Problem Representation and Mathematical Problem Solving of Students of Varying Math Ability

Jennifer Lee Krawec
University of Miami, j.krawec1@miami.edu

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/455

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
PROBLEM REPRESENTATION AND MATHEMATICAL PROBLEM SOLVING OF
STUDENTS OF VARYING MATH ABILITY

By

Jennifer Lee Krawec

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2010
A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

PROBLEM REPRESENTATION AND MATHEMATICAL PROBLEM SOLVING OF
STUDENTS OF VARYING MATH ABILITY

Jennifer Lee Krawec

Approved:

Marjorie Montague, Ph.D.
Professor of Teaching and Learning

Terri A. Scandura, Ph.D.
Dean of the Graduate School

Batya Elbaum, Ph.D.
Associate Professor of Teaching
and Learning

Wendy Cavendish, Ph.D.
Assistant Professor of Teaching
and Learning

Soyeon Ahn, Ph.D.
Assistant Professor of Educational
and Psychological Studies
The purpose of this study was to examine differences in math problem solving among students with learning disabilities (LD), low-achieving (LA) students, and average-achieving (AA) students. The primary interest was to analyze the problem representation processes students use to translate and integrate problem information as they solve math word problems. Problem representation processes were operationalized as (a) paraphrasing the problem and (b) visually representing the problem. Paraphrasing accuracy (i.e., paraphrasing relevant information, paraphrasing irrelevant linguistic information, and paraphrasing irrelevant numerical information), visual representation accuracy (i.e., visual representation of relevant information, visual representation of irrelevant linguistic information, and visual representation of irrelevant numerical information), and problem-solving accuracy were measured in eighth-grade students with LD (n = 25), LA students (n = 30), and AA students (n = 29) using a researcher-modified version of the Mathematical Processing Instrument (MPI). Results indicated that problem-solving accuracy was significantly and positively correlated to relevant information in both the paraphrasing and the visual representation phases and significantly negatively correlated to linguistic and numerical irrelevant information for
the two constructs. When separated by ability, students with LD showed a different profile as compared to the LA and AA students with respect to the relationships among the problem-solving variables. Mean differences showed that students with LD differed significantly from LA students in that they paraphrased less relevant information and also visually represented less irrelevant numerical information. Paraphrasing accuracy and visual representation accuracy were each shown to account for a statistically significant amount of variance in problem-solving accuracy when entered in a hierarchical model. Finally, the relationship between visual representation of relevant information and problem-solving accuracy was shown to be dependent on ability after controlling for the problem-solving variables and ability. Implications for classroom instruction for students with and without LD are discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
</tbody>
</table>

## CHAPTER 1
INTRODUCTION ........................................... 1

- Mathematical Problem Solving ........................................... 4
- Problem Representation .................................................. 6
- Paraphrasing ............................................................. 8
- Visual Representation ................................................... 9

Students with Learning Disabilities ................................... 12
- Math LD ................................................................. 14
- Computation ............................................................ 15
- Problem Solving .......................................................... 16

Summary ................................................................. 18

Statement of the Problem .................................................. 18

## CHAPTER 2
REVIEW OF THE LITERATURE ......................................... 21

- Students with LD and Low-Achieving Students ....................... 21
- Reading ................................................................. 23
- Mathematics ............................................................ 25
- Paraphrasing ............................................................. 28
- Visual Representation ................................................... 34
  - Categories of Visual Representation .............................. 35

Overview ................................................................. 38

Statement of Purpose ..................................................... 40

Research Questions ....................................................... 40

Hypotheses ............................................................... 41

## CHAPTER 3
METHODS ............................................................... 44

- Participants ............................................................. 44
- Measures ................................................................. 46

Mathematical Processing Instrument .................................... 47
  - Scoring ................................................................. 49
  - Accuracy of Paraphrasing ............................................. 49
  - Accuracy of Visual Representation .................................. 52
  - Problem-Solving Accuracy ............................................. 52

Test of Mathematical Ability ............................................ 52
  - Scoring ................................................................. 53
Procedures ................................................................. 55
Consent ................................................................. 55
Background Information ........................................... 55
Assessment ............................................................. 55
  Administration ....................................................... 56
Coding ................................................................. 56
  Reliability ........................................................... 57
Design and Analysis ................................................ 57

