed, and effect sizes (d or η^2) are reported where appropriate.

Solution Accuracy

Table 6 summarizes the overall means and standard deviations for solution accuracy (proportion correct) for the four scaffold conditions. The omnibus test of the effect of rewording and gesture scaffolds on arithmetic word problem solution accuracy of children with low oral language skill yielded a significant main effect of rewording ($F(1, 19) = 9.40$, $p < .01$, $\eta^2 = .33$), indicating that rewording a *Compare 6* word problem facilitated problem comprehension and solution. There was a marginal main effect of gesture ($F(1, 19) = 3.68$, $p < .07$, $\eta^2 = .16$) and the rewording * gesture interaction was not significant ($F(1, 19) = .85$, $p < .37$, $\eta^2 = .04$), indicating that the procedural gesture scaffold was not facilitative, and that gesture condition (With and Without Gesture) did not differentially influence performance in the two wording conditions (Traditional wording and Rewording).

Post hoc pairwise comparisons were conducted to examine differences between the four scaffolding conditions (See Figure 7). Two of the scaffold comparisons yielded significant results: Traditional wording without gesture versus rewording without

Table 6. Means and standard deviations for the four scaffold conditions for all 20 participants in Phase Three.

	All participants (n = 20)	
Measure	M	(SD)
Traditional Wording - Without Gesture	0.10	(0.20)
Reworded - Without Gesture	0.31	(0.38)
Traditional Wording - With Gesture	0.22	(0.32)
Reworded - With Gesture	0.36	(0.40)

gesture ($t(19) = -3.05$, $p < .01$, $d = -0.73$) and traditional wording without gesture versus rewording with gesture ($t(19) = -3.58$, $p < .01$, $d = -0.86$). The comparison of the traditional wording with gesture scaffold to the rewording with gesture scaffold was marginally significant ($t(19) = -2.05$, $p = .05$, $d = -0.39$). Taken together these results confirm that rewording the *Compare 6* word problems was a more facilitative scaffold than adding a procedural gesture.

Post hoc Blocking Factor

Recall that all 20 participants who continued into the current scaffolding phase of the study had achieved an LCOMP score below the 25th percentile, and were considered most at risk for having difficulty solving arithmetic word problems presented in a listening environment. While testing participants, however, I noticed that the children at the lower end of the language skill continuum seemed to be less able to take advantage of the rewording and gesture scaffolds than children with somewhat higher (but still below the 25th percentile) language skill. The group of participants could thus be divided into

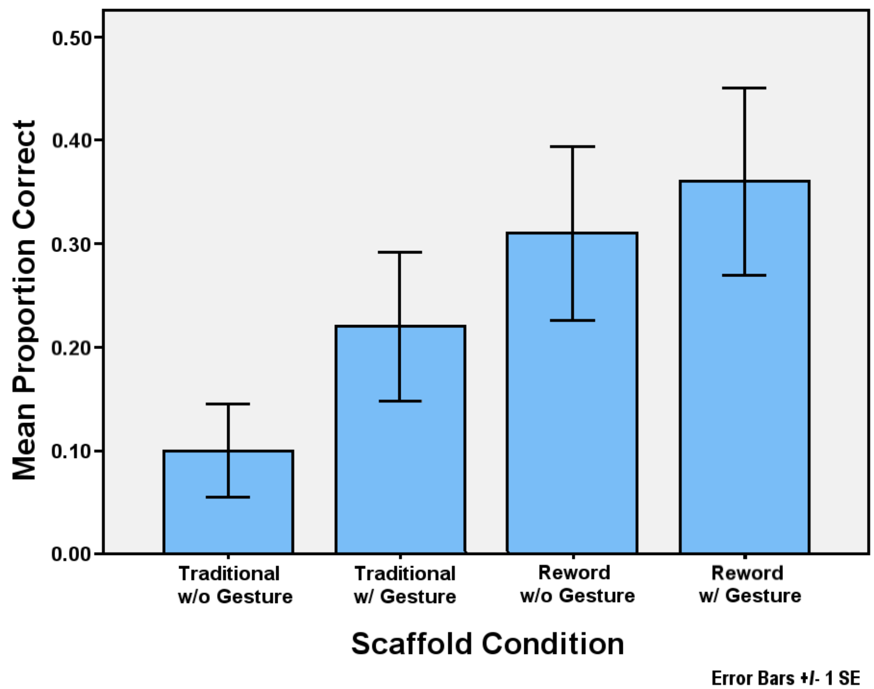

Figure 7. Overall Proportion Correct for the Four Scaffold Conditions in Phase Three.

Note: The range of chance performance is < 0.33 proportion correct.

two groups: 1) a group of children with the diagnosis of language impairment (LI), as determined by the CLRC's language diagnostic criteria (see below), and 2) a group of children with low normal language skill (LN) whose LCOMP scores fell between the LI cutoff and the 25th percentile.

For the purposes of the larger CLRC study, from which all participants in the current study were recruited, language impairment was defined as having an LCOMP *z*-score < -1.14 and a NVIQ composite *z*-score >= -2 (J. B. Tomblin, personal communication, March, 2008). For the purpose of the following analyses, these criteria were used for the LI group. Children in the LN group had LCOMP *z*-scores between -1.14 and -.67, and NVIQ composite *z*-scores >= -2.

To test the hypothesis that children in the LI group are less able to take advantage of the rewording and gesture scaffolds than children in the LN group, a two-way within subjects analysis of variance (ANOVA), with a between subjects post hoc blocking factor (LI, LN) was conducted. The within subjects factors were rewording with two levels (traditional and reworded) and gesture with two levels (with gesture and without gesture). The dependent variable was proportion correct on each of the scaffolding conditions.

Table 7 summarizes the means and standard deviations for solution accuracy (proportion correct) for both groups for the four scaffold conditions. The omnibus test of the effect of rewording and gesture scaffolds on arithmetic word problem solution accuracy yielded a significant main effect of rewording ($F(1, 18) = 15.45$, $p = .001$, $\eta^2 = .46$), indicating again that rewording a *Compare 6* word problem facilitated problem comprehension and solution. (Mauchly's Test of Sphericity was not significant for a one-way repeated measures ANOVA with one factor and four levels within that factor; $p = .46$). As before, the main effect of gesture was marginal ($F(1, 18) = 3.48$, $p = .08$, $\eta^2 = .16$), suggesting that the gesture scaffold was not facilitative. There was a significant group * rewording interaction, indicating that the two groups performed differently on the two rewording conditions. The group * gesture ($F(1, 18) = .01$, $p = .94$, $\eta^2 = .00$), the rewording * gesture ($F(1, 18) = .84$, $p = .37$, $\eta^2 = .05$), and the three-way rewording * gesture * group ($F(1, 18) = .05$, $p = . 38$, $\eta^2 = .00$) interactions were not significant. See Figure 8 for graphs of the interaction analyses. As expected, the test of between-subjects effects was significant ($F(1, 18) = 26.74$, $p < .001$, $\eta^2 = .60$), with the LN group performing better than the LI group. In summary, only rewording facilitated problem solution, but the LN and LI groups performed differently on the problem wording conditions.

Group by Rewording Interaction

To explore the significant group * rewording interaction, tests of simple effects were performed. First I conducted two paired *t*-tests, one for each group, to compare the two reworded conditions (*With Gesture* and *Without Gesture*) to the two traditionally-worded conditions (*With Gesture* and *Without Gesture)*. Using averages of the

Table 7. Means and standard deviations for the four scaffold conditions with a post hoc blocking factor.

	LN (n = 9)		LI (n = 11)	
Measure	M	(SD)	M	(SD)
Traditional Wording / Without Gesture	0.20	(0.26)	0.02	(0.06)
Reworded / Without Gesture	0.58	(0.42)	0.09	(0.10)
Traditional Wording / With Gesture	0.33	(0.35)	0.13	(0.29)
Reworded / With Gesture	0.62	(0.41)	0.15	(0.25)

proportion correct on the reworded (With Gesture and Without Gesture) problems and the traditionally-worded (*With Gesture* and *Without Gesture*) problems, the paired *t*-test for the LN group revealed that the reworded problems were significantly easier than the traditionally-worded problems ($t(8) = 3.38$, $p = .01$, $d = 1.05$). For the LI group, however, the difference between problem types was not significant ($t(10) = -1.34$, $p =$

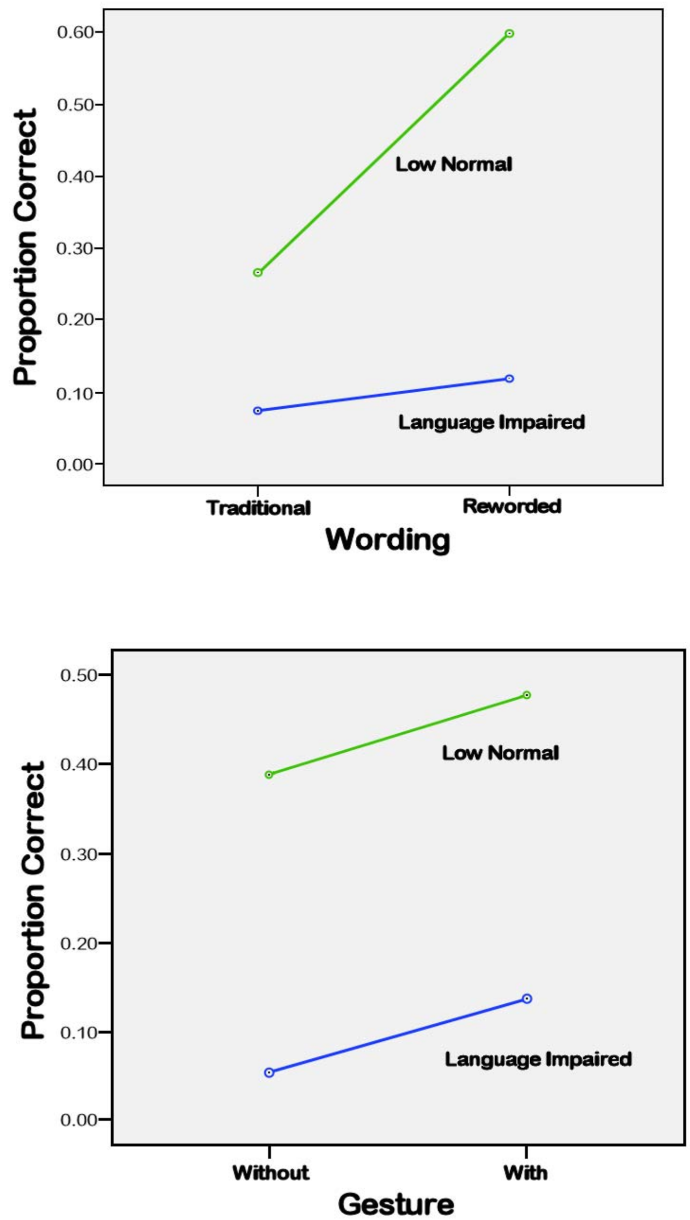

Figure 8. Proportion Correct for Rewording and Gesture Scaffolds.

.21, d = -0.27). Therefore the LN group, but not the LI group, benefited from rewording of the *Compare 6* word problems.

To explore whether there were group differences within each of the two wording conditions, (e.g. within the reworded problem condition, was there a difference in performance between the LN and LI groups), I conducted two independent samples *t*-tests, one for each wording condition. Because Levene's test of equality of variances was significant for both wording conditions (Traditionally-worded: $F(1, 18) = 5.43$, $p < .05$; Reworded: $F(1, 18) = 9.90$, $p = .01$), results for equal variances not assumed are reported. For the reworded condition there was a significant difference between the two groups, with the LN group performing better than the LI group (t (10.26) = 3.38, $p < .01$, d =1.79). For the traditionally-worded condition there was a marginal difference between the two groups ($t(14.13) = 2.08$, $p = .06$, $d = 0.99$). These results provide additional evidence that the LN group was able to take greater advantage of the problem rewording than the LI group.

Because of the significant result for the LN group, pairwise comparisons were conducted to examine differences between the four scaffolding conditions (see Figure 9). For the LN group, three of the scaffold comparisons yielded significant results: Traditional wording without gesture versus rewording without gesture ($t(8) = -2.88$, $p < .05$, $d = -1.11$), traditional wording without gesture versus rewording with gesture ($t(8) = -3.33$, $p = .01$, $d = -1.26$), and traditional wording with gesture versus rewording with gesture ($t(8) = -2.39$, $p < .05$, $d = -0.77$). The comparison of the traditional wording with gesture scaffold to the rewording without gesture scaffold was marginally significant ($t(8) = -2.14$, $p = .07$, $d = -0.64$). For the LN group of first graders, rewording was the strongest scaffold. However the gesture scaffold might have exerted some positive influence, because there was not a significant difference between the traditional wording plus gesture and the rewording without gesture scaffolds.

Two Sets of Number Facts

Recall that two different sets of number facts (e.g. augend + addend) were created. Half of the children received one set of facts during the first visit and the other half received the second set of facts. In order to ensure that differences in solution accuracy across problem types did not arise from differences in difficulty of the particular number pairs used within the two visits, four independent samples *t*-tests were conducted (two gesture conditions; two wording conditions within each gesture condition) to test whether the two sets of number facts differed in difficulty. For all four *t*-tests there were no significant differences between sets of number facts (all *p*s > .44).

Gesture Order Effect

During Visits 1 and 2 of Phase Three, session order was counterbalanced, with half of the children receiving the two *With Gesture* scaffold conditions in their first visit, and the remaining children receiving the *With Gesture* scaffolds during their second visit. In order to test a gesture order effect, I conducted an independent samples *t*-test on the *Without Gesture* data, comparing Visit 1 to Visit 2. There was no significant difference in performance between those children who received *With Gesture* second and those who received *With Gesture* first ($t(38) = -.10$, $p = .922$, $d = -0.03$).

Practice and Learning Effects

To test whether all children in Phase Three improved from Visit 1 to Visit 2 (i.e. a practice effect), I conducted a paired samples *t*-test to compare Visit 1 to Visit 2, using an average of traditional wording and rewording scaffolds for each visit. There was no significant difference in performance from Visit 1 to Visit 2 ($t(19) = -.52$, $p = .61$, $d = -0.08$), indicating that the participants did not improve their solution accuracy as a result of a practice effect.

In addition, because data collection for all three phases of this study extended from May to the end of September, it was possible that children learned on their own or as a result of classroom instruction. To assess whether learning had occurred during this time period, independent of any rewording or gesture scaffold effects, a paired samples *t*-test was conducted, comparing the Phase Two baseline (*Compare 6* problem set) to the Phase Three baseline (*Compare 6 – Tradition Wording Without Gesture* problem set) for

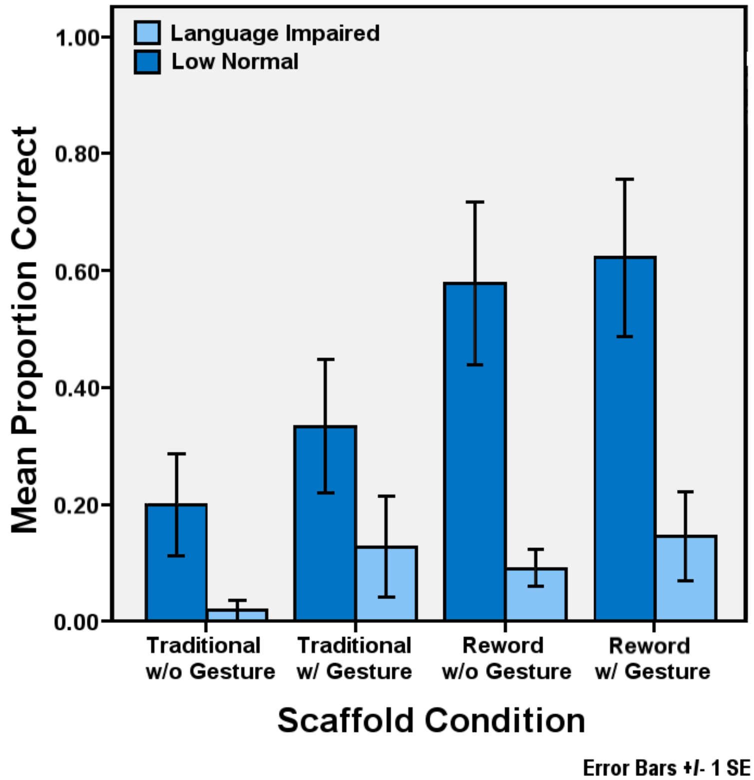

Figure 9. Proportion Correct for the Four Scaffolding Conditions in Phase Three.

Note: The range of chance performance is < 0.33 proportion correct.

each child. There was no significant difference between the *Compare 6* problem sets at the two time points ($t(19) = 0.00$, $p = 1.00$, $d = 0.00$), indicating that children did not learn on their own across the span of the study.

'Question First' Rewording versus Traditional Wording

Recall that the rewording scaffold consisted of several elements: 1) an introductory statement to orient children to the problem statement and characters, 2) *Compare 3* wording ('more than' as opposed to 'fewer than' from the *Compare 6* traditional wording), and 3) 'question first' rewording, to provide up front information about the unknown quantity that the child was expected to solve, as well as a rationale under which the child could view the problem structure. Previous research (Thevenot, et al, 2007) indicated that placing the question first in basic arithmetic word problems facilitated solution accuracy in children with low mathematical skill.

The design of the current study allowed a similar analysis of potential facilitative effects of placing the question first, by comparing the results of Phase Two *Compare 3* problems with Phase Three *Compare 6* reworded but without gesture problems. A paired samples t-test of these two problem types was not significant ($t(19) = -.68$, $p = .51$, $d = -0.83$), counter to Thevenot and colleague's (2007) findings.

Potential Influencing Factors

To test whether there were group differences in LCOMP, NVIQ, visual memory, or auditory memory skill that might have influenced group performance on word problem solution accuracy, a series of independent samples t-tests were conducted. See Table 5 for group means and standard deviations. Although the LI group means were lower than the LN group means on all measures except the Simon task, t-tests of the four dependent variables revealed a significant difference only for the LCOMP ($t(11.9) = 4.21$, $p = .001$, $d = 1.70$), where the LN group performed better than the LI group. Results for 'equal variance not assumed' are reported for LCOMP because Levene's Test for Equality of Variances was significant ($p = .01$). Results for the remaining t-tests were not significant: NVIQ ($p = .36$), Simon-High ($p = .96$) and Digit Span ($p = .32$). These results indicate that, of the measures tested in this study, only a language composite measure differentiated the children with LI from the children in the LN group.

In summary, the children in the LN group responded to the rewording scaffold, as measured by solution accuracy (proportion correct); the children in the LI group did not respond to the rewording scaffold. Oral language skill was the one measure that differentiated the two groups. The gesture scaffold results were marginal for all participants in Phase Three, as measured by solution accuracy. So, in this study, language skill seemed to be the one factor that allowed some of the participants in Phase

Three who were at risk for having difficulty solving arithmetic word problems to take advantage of the rewording scaffold while solving *Compare 6* problems.

Discussion

Response to Scaffolds

It is important to recall that all 20 first graders who participated in Phase Three were in the lowest quartile of language users, and hence were assumed to be at risk for having difficulty solving orally-presented word problems in a classroom setting. The goal of Phase Three was to evaluate potential classroom adaptations by testing whether rewording and gesture scaffolds for orally-presented *Compare 6* arithmetic word problems would facilitate comprehension and solution accuracy in children with low oral language skill. After establishing in Phase One that first graders with poor oral language skill had greater difficulty solving orally-presented arithmetic word problems than their peers, but could still access an underlying knowledge of mathematical set relations in a nonverbal arithmetic task, I hypothesized that a rewording scaffold would mitigate the influence of a linguistic bottleneck on these children's ability to solve arithmetic word problems and a gesture scaffold would help them access their underlying nonverbal knowledge of mathematical set relations. Specifically, I predicted that the rewording scaffold would make the linguistic message more explicit and available to children with low oral language skill, and would increase solution accuracy. In addition, I predicted that a procedural gesture scaffold would help these children access the underlying nonverbal problem structure, and would increase solution accuracy. Evidence of facilitative scaffolding would be seen in increased solution accuracy in the reworded and/or the gesture conditions in comparison to the baseline traditional wording condition. Any difference in the relative influence of the gesture and rewording scaffolds would be evident in a comparison of the reworded-alone condition to the reworded with gesture condition. If the gesture and rewording scaffolds were not facilitative, I predicted that solution accuracy would not differ significantly across the four conditions. On the other hand, facilitative scaffolds were predicted to result in a significant increase in the number of problems solved accurately.

These predictions were only partially supported. The children in the LN group responded to the rewording scaffold, as measured by solution accuracy (proportion correct); the children in the LI group did not respond to the rewording scaffold. The procedural gesture scaffold was only marginally facilitative. So, in a group of children who were in the lowest quartile of language users, the rewording scaffold was more powerful than the gesture scaffold. Possible reasons why children with low oral language skill responded more to a rewording scaffold than to a gesture scaffold will be discussed in a later section.

Group Differences in Response to Scaffolds

Given my observation that the children at the lower end of the language skill continuum seemed less able to take advantage of the rewording and gesture scaffolds than children with somewhat higher language skill, I compared the performance of the children who had been diagnosed with language impairment (LI) to the performance of the children whose language fell within normal limits, but was still below the 25th percentile (LN). The two groups responded differently to the scaffolding conditions.