CHAPTER 4
RESULTS ........................................................................ 63
Preliminary Analyses .................................................. 64
  Question 1: Relationships among Problem-Solving Variables . 68
    Average-Achieving Students ..................................... 71
    Low-Achieving Students ......................................... 73
    Students with LD .................................................. 74
    Summary .................................................................. 75
  Question 2: Differences in Problem-Solving Variables by Ability Group ............................................. 76
  Question 3: Effects of Problem-Solving Processes on Accuracy ......................................................... 82
    Interactions ........................................................... 84
    Summary .................................................................. 88

CHAPTER 5
DISCUSSION ................................................................. 91
  Question 1: Relationships among Problem-Solving Variables ......................................................... 92
  Question 2: Differences in Problem-Solving Variables by Ability Group ............................................ 97
  Question 3: Effects of Problem-Solving Processes on Accuracy ....................................................... 102
    Limitations .......................................................... 105
    Implications for Instruction ...................................... 107
    Directions for Future Research ................................. 111

APPENDICES ............................................................... 113
REFERENCES ............................................................... 132
# LIST OF FIGURES

| CHAPTER 1 | Figure 1.1 | Conceptual Framework: Math Problem-Solving Process | 6 |
| CHAPTER 1 | Figure 1.2 | FCAT Math Extended Response | 7 |
| CHAPTER 3 | Figure 3.1 | Modification of the original MPI | 49 |
| CHAPTER 3 | Figure 3.2 | Paraphrasing and Visual Representation Scoring Example | 51 |
| CHAPTER 3 | Figure 3.3 | Pictorial – Schematic Continuum | 54 |
| CHAPTER 4 | Figure 4.1 | Significant Relationships by Ability Group | 76 |
| CHAPTER 4 | Figure 4.2 | Interaction Between Ability and Visual Representation of Relevant Information | 88 |
| CHAPTER 4 | Figure 4.3 | Hierarchical Multiple Regression Model | 89 |
| CHAPTER 5 | Figure 5.1 | Effects of Misinterpretation of Information | 94 |
LIST OF TABLES

CHAPTER 3

Table 3.1 Participant Information by Ability Group ................................. 46

CHAPTER 4

Table 4.1 Sample Characteristics: SES and Ethnicity ......................... 64
Table 4.2 Means and Standard Deviations: Student Age .................... 65
Table 4.3 Means and Standard Deviations: Math FCAT, Reading
FCAT, Computation .......................................................... 66
Table 4.4 One-Way ANOVAs Summary Statistics ............................. 67
Table 4.5 One-Way ANOVAs Pairwise Comparisons .......................... 68
Table 4.6 Correlation Matrix for Problem-Solving Variables:
All Groups ...................................................................... 70
Table 4.7 Correlation Matrix for Problem-Solving Variables:
By Group ....................................................................... 72
Table 4.8 Means and Standard Deviations: Problem-Solving
Variables ...................................................................... 79
Table 4.9 A Univariate Effect from MANOVA ................................. 80
Table 4.10 Univariate Pairwise Comparisons on Problem-Solving
Variables ...................................................................... 81
Table 4.11 Hierarchical Regression: Model Summary Statistics .......... 83
Table 4.12 Hierarchical Regression Coefficients, Standard Errors,
and Beta Values ................................................................ 85
Table 4.13 Model Summary of Interactions Separately Entered in
the Fourth Block .............................................................. 86
Chapter 1

Introduction

The ability to solve mathematical word problems has long been recognized as an essential component of math competency, but only recently is it being reflected in school curricula. As far back as 1980, the National Council of Teachers of Mathematics (NCTM) proposed an overhaul of existing math instruction in schools across the United States. An Agenda for Action: Recommendations for School Mathematics of the 1980s outlined eight fundamental changes with an instructional focus on problem solving ranking first. The report described the existing curriculum as embodying a back-to-basics philosophy where math competence is erroneously tied to foundational computational skills and it called for a shift in focus to problem analysis and interpretation (NCTM, 1980). Members of NCTM recognized the need to link instruction to the increasing requirements of the job market. Basic computational fluency was being replaced with faster, more accurate technology available in the classroom (i.e., calculators), thus increasing the value placed on proficiency in higher order conceptual mathematics. The Council cited problem-solving ability as the measure of both personal and national mathematical competence (NCTM, 1980).

These recommendations for increased mathematical problem-solving competence were echoed by NCTM in 1989 (NCTM, 1989) and again in 2000 (NCTM, 2000); twenty years after their initial report NCTM still found itself advocating for the advancement of the curriculum beyond rote acquisition of procedural and declarative knowledge. In Principles and Standards for School Mathematics (NCTM, 2000), the Council reiterated its position by again denouncing the product-oriented curriculum, holding it accountable
for the number of students neither committed to nor engaged in learning valuable mathematics skills. In order to resolve this situation, the Council proposed five processes that should be incorporated into the mathematics curriculum at every grade level: problem solving, reasoning and proof, connections, communication, and representation. Accompanying content and performance standards for these five processes were provided for kindergarten through to the twelfth grade as a way to establish uniform expectations and measure student progress.

The position of the NCTM was further validated with the release of the *Workforce Readiness Report Card*, a report generated by a coalition of organizations concerned with the preparation of the most recent workforce (Conference Board, Inc., 2006). Of great concern were the increasingly global availability of labor and the decreasing competency of American employees to compete with the international market. As the “nation’s long-term ability to succeed… hinges on the abilities of today’s students” (p. 11), the report focused on employers’ ratings of recently-graduated-students’ employee performance on a priori-identified skills. Of a sample of 400 employers across the U.S., more than half rated high school graduates as “deficient” in the basic mathematical knowledge and skills and two-thirds rated the graduates as deficient in higher-order critical thinking and problem solving; these latter applied skills rank in employers’ top skills necessary for workforce success. Though competence in the basic skills is obviously a necessary prerequisite for all the high-ranking applied skills listed by employers (i.e., professionalism/work ethic, oral and written communication, teamwork/collaboration, and critical thinking/problem solving), a clear relationship exists specifically between mathematical skills and critical thinking/problem-solving proficiency (Conference Board
Inc., 2006). In summary, the study recognized that students’ ability to utilize basic mathematical skills comprises the foundation of their applied skills in problem solving, and that those critical thinking/problem-solving skills were key factors in evaluating their success (or, more often, their lack of success) in the workplace. Ultimately, the conclusion was that high school graduates were not sufficiently prepared to perform the entry level jobs they were expected to fill. Further, critical thinking/problem solving was most frequently identified (77.8% of respondents) as a skill likely to increase in importance over the next five years (p. 49).

Due to both the 2000 publication of *Principles and Standards* and the increasing expectations of the American workforce, there has been an observable change in the way mathematics is viewed in the U.S.; this shift is reflected in the national move to assess higher order thinking in mathematics as well as in the growing body of research in mathematical problem solving (Jitendra, Griffin, Deatline-Buchman, & Sczesniak, 2007). As these reasoning and processing skills become a substantial component of the math curriculum, the ability to solve math problems is increasingly essential to academic success. Results on student proficiency from the most recent *Trends in International Mathematics and Science Study* (TIMSS; Gonzales et al., 2008) have revealed both good and bad news. This ongoing study collects data every four years comparing the performance of fourth- and eighth-grade American students to that of students in other countries. On the positive side, the latest data show that U.S. student scores in fourth and eighth grade have increased since the 1995 assessment by 11 points and 16 points respectively (Gonzales et al., 2008). Unfortunately, so have the scores of competing nations; in the last four years U.S. students have made no relative gains and are still
performing in the middle of the pack compared to other participating countries. Thus, while some progress has been made, American students continue to struggle in mathematics. Research has substantiated this observation and shown that word problems are particularly difficult for students who are low achieving or have LD (Fuchs & Fuchs, 2002; Geary, 2003; Hanich, Jordan, Kaplan, & Dick, 2001; Montague & Applegate, 1993). In order to better conceptualize the potential impediments these students face, an explication of the problem-solving process follows.

**Mathematical Problem Solving**

In the context of this study, the definition of math word problems contains several elements: a) interpretation and analysis, not merely computational operations embedded in word form (Carpenter, Ansell, Franke, Fennema, & Weisbeck, 1993; Cawley & Miller, 1986; Passolunghi, Mazocchi, & Fiorillo, 2005); b) both single and multiple steps (Fuchs, Fuchs, & Prentice, 2004; Montague & Applegate, 1993); and c) contextually straightforward as well as more complex forms which include irrelevant information (Fuchs & Fuchs, 2002; Passolunghi et al., 2005).