<u>Low Normal Language Skill Group</u>

The children with low normal language skill benefited from the rewording scaffold, as measured by an increase in solution accuracy. Applying rewording alone and applying both rewording and gesture scaffolds resulted in increased accuracy beyond the baseline traditional wording condition. Rewording plus gesture also contributed to a significant increase in solution accuracy in comparison to the traditional wording plus gesture condition, indicating that a gesture scaffold did not interfere with the effects of rewording. In fact, gesture might have provided some benefit of its own (an indication of the marginal results for a gesture scaffold), because the comparison of traditional wording plus gesture to the condition with both scaffolds was not significant. It is also possible that this comparison indicates that adding a gesture to the rewording condition interfered with the effect of rewording, but given the other strong indications of a positive effect of rewording, and the marginal effect of gesture, it seems more likely that gesture did have some positive influence on the LN group.

Language-Impaired Group

The LI group of first graders presented a different pattern of response to the rewording and gesture scaffolds. Neither scaffold was facilitative for this group. Clinical implications for these results will be discussed in more detail in the final section of this chapter, but here we have some indication that these two groups might respond differently to rewording and gesture scaffolds in a classroom setting.

Performance above Chance

To confirm whether first graders with low oral language skill actually performed above chance in any of the scaffolding conditions, I computed the range of chance for a set of five word problems and compared this range to the mean number of problems solved correctly for both groups under all four scaffolding conditions. Computing the range of chance performance for an arithmetic word problem task is not as straight forward as computing the range for a closed set task. Given the evidence from prior studies as well as from the current study, I assumed that first graders could potentially provide approximately 8 different responses to a *Compare 6* word problem, resulting in above chance performance if they solved more than 1.67 problems correctly in a set of five problems. Using this cut-off point, the two groups combined performed above chance only on the rewording plus gesture condition. The LN group performed above chance on both rewording conditions (rewording without gesture, and rewording plus gesture), whereas the LI group did not perform above the level of chance on any of the four scaffolding conditions. This evidence confirms the facilitative effect of the rewording scaffold on the ability of first graders with low oral language skill to solve arithmetic word problems, especially if they do not have a diagnosis of language impairment.

Other Implications of the Rewording Scaffold Results

Part of the rewording scaffold for the *Compare 6* problems in Phase Three included substituting the *Compare 3* 'more than' wording for the traditional *Compare 6* 'less than' wording. In addition, *Compare 3* wording eliminated the potential linguistic roadblock of pronoun anaphora that was present in the traditional *Compare 6* wording. Both of these factors made the reworded *Compare 6* problems more similar linguistically to the traditionally worded *Compare 3* problem type. The rewording for *Compare 6* problems, however also included: 1) an introductory sentence, orienting children to the specific problem situation and characters, 2) placing verbs in optional settings to make the problem wording more explicit, and 3) placing the question early in the problem statement, to provide up front information about the unknown quantity, as well as a rationale under which to view the problem structure.

In their 2007 study, Thevenot and colleagues demonstrated that the use of 'question first' rewording alone facilitated arithmetic word problem solution in fourth graders, especially those with low math skills. In this study, however, the test of 'question first' rewording (comparing the results of Phase Two *Compare 3* problems with Phase Three reworded *Compare 6* problems) was not significant. Recall, however, that Thevenot and colleagues specifically used a 'question first' rewording, whereas I used 'question first' in addition to an introductory sentence and optional verbs. Because I did not find a significant difference in the PLU group between the Phase Two *Compare 3* problems and the Phase Three reworded *Compare 6* problems, it appears that the significant effect of rewording was driven by the change from *Compare 6* wording to

Compare 3 wording. Given the difference in design and results between this study and the Thevenot et al (2007) study, the components of problem rewording should be explored in future research in order to pinpoint more specifically the reason(s) for a rewording effect in Phase Three of this study.

Marginal Effect of the Gesture Scaffold

My prediction that a gesture scaffold would facilitate solution accuracy in children with low oral language skill was not supported in this study. There are a number of possible reasons for this result. First, the gesture I used for the *Compare 6* problems in Phase Three was a procedural gesture, consisting of several gestures indicating that two sets needed to be combined (i.e. an addition procedure) in order to solve this problem type accurately. Much of the previous work on the effects of teachers' gestures on children's learning has focused on the use of conceptual gestures, such as the use of a single gesture to represent the concept of equivalence across the equal sign in a mathematical equation (e.g. Church, Ayman-Nolley, and Vasich, 2007; Ellis Weismer and Hesketh, 1993; McNeil, Alibali, and Evans, 2000). Therefore, I hypothesize that my procedural gesture carried a higher cognitive load, making it less accessible to children in the lowest quartile of language skill. My limited pilot data, however, suggested that children with higher language skill were able to access a procedural gesture and integrate it with the problem wording.

This hypothesis might be supported by Mainela-Arnold and colleagues' (2006) proposal that it was easier for children with language impairment to express a perceptually-based 'external' explanation for their conservation judgments in gesture than it was for them to express a higher level 'internal' explanation in gesture. They also suggested that actually producing more external gestures might then have caused the children with SLI to focus more on superficial perceptual features rather than the higher-level cognitive features of the conservation task. In a similar fashion, in the current study the children's manipulation of the characters and objects might have focused their attention on quantities and superficial features rather than on the relations between those quantities (i.e. the children couldn't inhibit their attention to objects in order to attend to and integrate my procedural gesture).

A second reason why there was not a significant effect of gesture might have been because children in the lowest quartile of oral language skill had difficulty integrating multiple modalities (i.e. verbal and gesture). Cain and Oakhill (1999) found that children who were poor reading comprehenders had difficulty constructing accurate mental models from the text. The results of the current study suggested that the children with language impairment also had difficulty constructing accurate mental models of the word problems that they heard. Theoretically we would predict that these children would benefit from the addition of a gesture scaffold that would help them access a stronger nonverbal representation of the problem structure and the solution procedure. However Cain and Oakhill also concluded that their poor reading comprehenders had difficulty integrating new information within an existing representation (also see Cain, Oakhill, Barnes, & Bryant, 2001; Kintsch, 1998). Therefore, in the current study it was possible that adding gestural information to an already heavy linguistic load overwhelmed the children with low oral language skill.

I did see some anecdotal evidence that *individual* children responded to the procedural gesture scaffold. For at least some of these children, however, although the gesture did seem to alter their mental model of the word problem structure, this altered representation still resulted in an incorrect solution strategy and therefore affected solution accuracy.

A third reason for the gesture scaffold findings in this study might have been related to the fact that I intentionally did not provide problem-specific feedback during any of the three phases. Because this was not an intervention study, but rather was an experimental study designed to test the effects of rewording and gesture scaffolds, explicit feedback on children's performance on any of the conditions would have confounded the results. Therefore, in order for children in this study to benefit from a gesture scaffold, they would have had to process the gesture implicitly. It is possible that children with language impairment require explicit instruction in a gesture scaffold in order to access the information that it contains and integrate this information with the verbal message. I did see some anecdotal evidence that, because no explicit feedback was provided, some of the children with lower language skill got 'stuck' in an incorrect strategy. This observation might partially account for the fact that children in the LI group only got .1 out of 5 problems correct on average in the traditional wording baseline condition.

A fourth reason why there was not a significant effect of gesture might have been because I chose to test gesture in a higher level connected language task. Previous studies that found a positive effect of gesture on language comprehension in children with language impairment used word learning and spoken word recognition paradigms rather than connected language (Ellis Weismer and Hesketh, 1993; McGregor and Rost, 2007).

And finally, Goldin-Meadow and colleagues (1999) reported that third and fourth grade children reiterated incorrect information that they had seen presented in gesture mode, despite the fact that they heard the correct information in the teacher's accompanying speech. It is possible that my procedural gesture misled the children in the current study. They did respond to rewording though, so if my gesture was misleading, the children might have responded with lower accuracy to the rewording + gesture condition.

Given my limited pilot data suggesting that kindergarten and first grade children with higher oral language skills do benefit from rewording and gesture scaffolds, and given that the GLU group in Phase Two also struggled with *Compare 6* word problems, it would have been interesting to test whether the children in the GLU group who did not perform above chance on this problem type responded to the gesture scaffolding conditions. This group was not followed into Phase Three for two reasons: 1) the focus of this study was to investigate the influence of rewording and gesture scaffolds in a group of children with low oral language skill, and 2) seeing another group of children for two more visits would have extended data collection far enough into their second grade year to raise concerns that their language skills were developing and/or that they were developing mathematical schemata due to classroom instruction. Potential development in either of these areas would have obscured any interpretation of significant results.

Other Potential Influencing Factors

Schema theory is a component of a number of models of learning and problem solving (Marshall, 1995). Within the realm of mathematical word problems, schema theory proposes that repeated exposure to a given problem type eventually results in the construction of a productive and accurate schema. Given the existing evidence that children with language impairment have difficulty with abstraction and generalization in verbal tasks (Masterson, Evans & Aloia, 1993; Nippold, 1994), the question remains whether any of these factors also influenced the LI group's relatively poor performance on the word problem tasks in this study. As mentioned in the previous section, there is evidence that children with language impairment have difficulty inhibiting irrelevant information in a variety of tasks (Norbury, 2005; Samelson, McMurray, Lee, & Tomblin, 2007). There was anecdotal evidence in this study that the children who had been diagnosed with language impairment tended to focus more on the characters and objects that were physically in front of them than on the verbal problem statements. Under schema theory, the characters and objects in a word problem are the least critical pieces of information in the word problem schema, in contrast to the quantities and the relationships between those quantities (Marshall, 1995). It is possible that the children with language impairment in this study might have had difficulty inhibiting their attention to the manipulatives in front of them and focusing their attention on the verbal and/or gestural cues because they were still at a more immature stage in the developmental progression of allocation of attention (McNeil, et al, 2000, p. 145). Use of manipulatives and characters might have encouraged the children to provide given number errors (assigning given numbers to each character), rather than to add the two quantities, despite my procedural gesture that encouraged them to join the quantities.

Considering the potential factors of difficulty with attentional processes, difficulty with the inhibition of irrelevant information, and difficulty with abstracting and generalizing, we would then expect children with language impairment to have difficulty building an accurate mental model of an arithmetic word problem and then abstracting an accurate schema from repeated exposure to a given type of word problem.

Zone of Proximal Development

A key component of Vygotsky's theory of learning and child development is the concept of 'zone of proximal development' (ZPD). Vygotsky defined a child's ZPD as the distance between the child's actual level of independent problem solving skill and the level a child can attain with the assistance of a capable adult or peer (1978). One of the challenges in this study was the task of deciding which arithmetic word problem type would continue into Phase Three. Prior research (Riley & Greeno, 1988) indicated that between the end of first grade and the end of second grade children continue to develop a number of word problem schemata, as evidenced by an increase in their ability to solve a greater variety of word problems. It was necessary to select a problem type that was difficult enough to allow room for scaffolding effects and that was not likely to be part of a first grader's arsenal of word problem schemata. On the other hand the problem type could not be so difficult that it would be beyond the ZPD for first graders with low oral language skill. Given the relatively short period of time within which children with good language skill develop basic arithmetic word problem schemata, the challenge was to avoid both floor and ceiling effects. As a result, my choice of *Compare 6* word problems for Phase Three probably placed the rewording scaffold into the ZPD of the LN group, but not into the ZPD of the children with language impairment. It is also possible in this study that the gesture scaffold was outside the ZPD for both groups of children with low oral language skill.

Through the years, the term ‘scaffold’ has become linked to Vygotsky’s learning theories (Stone, 1998). It is important to note that, although I have used the terms ‘scaffold’ and ‘scaffolding’ liberally throughout this paper, the fact that I did not provide explicit feedback and guidance while the children were solving the word problems means that my interpretation and use of these terms for the purpose of this study does not conform with a truly Vygotskyan definition.

Clinical Implications

The goal of Phase Three was to evaluate potential classroom adaptations by investigating the effects of two types of scaffolds on arithmetic word problem solving ability in those children with low oral language skill who had difficulty solving orally-presented word problems. In this section, rewording and gesture adaptations are discussed, specifically from the viewpoint of a speech and language pathologist whose task is to serve at-risk children directly while also providing support for the classroom educator.

Under the Response to Intervention (RTI) initiative, speech and language pathologists (SLPs) are often expected to help monitor the status of all at-risk students, not only those with a speech and language IEP. Because the RTI model focuses on providing high quality, evidenced-based instruction within the classroom environment, followed by systematically applied tiers of more specialized instruction for at-risk children, SLPs now provide services in both traditional intervention and classroom-based intervention models (Justice, 2006; Ukrainetz, 2006).

Because the current study demonstrated that children with a diagnosis of language impairment and children with low normal oral language skill responded differently to rewording and gesture scaffolds, and because under an RTI model both of these groups of children would receive services from an SLP, we need to consider the specific types of service each of these groups might require. The results of the current study suggested that children with low oral language skill, particularly if they have a diagnosis of language impairment, do not benefit from implicitly presented rewording and gesture scaffolds. Therefore, in a classroom listening environment, we would not expect these children to comprehend more difficult language content without explicit intervention directed to their language level. Educators and SLPs need to be especially sensitive to the language and conceptual level of curriculum materials, as well as to the language level used by the classroom teacher when he/she presents new information. This would be particularly true for math instruction, because instruction in this content area is often less individualized than in reading instruction.

Within the area of math instruction, simply rewording the curriculum content might not be sufficient for children with language impairment. They might require more explicit instruction beyond the rewording of arithmetic word problems. In addition, this study provided evidence that children in the lowest quartile of language skill might not benefit from information presented implicitly in the teacher's gestures. Awareness by both the classroom teacher and the SLP of these potential roadblocks to comprehension in children with low oral language skill will help both groups of professionals design more targeted and therefore more effective interventions.

CHAPTER VII

SUMMARY AND FUTURE DIRECTIONS

Summary

This study consisted of three phases. Phase One supported previous research (Jordan, Levine, & Huttenlocher, 1995) and established that first graders with poor oral language skill had more difficulty solving orally-presented arithmetic word problems than children with good oral language skill. Children with poor oral language skill, however, were able to solve nonverbal arithmetic problems as well as their peers with good oral language skill. In Phase Two I examined the ability of first graders with good and poor oral language skill to solve *Compare 3, Compare 5,* and *Compare 6* arithmetic word problems. All three problem types were difficult for most of the participants. Based on the rationale discussed in Chapter VI, I selected *Compare 6* word problems for Phase Three, where I examined the effects of two scaffolds: 1) a rewording scaffold that was hypothesized to mitigate the influence of a linguistic bottleneck on children's ability to solve verbal arithmetic word problems, and 2) a gesture scaffold that was hypothesized to access and support the underlying knowledge of mathematical set relations of children with low language skills. In a group of children who were in the lowest quartile of language users, the rewording scaffold was more powerful than the gesture scaffold. Furthermore, the children with low normal language skill responded to the rewording scaffold, as measured by solution accuracy (proportion correct), but the children with language impairment did not respond to the rewording scaffold. The procedural gesture scaffold was only marginally facilitative, and possible reasons for this result were discussed in the previous chapter.

The main message of this study seems to be that language skill matters when children are asked to solve arithmetic word problems. Furthermore, particularly for children with language impairment, despite their underlying knowledge of mathematical set relations, a linguistic bottleneck still impeded their ability to solve reworded problems. Despite the finding by Jordan and colleagues (1995) and the finding in Phase One of the current study that children with low oral language skill were able to access nonverbal knowledge of set relations, it is possible that other cognitive general factors such as attentional skill, the ability to form abstractions, and the ability to integrate multiple modalities influenced the findings of the current study. In addition, language-based factors such as the ability to form inferences and the ability to abstract a verbal word problem schema possibly influenced solution accuracy. It remains to be tested whether explicit gesture or word problem strategy instruction will facilitate solution accuracy for arithmetic word problems in children with low oral language skill.

The current study extended prior research on children's ability to solve arithmetic word problems by specifically analyzing word problem skill in first graders whose oral language skills were in the lowest quartile of language functioning. Moreover, this study identified differences in problem solving skill between a group of children who had been identified as language impaired and a group of children whose oral language skills were in the low normal range. The implications of these differences in word problem solving skill were discussed with regard to classroom curriculum and instruction adaptations, specifically from the viewpoint of a speech and language pathologist whose task is to directly serve at-risk children while also providing support for the classroom educator. In summary, classroom and curriculum adaptations may be effective for children with low oral language skill, but special consideration must be given to children who have been diagnosed with language impairment.

Future Directions

The results of the current study, along with the implications for children with a diagnosis of language impairment, lead to a number of directions for future research:

1) First, anecdotal evidence from this study suggested that some of the children responded to the procedural gesture scaffold. Although the gesture seemed to alter their mental model of the word problem structure, this altered representation sometimes resulted in a different but still incorrect solution strategy. Future research should include an analysis of error patterns and solution strategies for all four scaffolding conditions, in order to assess the influence of rewording and/or gesture on the types of errors and the solution strategies children choose to use.

2) It would be interesting to teach my procedural gesture explicitly to a group of first graders with language impairment. Broaders, Cook, Mitchell, and Goldin-Meadow (2007), taught third and fourth graders a gesture and then told the students to use that gesture while explaining how they solved mathematical equivalence problems. They found that students who were told to use the gesture added new and correct solution strategies, even though they were previously unable to solve the problems. It remains to be seen whether this approach would be facilitative in a word problem solving task in children with language impairment. Based on Cook and Goldin-Meadow's (2006) finding that being explicitly taught a gesture, and then using that gesture, helps shape a child's mental representation of the problem over a longer period of time, it would also be interesting to see if children with language impairment are able to retain and generalize this new mental representation.

3) Future research should also extend the current study to children with higher oral language skill. The results from this extension would verify if my procedural gesture is facilitative, at least for good language users.

4) Perhaps a conceptual gesture would be more facilitative than my procedural gesture. In order to explore this possibility, a future study would have to determine the ZPD for a different type of word problem that would better accommodate a conceptual gesture, as well as determine the appropriate participant age range and language level.

5) Now that I have established that *Compare 3, Compare 5,* and *Compare 6* arithmetic word problems are difficult for children with language impairment and that these children don't make use of scaffolds implicitly, it is important to determine if differences exist between children with language impairment and children with low normal language skill in their ability to make use of context, to focus attention on relevant information, to make inferences, abstractions and generalizations, in the context of different types of arithmetic word problems.

6) Rewording the *Compare 6* problems was facilitative for children with low normal oral language skill, but not for children with language impairment. Both groups, however, performed with a higher level of accuracy on the filler problems than on the *Compare 6* problems. These fillers were based on 'action' schemata requiring a separation of a smaller set from a larger set, and contained verbs of change of possession (e.g. 'Batman stole three of Spiderman's Easter eggs.'). In contrast, the *Compare 3, Compare 5,* and *Compare 6* problem types were based on 'static' schemata that required a comparison of two sets, where the semantic structure of the problem focused on the lexical items 'more than' or 'less than'. Future research should investigate whether verb semantics is a significant factor in the comprehension of arithmetic word problems in children with language impairment.

7) And finally, additional work is needed to explore the relationship between mathematical cognition, cognitive general abilities, and language ability in three groups: children with good oral language skill, children with low normal language skill, and children with language impairment. We should then use the results of these explorations to refine and test scaffolds and interventions that will improve the ability of children with low oral language skill to comprehend and integrate linguistic and gestural information in a classroom setting.

APPENDIX A

ARITHETIC WORD PROBLEM TYPES

Problem Type	Example
Combine 1: Sum unknown	Joe has 2 marbles. Tom has 6 marbles. How many marbles do they have altogether?
Combine 2: Subset unknown	Joe has 2 marbles. Tom has some marbles. They have 6 marbles altogether. How many marbles does Tom have?
Combine 5: Sum is known. Subset unknown; Decrease	Joe and Tom have 6 marbles altogether. Joe has 2 marbles. How many does Tom have?
Change 1: Result unknown; Increase	Joe had 2 marbles. Then Tom gave him 6 marbles. How many marbles does Joe have now?
Change 2: Result unknown; Decrease	Joe had 6 marbles. Then he gave Tom 2 marbles. How many marbles does Joe have now?
Change 3: Change unknown; Increase	Joe had 2 marbles. Then Tom gave him some marbles. Now Joe has 6 marbles. How many marbles did Tom give him?
Change 4: Change unknown; Decrease	Joe had 6 marbles. Then he gave Tom some marbles. Now Joe has 2 marbles. How many marbles did he give Tom?
Change 5: Start unknown; Increase	Joe had some marbles. Then Tom gave him 2 marbles. Now Joe has 6 marbles. How many marbles did Joe have in the beginning?
Change 6: Start unknown; Decrease	Joe had some marbles. Then he gave Tom 2 marbles. Now Joe has 6 marbles. How many marbles did Joe have in the beginning?
Compare 1: Difference unknown; more than	Joe has 2 marbles. Tom has 6 marbles. How many more marbles does Tom have than Joe?
Compare 2: Difference unknown; fewer than	Joe has 6 marbles. Tom has 2 marbles. How many fewer marbles does Tom have than Joe?
Compare 3: Compared quantity unknown; more than (Add)	Joe has 2 marbles. Tom has 6 more marbles than Joe. How many marbles does Tom have?
Compare 4: Compared quantity unknown; fewer than (Subtract)	Joe has 6 marbles. Tom has 2 fewer marbles than Joe. How many marbles does Tom have?
Compare 5: Compared quantity unknown; more than (Subtract)	Joe has 6 marbles. He has 2 more marbles than Tom. How many marbles does Tom have?
Compare 6: Compared quantity unknown; fewer than (Add)	Joe has 2 marbles. He has 6 fewer marbles than Tom. How many marbles does Tom have?