According to Mayer (1985), there are two major phases involved in mathematical word problem solving: *problem representation* and *problem solution*. Problem representation is composed of two substages: *problem translation*, which relies on linguistic skills needed to comprehend what the problem is saying, and *problem integration*, which depends on the ability to mathematically interpret the relationships among the problem parts to form a structural representation. For the purposes of this study, the first substage, *problem translation*, will be synonymous with paraphrasing, and the second substage, *problem integration*, will be synonymous with visual representation.
The second general phase, problem solution, is composed of the substages solution planning, determining which operations to use and the order in which to use them, and solution execution, carrying out the planned computations in order to solve the problem. Mayer’s model illustrates why mathematical word problems are such a struggle for students of all ages; that is, each phase of the problem-solving process is complex, and the correct solution depends on the accuracy of each of the preceding substages (Jitendra et al., 2007). Mayer’s model of problem solving provides the conceptual framework for this study to investigate processing differences among students.

Through research on cognitive strategy instruction, Montague and colleagues have validated seven cognitive processes which support successful problem solving and can be placed sequentially into Mayer’s model (Montague, 1992; Montague, 1997; Montague & Applegate, 1993; Montague & Applegate, 2000; Montague & Bos, 1986). Montague has developed these processes into a program for solving math word problems called Solve It! (Montague, 2003). To further conceptualize the steps of the problem-solving process and build from this basic framework, Montague’s model has been integrated with Mayer’s model to illustrate both the phases of the process and the functions required to carry them out (see Figure 1.1). Specifically, problem translation consists of students reading the problem for understanding and then paraphrasing the problem in their own words. Problem integration is when the student visualizes the problem by making a schematic representation. The student then hypothesizes or makes a plan to solve the problem and estimates a reasonable answer during the solution planning stage. The final stage, solution execution, is when students compute or do the arithmetic, and then check to make sure everything is right.
Figure 1.1. Conceptual framework of the math problem-solving process. This figure illustrates an integration of the work of Mayer (1985) and Montague (2003).

**Problem representation.** The overwhelming majority of problem-solving research has focused on the latter part of Mayer’s model, problem solution (e.g., Fuchs & Fuchs, 2002; Hanich et al., 2001; Jitendra & Hoff, 1996; McNear, 1990; Montague, Warger, & Morgan, 2000). It is not just in research that student proficiency is measured by solution accuracy; in schools mathematical ability is generally determined by this phase. For example, the Florida annual assessment of student performance (Florida Comprehensive Assessment Test; FCAT) allots points for students’ ability to justify their answers on “extended response questions” where students are required to explain their answer. Though extended response questions are purported to measure students’ “understanding of the mathematics concepts and/or procedures embodied in the task”
(Florida Department of Education, 2004), an examination of these types of questions reveal that they are more accurately a measure of students’ ability to express through language why their answer was right, not how they arrived at it (see Figure 1.2). Math word problems on the FCAT are used as a means of gauging mathematical competency, where points are given for the final product without regard to process. The drawback to this product-focus is that while the answer is a sufficient indicator of accuracy, it ignores students’ processing. There may actually be broad variability in skill among students with incorrect answers depending on the problem-solving phase with which they struggled.

![Pattern for Multiplying and Adding Table](image)

Figure 1.2. Example of FCAT Mathematics extended response question (Florida Department of Education, 2004).

Problem representation is critical to successful problem solving; Silver’s (2000) analysis of the NCTM standards cited representing a problem as a necessary prerequisite to deep understanding. Accordingly, the focus of this study is on the accuracy of student performance in the first two phases of the problem-solving process. Both paraphrasing
(i.e., problem translation) and visually representing (i.e., problem integration) the problem provide concrete evidence of how students conceptualize what they read. Misconceptions that are consistent within and/or across groups can thus be analyzed.

**Paraphrasing.** Problem translation in Mayer’s (1985) problem-solving model involves transforming the words of the problem into an understandable form. It is a step that is essential for eventual problem-solving success, yet research on paraphrasing in mathematics is almost nonexistent. The argument for paraphrasing as an effective comprehension strategy is much more developed in the field of reading (e.g., Ellis & Graves, 1990; Gillam, Fargo, & Robertson, 2009; Hoyt, 1999; Schumaker, Denton, & Deshler, 1984). However, some instructional interventions in mathematics have included paraphrasing as a strategy for comprehending word problems. For example, in *Solve It!* (Montague, 2003) students are instructed to paraphrase the math word problem in order to monitor their understanding of the text. Unfortunately, students’ ability to carry out this strategy was not explicitly measured in the *Solve It!* research; it is therefore impossible to determine the impact of students’ comprehension of word problems on their ability to solve them.