Adapted from Okamoto, 1996, p. 411

APPENDIX B

LANGUAGE MEASURES INCLUDED IN THE LANGUAGE COMPOSITE SCORE (LCOMP)

Four to Six Year-Old Protocol:

Test of Language Development-2 Primary (TOLD-2:P; Newcomer & Hammill, 1988)

Subtests:

- Picture Vocabulary
- Oral Vocabulary
- Grammatical Understanding
- Grammatical Completion
- Sentence Imitation

Seven to Eight Year-Old Protocol:

Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981)

Clinical Evaluation of Language Fundamentals – Third Edition (CELF-III; Semel, Wiig, & Secord, 1995) Subtests:

- Concepts and Following Directions
- Sentence Structure
- Word Structure
- Recalling Sentences

APPENDIX C

PHASE ONE: VERBAL AND NONVERBAL ARITHMETIC PROBLEM SETS

Nonverbal Arithmetic Problems:

1. $1 + 3 = 4$
2. $5 – 3 = 2$
3. $7 – 4 = 3$
4. $4 + 2 - 1 = 5$
5. $4 – 1 = 3$
6. $5 + 4 = 9$
7. $2 + 1 = 3$
8. $6 + 4 – 2 = 8$
9. $9 – 5 = 4$
10. $3 + 4 = 7$
11. $3 – 2 = 1$
12. $2 + 6 – 3 = 5$
13. $2 + 3 = 5$
14. $8 – 5 + 3 = 6$

Protocol for Nonverbal Arithmetic Problem Task:

Use no verbal number or process (e.g. ‘add’, ‘subtract’) labels during the task, and do not explicitly ask the child to provide a verbal response. However spontaneous verbal responses from the child are fine and the child can solve the problem in any way that he chooses.

Setup: Place the pile of disks at the right edge of the mat, between you and the child. Place the opaque cover at the left edge of the mat, closer to you.

We're going to play a number game. You can figure out the answer any way that you want. You can use your fingers, you can use these plastic disks, you can count, or you can do it in your head. Take your time. You don't have to answer fast.

1 repetition is allowed, if requested by the child, or if the child is not attending.

Demonstration: Place 3 disks on the mat in a horizontal line, in front of you and slightly to your left. *See this?* (Pull your hand back and pause slightly) *Now watch what I do.* Cover your disks with the box. Place 3 disks on the mat in a horizontal line, in front of the child and slightly to his right. Lift your cover. *See?* .... *Yours is just like mine.* Repeat the placement with 4 disks. *See this?* (Pull your hand back and pause slightly) *Now watch what I do.* Cover your disks with the box. *Now make yours just like mine.* Lift your cover. *Yes, yours is just like mine. Let's try another one.*

Addition Problems: (Practice: 1+1)

Place a horizontal line of disks corresponding to the augend on your mat. *See this?* (Pull your hand back and pause slightly) *Now watch what I do next.* Cover the disks. Place a second line of disks corresponding to the addend to the right side of the box, and in view of the child. Slide the disks under the cover through the opening, one at a time. *Make yours just like mine.* Lift your cover. *Yes, yours is just like mine.*

Subtraction Problems: (Practice 2-1)

Place a horizontal line of disks corresponding to the large set on your mat. *See this?* (Pull your hand back and pause slightly) *Now watch what I do next.* Cover the disks. Slide the number of disks corresponding to the subtrahend out from under the cover, one by one. *Now make yours just like mine.* Lift your cover. *Yes, yours is just like mine.*

We are going to keep playing this game. Some of the problems will be easy and some will be harder. Don't worry if you don't get them all right. Just do your best. You *have to try to make yours look just like mine. Remember you can figure out the answers any way that you want.*

<u>Verbal Arithmetic Problem Set and Protocol</u>

Setup: Place the pile of disks at the right edge of the mat, between you and the child. *We're going to do some word problems. You can figure out the answer any way that you want. You can use your fingers, you can use these plastic disks, you can count, or you can do it in your head. Take your time. You don't have to answer fast. Let's do two practice problems.*

Read with natural inflection. One repetition is allowed, if requested by the child, or if the child is not attending.

Practice – Addition: 1+1 Cookie Monster and Linda are eating lunch. Cookie Monster has one cookie. Linda gives him one more cookie. How many cookies does Cookie Monster have altogether?

Practice – Subtraction: 2-1 Susan and Oscar are playing with their marbles. Susan has 2 marbles. Oscar takes away 1 of her marbles. How many marbles does Susan have left?

1. 1+3 Lydia and John like snack time. Lydia has one cracker. John gives her 3 more crackers. How many crackers does Lydia have altogether?
2. 5-3 Raymond and Lisa are counting bouncy balls. Raymond has 5 balls. Lisa takes away 3 of his balls. How many bouncy balls does Raymond have left?
3. 7-4 Elmo and Zoe are eating potato chips. Elmo has 7 chips. Zoe takes away 4 of his chips. How many chips does Elmo have left?

4. 4+2-1	Dora and Boots go to the circus. Dora buys 4 balloons. Boots gives her 2 more balloons. Then he takes away one of her balloons. How many balloons does Dora have left?
5. 4-1	Mark and Emily have pencils at school. Mark has 4 pencils. He gives one pencil to Emily. Now how many pencils does Mark have?
6. 5+4	Mark and Sarah go to the library. Mark finds 5 books that he wants to read. Sarah gives him 4 more books. How many books does Mark have altogether?
7. 2+1	Andy and Julia like popsicles. Andy has 2 popsicles. Julia gives him one more popsicle. How many popsicles does Andy have altogether?
8. 6+4-2	Bert and Linda are catching frogs. Bert catches 6 frogs. Linda gives him 4 more frogs. Then she takes one of the frogs back. How many frogs does Bert have now?
9. 9-5	Kim and Alex are collecting pennies. Kim has 9 pennies. Then she gives 5 pennies to Alex. How many pennies does she have left?
10. 3+4	Sally and John like pet fish. Sally has 3 fish. John buys her 4 more fish. How many pet fish does Sally have now?
11. 3-2	Big Bird and Julie found some squishy worms. Big Bird has 3 worms. Julie takes away 2 of his worms. How many squishy worms does Big Bird have left?
12. 2+6-3	Larry and Annie are eating jelly beans. Larry has 2 jelly beans. Annie gives him 6 more. Then she takes 3 jelly beans back. How many jelly beans does Larry have left?

13. 2+3	Diego has a birthday party. He gets 2 presents. Then Dora gives him 3 more presents. How many presents does Diego have altogether?
14. 8-5+3	SpongeBob and Madeline collect shells. SpongeBob has 8 shells. Madeline takes away 5 of his shells. Then she gives him 3 shells back. How many shells does SpongeBob have now?

APPENDIX D

PHASE TWO: WORD PROBLEM SETS

Set 1

Setup: Ricky Raccoon, lying on left corner of table, facing child. For each problem, place the first-mentioned character to your left and the second-mentioned character to your right. Place the objects between and slightly closer to the child. Spread each set of objects slightly, so that none are touching. As you place the characters and objects, chat with the child about each, to make sure the child is familiar with character and object names. Set out the materials for the first practice problem. Other character names can be substituted if the child is more familiar with another name (e.g. bunnies/rabbits)

I brought my friend Ricky with me today. He is learning how to do math word problems in school. But he needs a friend to teach him how to do these problems. I thought that you would be able to help him. When we are all done with the problems, you can give Ricky a couple of practice problems, to see if he was listening.

I will read each problem 2 times. Wait to tell me your answer until I have finished reading the problem 2 times. Then I want you to help Ricky, and tell him how to do the problem.

You can figure out the answer to these problems any way that you want. You can use your fingers, you can use the objects that will be out here, you can count, or you can do it in your head. Take your time. You don't have to answer fast.

Let's do 2 practice problems. (No gesture. Read each problem 2 times, with natural inflection.)

Practice #1. Alligator ate 3 ice cream cones. Tiger ate 1 ice cream cone. How many ice cream cones did they eat altogether? (3+1=4)

(Explain problem if incorrect) *Ricky says: How did you figure out that one?*

Practice #2. Mouse and Lady Bug are catching fish. Mouse caught 4 fish. Lady Bug took away one of mouse's fish. How many fish does Mouse have left? (4-1=3)

(Explain problem if incorrect) *Ricky says: How did you figure out that one?*

Let's do some more problems. Some of the problems will be easy and some will be harder. Don't worry if you don't get them all right. Just do your best. Remember you can figure out the answers any way that you want.

While reading the word problems, wait for child to move manipulatives before continuing to read the problem. Remind the child not to answer until you have read the problem 2 times.

1. Bunny caught 5 frogs. He caught 2 fewer frogs than Tiger. How many frogs did Tiger catch? (5+2=7) *Compare 6*

Ricky says: *How did you figure out that one?****

2. Mario has 9 legos. He has 7 more legos than Tweety. How many legos does Tweety have? (9-7=2) *Compare 5*

Ricky says: *How did you figure out that one?****

3. Curious George has 8 pet turtles. He has 2 fewer turtles than Elephant. How many turtles does Elephant have? (8+2=10) *Compare 6*

Ricky says: *How did you figure out that one?****

F. Donald Duck has 6 Easter eggs. Tigger steals 1 Easter egg from Donald. How many eggs does Donald Duck have left? (6-1=5)

Ricky says: *How did you figure out that one?****

4. Pooh ate 6 carrots. Minnie ate 3 more carrots than Pooh. How many carrots did Minnie eat? (6+3=9) *Compare 3*

Ricky says: *How did you figure out that one?****

5. Lion has 6 fish. He has 4 fewer fish than Panda. How many fish does Panda have? (6+4=10) *Compare 6*

Ricky says: *How did you figure out that one?****

6. Big Bird has 10 rings. He has 2 more rings than Oscar. How many rings does Oscar have? (10-2=8) *Compare 5*

Ricky says: *How did you figure out that one?****

7. Thing One has 5 rabbits. Lambie has 4 more rabbits than Thing One. How many rabbits does Lambie have? (5+4=9) *Compare 3*

Ricky says: *How did you figure out that one?****

F. The girl construction worker has 4 pipes. She gives 1 pipe to the boy worker. How many pipes does the girl have left? (4-1=3)

Ricky says: *How did you figure out that one?****

8. Brown puppy has 3 bones. Spot has 4 more bones than Brown puppy. How many bones does Spot have? (3+4=7) *Compare 3*

Ricky says: *How did you figure out that one?****

9. Duck ate 7 ice cream cones. He ate 3 fewer ice cream cones than Bunny. How many ice cream cones did Bunny eat? (7+3=10) *Compare 6*

Ricky says: *How did you figure out that one?****

Break…

10. Spiderman caught 4 bees. Batman caught 2 more bees than Spiderman. How many bees did Batman catch? (4+2=6) *Compare 3*

Ricky says: *How did you figure out that one?****

11. Pig has 7 baseballs. He caught has 3 more baseballs than Monkey. How many baseballs does Monkey have? (7-3=4) *Compare 5*

Ricky says: *How did you figure out that one?****

F. Goofy bought 3 baby chicks. Garfield took 2 of Goofy's chicks. How many chicks does Goofy have left? (3-2=1)

Ricky says: *How did you figure out that one?****

12. Warthog counted 6 smiley faces. Hippo counted 2 more smiley faces than Warthog. How many smiley faces did Hippo count? (6+2=8) *Compare 3*

Ricky says: *How did you figure out that one?****

13. Elephant picked 8 flowers. He picked 3 more flowers than Shrek. How many flowers did Shrek pick? (8-3=5) *Compare 5*

Ricky says: *How did you figure out that one?****

14. Mickey counted 5 buttons. He counted 4 fewer buttons than Tigger. How many buttons did Tigger count? (5+4=9) *Compare 6*

Ricky says: *How did you figure out that one?****

15. Lion has 5 jacks. He has 2 more jacks than Pig. How many jacks does Pig have? (5-2=3) *Compare 5*

Ricky says: *How did you figure out that one?****

***Alternative questions to elicit problem-solving rationale:

What were you thinking about?

How did you decide/figure out that x was your answer?

Can you help Ricky understand how to do that problem?

Set 2

Setup: Ricky Raccoon, lying on left corner of table, facing child. For each problem, place the first-mentioned character to your left and the second-mentioned character to your right. Place the objects between and slightly closer to the child. Spread each set of objects slightly, so that none are touching. As you place the characters and objects, chat with the child about each, to make sure the child is familiar with character and object names. Set out the materials for the first practice problem. Other character names can be substituted if the child is more familiar with another name (e.g. bunnies/rabbits)

I brought my friend Ricky with me today. He is learning how to do math word problems in school. But he needs a friend to teach him how to do these problems. I thought that you would be able to help him. When we are all done with the problems, you can give Ricky a couple of practice problems, to see if he was listening.

I will read each problem 2 times. Wait to tell me your answer until I have finished reading the problem 2 times. Then I want you to help Ricky, and tell him how to do the problem.

You can figure out the answer to these problems any way that you want. You can use your fingers, you can use the objects that will be out here, you can count, or you can do it in your head. Take your time. You don't have to answer fast.

Let's do 2 practice problems. (No gesture. Read each problem 2 times, with natural inflection.)

Practice #1. Alligator ate 3 ice cream cones. Tiger ate 1 ice cream cone. How many ice cream cones did they eat altogether? (3+1=4)

(Explain problem if incorrect) Ricky says: *How did you figure out that one?****

Practice #2. Mouse and Lady Bug are catching fish. Mouse caught 4 fish. Lady Bug took away one of mouse's fish. How many fish does Mouse have left? (4-1=3)

(Explain problem if incorrect) Ricky says: *How did you figure out that one?****

Let's do some more problems. Some of the problems will be easy and some will be harder. Don't worry if you don't get them all right. Just do your best. Remember you can figure out the answers any way that you want.

While reading the word problems, wait for child to move manipulatives before continuing to read the problem. Remind the child not to answer until you have read the problem 2 times.

1. Bunny caught 4 frogs. Tiger caught 3 more frogs than Bunny. How many frogs did Tiger catch? (4+3=7) *Compare 3*

Ricky says: *How did you figure out that one?****

2. Mario has 9 legos. He has 4 more legos than Tweety. How many legos does Tweety have? (9-4=5) *Compare 5*

Ricky says: *How did you figure out that one?****

3. Curious George has 7 pet turtles. He has 4 more turtles than Elephant. How many turtles does Elephant have? (7-4=3) *Compare 5*

Ricky says: *How did you figure out that one?****

F. Donald Duck has 6 Easter eggs. Tigger steals 1 Easter egg from Donald. How many eggs does Donald Duck have left? (6-1=5)

Ricky says: *How did you figure out that one?****

4. Pooh ate 6 carrots. Minnie ate 2 more carrots than Pooh. How many carrots did Minnie eat? (6+2=8) *Compare 3*

Ricky says: *How did you figure out that one?****

5. Lion has 8 fish. He has 3 more fish than Panda. How many fish does Panda have? (8-3=5)) *Compare 5*

Ricky says: *How did you figure out that one?****

6. Big Bird has 5 rings. He has 3 fewer rings than Oscar. How many rings does Oscar have? (5+3=8) *Compare 6*

Ricky says: *How did you figure out that one?****

7. Thing One has 6 rabbits. He has 3 fewer rabbits than Lambie. How many rabbits does Lambie have? (6+3=9) *Compare 6*

Ricky says: *How did you figure out that one?****

F. The girl construction worker has 4 pipes. She gives 1 pipe to the boy worker. How many pipes does the girl have left? (4-1=3)

Ricky says: *How did you figure out that one?****

8. Brown puppy has 3 bones. Spot has 2 more bones than Brown puppy. How many bones does Spot have? (3+2=5) *Compare 3*

Ricky says: *How did you figure out that one?****

9. Duck ate 7 ice cream cones. He ate 3 fewer ice cream cones than Bunny. How many ice cream cones did Bunny eat? (7+3=10) *Compare 6*

Ricky says: *How did you figure out that one?****

Break…

10. Spiderman caught 4 bees. He caught 2 fewer bees than Batman. How many bees did Batman catch? (4+2=6) *Compare 6*

Ricky says: *How did you figure out that one?****

11. Pig has 4 baseballs. Monkey has 6 more baseballs than Pig. How many baseballs does Monkey have? (4+6=10) *Compare 3*

Ricky says: *How did you figure out that one?****

F. Goofy bought 3 baby chicks. Garfield took 2 of Goofy's chicks. How many chicks does Goofy have left? (3-2=1)

Ricky says: *How did you figure out that one?****

12. Warthog counted 6 smiley faces. He counted 4 fewer smiley faces than Hippo. How many smiley faces did Hippo count? (6+4=10) *Compare 6*

Ricky says: *How did you figure out that one?****

13. Elephant picked 10 flowers. He picked 2 more flowers than Shrek. How many flowers did Shrek pick? (10-2=8) *Compare 5*

Ricky says: *How did you figure out that one?****

14. Mickey counted 9 buttons. He counted 7 more buttons than Tigger. How many buttons did Tigger count? (9-7=2) *Compare 5*

Ricky says: *How did you figure out that one?****

15. Lion has 5 jacks. Pig has 2 more jacks than Lion. How many jacks does Pig have? (5+2=7) *Compare 3*

Ricky says: *How did you figure out that one?****

***Alternative questions to elicit problem-solving rationale:

What were you thinking about?

How did you decide/figure out that x was your answer?

Can you help Ricky understand how to do that problem?

APPENDIX E

PHASE THREE: WORD PROBLEM SETS

Set A

Setup: Rex, lying on left corner of table, facing child. For each problem, place the first-mentioned character to your left and the second-mentioned character to your right. Place the objects between and slightly closer to the child. Spread each set of objects slightly, so that none are touching. As you place the characters and objects, chat with the child about each, to make sure the child is familiar with character and object names. Set out the materials for the first practice problem. Other character names can be substituted if the child is more familiar with another name (e.g. bunnies/rabbits)

Today I brought my friend Rex with me. He is in Rex's class and he is learning how to do math word problems in school too. He wants you to teach him how to do these problems. When we are all done with the problems, you can give Rex a couple of practice problems, to see if he was listening.

I will read each problem 2 times. Wait to tell me your answer until I have finished reading the problem 2 times. Then I want you to help Rex, and tell him how to do the problem.

You can figure out the answer to these problems any way that you want. You can use your fingers, you can use the objects that will be out here, you can count, or you can do it in your head. Take your time. You don't have to answer fast.

Let's do 2 practice problems. (Read each problem 2 times, with natural inflection.)

Practice #1. Cow has 4 pet ducks. He gave 1 duck to Bear. How many ducks does Cow have left? (4-1=3)

(Explain problem if incorrect) Rex says: *How did you figure out that one?*

Practice #2. Elephant has 5 squishy dinosaurs. Tiger took away 2 of elephant's dinosaurs. How many dinosaurs does Elephant have left? (5-2=3)

(Explain problem if incorrect) Rex says: *How did you figure out that one?*

Let's do some more problems. Some of the problems will be easy and some will be harder. Don't worry if you don't get them all right. Just do your best. Remember you can figure out the answers any way that you want.

While reading the word problems, wait for child to move manipulatives before continuing to read the problem. Remind the child not to answer until you have read the problem 2 times.

1. Baboon has 2 basketballs. He has 6 fewer basketballs than Zebra. How many basketballs does Zebra have? (2+6=8) *Compare 6* (C6)-Traditional wording

Rex says: *How did you figure out that one?****

F1. Grasshopper bought 4 candy hearts. Lady Bug ate 2 of Grasshopper's candy hearts. How many candy hearts does Grasshopper have left? (4-2=2) Filler

Rex says: *How did you figure out that one?****

2. Monkey ate 2 oranges. He ate 5 fewer oranges than Dinosaur. How many oranges did Dinosaur eat? (2+5=7) C6-Traditional wording

Rex says: *How did you figure out that one?****

3. Lion bought 5 horseshoes. He bought 4 fewer horseshoes than Pony. How many horseshoes did Pony buy? (5+4=9) C6-Traditional wording

Rex says: *How did you figure out that one?****

F2. Moose and Wolf love to play football. Moose has 5 footballs. He gave one of his footballs to Wolf. How many footballs does Moose have left? (5-1=4) Filler

Rex says: *How did you figure out that one?****

4. Giraffe and Dinosaur love to eat strawberries. You need to figure out how many strawberries Dinosaur ate. Giraffe ate 8 strawberries. Dinosaur ate 2 more strawberries than Giraffe ate. How many strawberries did Dinosaur eat? (8+2=10) C6-Reworded

Rex says: *How did you figure out that one?****

F3. Chicken and Rooster like to eat Lady Bugs for lunch. Chicken found 3 Ladybugs. Rooster took one of Chicken's Lady Bugs. How many Lady Bugs does Chicken have left? (3-2=1) Filler

Rex says: *How did you figure out that one?****

5. Cow and Monkey are carving Jack-O-Lanterns for Halloween. You need to figure out how many Jack-O-Lanterns Monkey carved. Cow carved 5 Jack-O-Lanterns. Monkey carved 2 more Jack-O-Lanterns than Cow carved. How many Jack-O-Lanterns did Monkey carve? (5+2=7) C6-Reworded

Rex says: *How did you figure out that one?****

6. Monkey and Rhino always eat apples at snack time. You need to figure out how many apples Rhino ate. Monkey ate 2 apples. Rhino ate 7 more apples than Monkey ate. How many apples did Rhino eat? (2+7=9) C6-Reworded

Rex says: How did you figure out that one?***

F4. It's Halloween and Knight and Pirate are catching ghosts. Knight caught 4 ghosts. Pirate stole one of Knight's ghosts. How many ghosts does Pirate have left?

(4-1=3) Filler

Rex says: *How did you figure out that one?****

Break…

7. Butterfly collected 3 shells. He collected 7 fewer shells than Snail. How many shells did Snail collect? (3+7=10) C6-Traditional wording

Rex says: *How did you figure out that one?****

F5. Pig has 8 pet black cats. He gave one of his black cats to Lamb. How many black cats does pig have now? (8-1=7) Filler

Rex says: *How did you figure out that one?****

8. Kermit has 6 Christmas presents. He has 4 fewer presents than Thing Two. How many presents does Thing Two have? (6+4=10) C6-Traditional wording

Rex says: *How did you figure out that one?****

F6. Spider and Bee were hunting for pretty fall leaves. Spider found 6 pretty leaves. He gave one leaf to Bee. How many leaves does Spider have left? (6-1=5) Filler

Rex says: *How did you figure out that one?****

9. It is Christmas time. Puppy and Grinch are counting jingle bells. You need to figure out how many jingle bells Grinch counted. Puppy counted 6 jingle bells. Grinch counted 3 more jingle bells than Puppy counted. How many jingle bells did Grinch count? (6+3=9) C6-Reworded

Rex says: *How did you figure out that one?****

10. It is snowing. Puppy and Frog are catching snow flakes. You need to figure out how many snow flakes Frog caught. Puppy caught 3 snow flakes. Frog caught 4 more snow flakes than Puppy caught. How many snow flakes did Frog catch?

(3+4=7) C6-Reworded

Rex says: *How did you figure out that one?****

***Alternative questions to elicit problem-solving rationale:

What were you thinking about?

How did you decide/figure out that x was your answer?

Can you help Rex understand how to do that problem?

Set B

Setup: Rex, lying on left corner of table, facing child. For each problem, place the first-mentioned character to your left and the second-mentioned character to your right. Place the objects between and slightly closer to the child. Spread each set of objects slightly, so that none are touching. As you place the characters and objects, chat with the child about each, to make sure the child is familiar with character and object names. Set out the materials for the first practice problem. Other character names can be substituted if the child is more familiar with another name (e.g. bunnies/rabbits)

Today I brought my friend Rex with me. He is in Rex's class and he is learning how to do math word problems in school too. He wants you to teach him how to do these problems. When we are all done with the problems, you can give Rex a couple of practice problems, to see if he was listening.

I will read each problem 2 times. Wait to tell me your answer until I have finished reading the problem 2 times. Then I want you to help Rex, and tell him how to do the problem.

You can figure out the answer to these problems any way that you want. You can use your fingers, you can use the objects that will be out here, you can count, or you can do it in your head. Take your time. You don't have to answer fast.

Let's do 2 practice problems. (Read each problem 2 times, with natural inflection.)

Practice #1. Brown Puppy caught 7 butterflies. He gave 2 butterflies to Spot. How many butterflies does Brown Puppy have left? (7-2=5)

(Explain problem if incorrect) Rex says: *How did you figure out that one?*

Practice #2. The Knight has 2 spinning tops. Pirate took 1 of the Knight's spinning tops. How many spinning tops does the Knight have left? (2-1=1)

(Explain problem if incorrect) Rex says: *How did you figure out that one?*

Let's do some more problems. Some of the problems will be easy and some will be harder. Don't worry if you don't get them all right. Just do your best. Remember you can figure out the answers any way that you want.

While reading the word problems, wait for child to move manipulatives before continuing to read the problem. Remind the child not to answer until you have read the problem 2 times.

1. Zebra has 7 basketballs. He has 2 fewer basketballs than Baboon. How many basketballs does Baboon have? (7+2=9) C6-T

Rex says: *How did you figure out that one?****

F1. Lady Bug had 5 candy hearts. Grasshopper stole 1 of Lady Bug's candy hearts. How many candy hearts does Lady Bug have left? (5-1=4) Filler

Rex says: *How did you figure out that one?****

2. Dinosaur ate 4 oranges. He ate 2 fewer oranges than Monkey. How many oranges did Monkey eat? (4+2=6) C6-T

Rex says: *How did you figure out that one?****

3. Horse and Lion are buying some new horseshoes. You need to figure out how many horseshoes Lion bought. Horse bought 3 horseshoes. Lion bought 2 more horseshoes than Horse bought. How many horseshoes did Lion buy? (3+2=5) C6-R

Rex says: *How did you figure out that one?****

F2. Wolf has 3 footballs. He gave 2 footballs to Moose. How many footballs does Wolf have left? (3-2=1) Filler

Rex says: *How did you figure out that one?****

4. Giraffe ate 3 strawberries. He ate 6 fewer strawberries than Dinosaur. How many strawberries did Dinosaur eat? (3+6=9) C6-T

Rex says: *How did you figure out that one?****

F3. Chicken caught 8 lady bugs. Rooster took 1 lady bug from Chicken. How many lady bugs does chicken have left? (8-1=7) Filler

Rex says: *How did you figure out that one?****

5. Cow and Monkey bought some Jack-O-Lanterns for Halloween. You need to figure out how many Jack-O-Lanterns Monkey bought. Cow bought 2 Jack-O-Lanterns. Monkey bought 5 more Jack-O-Lanterns than Cow bought. How many Jack-O-Lanterns did Monkey buy? (2+5=7) C6-R

Rex says: *How did you figure out that one?****

6. Monkey and Rhino love to eat apples. You need to figure out how many apples Rhino ate. Monkey ate 4 apples. Rhino ate 5 more apples than Monkey ate. How many apples did Rhino eat? (4+5=9) C6-R

Rex says: *How did you figure out that one?****

F4. Pirate made some Halloween ghosts. He gave 1 ghost to the Knight. How many ghosts does Pirate have left? (4-1=3) Filler

Rex says: *How did you figure out that one?****

Break…

7. Butterfly found 6 white shells on the beach. He found 2 fewer shells than Snail. How many shells did Snail find? (6+2=8) C6-T

Rex says: *How did you figure out that one?****

F5. Lamb made 6 black cat faces. He gave one black cat face to Pig. How many black cat faces does Lamb have left? (6-1=5) Filler

Rex says: *How did you figure out that one?****

8. Thing Two has 5 Christmas presents. He has 3 fewer presents than Kermit. How many presents does Kermit have? (5+3=8) C6-T

Rex says: *How did you figure out that one?****

F6. Bee and Fly collect pretty leaves. Bee found 4 leaves. Fly took 2 of Bee's leaves. How many leaves does Bee have left? (4-2=2)

Rex says: *How did you figure out that one?****

9. Puppy and Grinch are hunting for jingle bells. You need to figure out how many jingle bells Grinch found. Puppy found 3 jingle bells. Grinch found 7 more jingle bells than Puppy found. How many jingle bells did Grinch find? (3+7=10) C6-R

Rex says: *How did you figure out that one?****

10. Frog and Puppy like to play outside when it is snowing. Today they are counting snowflakes. You need to figure out how many snowflakes Puppy counted. Frog counted 4 snowflakes. Puppy counted 3 more snowflakes than Frog. How many snowflakes did Puppy count? (4+3=7) C6-R

Rex says: *How did you figure out that one?****

***Alternative questions to elicit problem-solving rationale:

What were you thinking about?

How did you decide/figure out that x was your answer?

Can you help Rex understand how to do that problem?

REFERENCES

Alibali, M. W. (2005). Mechanisms of change in the development of mathematical reasoning. In R. Kail (Ed.), *Advances in child development and behavior,* (Vol. 33, pp. 79-123). New York: Elsevier Academic Press.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*(6), 593-613.

Arvedson, P. J. (2002). Young children with specific language impairment and their numerical cognition. *Journal of Speech Language and Hearing Research, 45*(5), 970-982.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208.

Blessing, S. B. & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(3),* 792 – 810.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, 46*(3), 317-326.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology, General, 136(4),* 539-550.

Cain, K. & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5-6), 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory and Cognition, 29(6)*, 850-859.

Capone, N. C., & McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research, 47(1)*, 173-186.

Carpenter, T., & Moser, J. (1982). The development of addition and subtraction problem-solving skills. In T. Carpenter, J. Moser & T. Romberg (Eds.), *Addition and subtraction: A cognitive perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Carr, M., & Hettinger, H. (2003). Perspectives on mathematical strategy development. In J. Royer (Ed.), *Mathematical cognition* (pp. 33-68). Greenwich, CT: Information Age Publishing Inc.

Case, R., Okamoto, Y., Griffin, S., Siegler, R. S., & Kaeting, D. P. (1996). *The role of central conceptual structures in the development of children's thought.* Chicago, Ill.: Society for Research in Child Development.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36*(2), 151-164.

Catts, H., Fey, M., Tomblin, J., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45,* 1142-1157.

Chien, Y.-C., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition, 1,* 225-295.

Church, R., Ayman-Nolley, & Vasich. (2007, March). *The role of gesture in the creation of conceptual representation.* Presented at the Society for Research in Child Development, Boston, MA.

Church, R., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23,* 43-71.

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7*(2), 211-232.

Cowan, R., Donlan, C., Newton, E. J., & Lloyd, D. (2005). Number skills and knowledge in children with specific language impairment. *Journal of Educational Psychology, 97*(4), 732-744.

Culatta, B. & Wiig, E. H. (2006). Language disabilities in school-age children and youth. In N. B. Anderson & G. H. Shames (Eds.), *Human communication disorders: An introduction* (pp. 352-385). Boston, MA: Pearson Education, Inc.

Cummins, D. D. (1991). Children’s interpretations of arithmetic word problems. *Cognition and Instruction, 8*(3), 261-289.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*(4), 405-438.

De Corte, E. & Verschaffel, L. (1987). *The effect of semantic structure on first graders’ strategies for solving addition and subtraction word problems, 18(5)*. 363-381.

De Corte, E., Verschaffel, L., & Dewin, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology, 77*(4), 460-470.

Donlan, C. (1998a). *The development of mathematical skills.* Sussex, England: Psychology Press.

Donlan, C. (1998b). The importance of educational transitions. *International Journal of Language & Communication Disorders, 33*(2), 212-215.

Donlan, C., Bishop, D. V. M., & Hitch, G. J. (1998). Magnitude comparisons by children with specific language impairments: Evidence of unimpaired symbolic processing. *International Journal of Language & Communication Disorders, 33*(2), 149-160.

Donlan, C., Cowan, R., Newton, E., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition, 103,* 23-33.

Donlan, C., & Gourlay, S. (1999). The importance of non-verbal skills in the acquisition of place-value knowledge: Evidence from normally-developing and language-impaired children. *British Journal of Developmental Psychology, 17,* 1-19.

Donlan, C., & Newton, L. (2007, April). *Mathematical and reasoning skills of children with specific language impairments: Findings from the number talk and puzzle talk projects.* Paper presented at the Afasic Conference, Warwick, England.

Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience and education.* New York: Psychology Press.

Dunn, L. M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised.* Circle Pines, MN: American Guidance Service.

Ellis Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42(5)*, 1249-1260.

Ellis Weismer, S. & Hesketh, L. J. (1993). The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech and Hearing Research, 36(3),* 1013-1024.

Elman, J. (2007). *On words and dinosaur bones: Where is the meaning?* Keynote Speech presented at the 29th Annual Cognitive Science Society, Nashville, TN.

Evans, J. L., Alibali, M. W., & McNeil, N. M. (2001). Divergence of verbal expression and embodied knowledge: Evidence from speech and gesture in children with specific language impairment. *Language and Cognitive Processes, 16*(2-3), 309-331.

Fazio, B. B. (1994). The counting abilities of children with specific language impairment - a comparison of oral and gestural tasks. *Journal of Speech and Hearing Research, 37*(2), 358-368.

Fazio, B. B. (1996). Mathematical abilities of children with specific language impairment: A 2-year follow up. *Journal of Speech and Hearing Research, 39*(4), 839-849.

Fazio, B. B. (1999). Arithmetic calculation, short-term memory, and language performance in children with specific language impairment: A 5-year follow-up. *Journal of Speech Language and Hearing Research, 42*(2), 420-431.

Fey, M. E., Catts, H., Proctor-Williams, K., Tomblin, J., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47,* 1301-1318.

Fuson, K. C. (1988). *Children's counting and concepts of number.* New York: Springer-Verlag.

Fuson, K., Carroll, W., & Landis, J. (1996). Levels in conceptualizing and solving addition and subtraction compare word problems. *Cognition And Instruction, 14*(3), 345-371.

Garber, P., Alibali, M. W., & Goldin-Meadow, S. (1998). Knowledge conveyed in gesture is not tied to the hands. *Child Development, 69*(1), 75-84.

Geary, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child & Adolescent Psychiatry, 9*(Suppl. 2), 11-16.

Gelman, R., & Gallistel, C. (1978). *The Child's Understanding of Number.* Cambridge, MA: Harvard University Press.

Ginsburg, H. P. (1983). *The development of mathematical thinking.* NY: Academic Press, Inc.

Ginsburg, H. P. (1989). *Children's arithmetic: How they learn it and how you teach it.*. Austin, TX: Pro-Ed.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, 91*(4), 720-730.

Greer, B. (1990). Conceptual obstacles to the development of the concepts of multiplication and division. In H. Mandl, E. DeCorte, N. Bennett & H. Friedrich (Eds.), *Learning and instruction* (pp. 461-476). Oxford, England: Pergamon Press.

Griffin, S. (2003). The development of math competence in the preschool and early school years: Cognitive foundations and instructional strategies. In J. Royer (Ed.), *Mathematical cognition* (pp. 1-32). Greenwich, CT: Information Age Publishing Inc.

Hauser, M., & Spelke, E. (2004). Evolutionary and developmental foundations of human knowledge. In M. Gazzaniga (Ed.), *The cognitive neurosciences, III.* Cambridge, MA: MIT Press.

Hoover, W. & Gough, P. B. (1990). The simple view of reading. *Reading & Writing, 2*(2), 127-160.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development, 54,* 84-90.

Hughes, M. (1986). *Children and number: Difficulties in learning mathematics.* New York: Basil Blackwell, Inc.

Johnston, J. R. & Smith, L. B. (1989). Dimensional thinking in language impaired children. *Journal of Speech and Hearing Research, 32,* 33-38.

Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1995). Calculation abilities in young-children with different patterns of cognitive-functioning. *Journal of Learning Disabilities, 28*(1), 53-64.

Jordan, N., Hanich, L., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology, 85*(2), 103-119.

Justice, L. M. (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in the Schools, 37,* 284-297.

Justice, L. M. (2008, November). *How much is enough? The intensity evidence in language intervention.* Paper presented at the annual meeting of the American Speech, Language, Hearing Association, Chicago, IL.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163-182.

Kintsch, W. (1994a). Text comprehension, memory, and learning. *American Psychologist*, *49:4*, 294-303.

Kintsch, W. (1994b). The psychology of discourse processing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 721-739). San Diego: Academic Press.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge; New York, NY: Cambridge University Press.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109-129.

Klein, D. (2003). A brief history of American K-12 mathematics education in the 20th century. In J. Royer (Ed.), *Mathematical cognition* (pp. 175-225). Greenwich, CT: Information Age Publishing.

LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematical connections: A computer model of arithmetic word problem solving. *Cognitive Science, 20*(3), 357-407.

Leonard, L. B. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Development of calculation abilities in young children. *Journal of Experimental Child Psychology, 53*(1), 72-103.

Lewis, A. B., & Mayer, R. E. (1987). Students miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology, 79*(4), 363-371.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review. 109*(1), 35-54.

Mainela-Arnold, E., Evans, J. L., & Alibali, M. W. (2006). Understanding conservation delays in children with specific language impairment: Task representations revealed in speech and gesture. *Journal of Speech, Language, and Hearing Research, 49(6)*, 1267-1279.

Marshall S. P. (1995). *Schemas in problem solving.* Cambridge, MA: Cambridge University Press.

Masterson, J. J., Evans, L. H., & Aloia, M. (1993). Verbal Analogical Reasoning in Children With Language-Learning Disabilities. *Journal of Speech and Hearing Research, 36*(1), 76-82.

Mayer, R. (2003). Mathematical problem solving. In J. Royer (Ed.), *Mathematical cognition* (pp. 69-92). Greenwich, CT: Information Age Publishing.

McGregor, K., & Rost, G. (2007, November). *Do children with SLI take advantage of gestural and statistical cues to determine word referents?* Presented at the American Speech, Language, and Hearing Association Annual Convention, Boston, MA.

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior, 24*(2), 131-150.

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(2), 416-433.

Montgomery, J. W. (2003). Working memory and comprehension in children with language impairment: What we know so far. *Journal of Communication Disorders, 36(3)*, 221-231.

Nation, K., & Norbury, C. F. (2005). Why reading comprehension fails - insights from developmental disorders. *Topics in Language Disorders, 25*(1), 21-32.

Newcomer, P, & Hammill, D. (1988). *Test of Language Development-2 Primary*. Austin, TX: Pro-Ed.

Nippold, M. (1994). Third-order verbal analogical reasoning: A developmental study of children and adolescents. *Contemporary Educational Psychology, 19*(1), 101-107.

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90*(2), 142-171.

Nunes, T., Schliemann, A., & Carraher, D. (1993). *Street mathematics and school mathematics.* Cambridge, England: Cambridge University Press.

Okamoto, Y. (1996). Modeling children's understanding of quantitative relations in texts: A developmental perspective. *Cognition and Instruction, 14*(4), 409-440.

Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. In R. Case, Y. Okamoto, S. Griffin, A. McKeough, C. Bleiker, & B. Henderson (Eds.), *The role of central conceptual structures in the development of children's thought* (pp. 27-58). Monographs of the Society for Research in Child Development.

Paul, R. (2006). *Language disorders from infancy through adolescence: Assessment and intervention..* St. Louis: Mosby.

Perfetti, C. (1999). Cognitive research and the misconceptions of reading education. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading* (pp. 42-58). Oxford, UK: Blackwell Publishing.

Pisoni, D. B., & Cleary, M. (2004). Learning, memory, and cognitive processes in deaf children following cochlear implantation. In F. G. Zeng, A. N. Popper, & R. R. Fay (Eds.), *Springer handbook of auditory research: Vol. 20. Cochlear implants: Auditory prostheses & electric hearing*. New York: Springer-Verlag.

Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*(1), 49-101.

Riley, M., Greeno, J., & Heller, J. (1983). Development of children's ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). NY: Academic Press.

Rumelhart, D., Smolensky, P., McClelland, J., & Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (pp. 7-57). Cambridge, MA: MIT Press.

Samelson, V. M., McMurray, B., Lee, S. H., & Tomblin, J. B. (2007, March). *Eye movement during spoken word recognition in normal and language-impaired adolescents: Activation and inhibition effects.* Poster presented at the Society for Research in Child Development Biennial Convention, Boston, MA.

Schoenfeld, A., H. (1985). *Mathematical problem solving.* New York: Academic Press, Inc.

Semel, E., Wiig, E. H., & Secord, W. A. (1995). *Clinical Evaluation of Language Fundamentals – Third Edition.* San Antonio, TX: Psychological Corporation.

Share, D., & Stanovich, K. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education, 1,* 1-58.

Siegler, R., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. Kail (Ed.), *Advances in Child Development and Behavior* (pp. 1-42). San Diego, California: Elsevier.

Spaulding, T. J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research, 51(1)*, 16-34.

Stone, C. A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31(4)*, 344-364.

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(3)*, 471-791.

Thevenot, C., Devidal, M., Barrouillet, P., & Fayol, M. (2007). Why does placing the question before an arithmetic word problem improve performance? *Quarterly Journal of Experimental Psychology, 60*(1), 43-56.

Tomblin, J. B., Records, N. L., & Zhang, X. Y. (1996). A system for the diagnosis of language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39(4)*, 1284-1294.

Tomblin, J. B., Zhang, X. Y., Buckwalter, P., & O'Brien, M. (2003). The stability of primary language disorder: Four years after kindergarten diagnosis. *Journal of Speech, Language, and Hearing Research, 46*(6), 1283-1296.

Tomblin, J. (2005, June). *Adolescent outcomes of developmental language disorder in kindergarten.* Presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Ukrainetz, T. A. (2006). The implications of RTI and EBP for SLPs: Commentary on L. M. Justice. *Language, Speech, and Hearing Services in the Schools*, *37*, 298-303.

van der Lely, H. K. J., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition, 62*(3), 245-290.

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

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking & Learning, 1*(3), 195.

Vygotsky, L. S. (1987). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Wechsler, D. (1991). *Wechsler intelligence scale for children (3rd ed.).* San Antonio: The Psychological Corporation.

Woodward, J. (2004). Mathematics education in the United States: Past to present. *Journal of Learning Disabilities, 37*(1), 16-31.