While paraphrasing has been used in the fields of both reading and math as a comprehension strategy, it is also used in the research literature for another purpose, that is, to measure comprehension (Nettles, 2006). In the reading research, paraphrasing as a measure of student understanding is often recorded as a percentage of total information provided (e.g., Hagaman & Reid, 2008). This is because most often in reading, the more information a student is able to accurately paraphrase, the better their comprehension. Therefore, a simple percentage of total information provided is a legitimate means of
calculating student understanding. Often called retellings, this method offers a more thorough depiction of student understanding than questioning probes (Nettles, 2006). In math problem solving, however, paraphrasing as a measure of comprehension is assessed somewhat differently. Because problems may contain irrelevant information, students must be more selective in choosing the parts to paraphrase. Thus, relevancy of information within the word problem becomes an additional feature to be considered in math that is generally ignored in the reading literature. At present, there is a small but important body of mathematical research on students’ ability to discriminate between relevant and irrelevant information (e.g., Censabella & Noel, 2008; Englert, Cullata, & Horn, 1987; Fuchs & Fuchs, 2002; Littlefield & Rieser, 1993; Parmar, Cawley, & Frazita, 1996; Passolunghi et al., 2005). Students with LD have demonstrated deficits in the ability to suppress irrelevant information, particularly when it is numerical (Passolunghi et al., 2005). Measuring students’ ability to paraphrase when irrelevant information is included will clarify some potential hindrances to comprehension (and ultimately solution accuracy) for students with LD and low-achieving students.

**Visual representation.** In comparison to the relatively small body of research on the different aspects of paraphrasing, visual representation has received greater attention. Research on this construct and how it relates to problem solving has been conducted for 30 years. Current research operationalizes visual representation as “the construction and formation of internal images (e.g., mental images) and/or external images” (van Garderen & Montague, 2003, p. 246). Research has further divided visual representation into two categories: pictorial representations, which are primarily drawings of objects, and schematic representations, which are diagrams representing the spatial relationships
among problem parts (Hegarty & Kozhevnikov, 1999). Analyses of students’ visual representations have shown that successful problem solvers generally produce more schematic representations and thus have greater success solving problems, whereas poor problem solvers (including students with LD) more often rely on pictorial representations leading to inaccurate solutions (Guoliang, 2003; Hegarty & Kozhevnikov, 1999; van Garderen & Montague, 2003).

The types of visual representations used by students of differing abilities have thus been corroborated through several studies, but the current literature contains two limitations. First, it theorizes that visual representations can be placed in two discrete categories; however, it is more accurate to say that they exist along a continuum of accuracy in representing the problem information. It is possible for representations to have features of both groups, and thus not fit clearly into either category. By employing the idea of a continuum instead, completely pictorial representations would be on one end and completely accurate schematic representations would represent the opposite end. The more visual representations include appropriate relational and numerical components, the closer they would fall on the schematic end of the continuum. Conceptualizing visual representations on a continuous scale permits qualitative distinctions to emerge and should provide more accurate associations between representation and solution accuracy. The concept of continuum as opposed to category also addresses the second limitation of previous research on visual representation. That is, the current literature lacks a qualitative distinction within the category of schematic representation. The research incorrectly implies that all representations of word problems identified as schematic are equivalent. However, it is reasonable that representations that use correct numbers but are
relationally incorrect would lead to an inaccurate solution path and thus an incorrect answer. The use of a continuum would distinguish between theoretically schematic but incomplete (or inaccurate) representations and numerically and relationally accurate representations. Thus by utilizing a continuous scale, this study attempts to address these shortcomings in the research.

In summary, effective problem solving begins with the problem representation phase. Comprehension of the problem must logically precede the ability to accurately depict a visual representation of the problem parts, as students cannot be expected to correctly represent the relationships of the relevant information in a problem that they do not understand. These two phases of the math problem-solving process are essential as “integration of the relevant information into a coherent structure allows a correct and complete mental representation of the problem” (Passolunghi & Pazzaglia, 2005, p. 259). Math research still lacks clarification in exactly where students struggle in their ability to paraphrase and visually represent word problems. Additionally, processing differences among ability groups need further investigation. Because the performance of students with LD and low-achieving students appears to be similar, it is assumed that the processes they use are similar; empirical evidence to substantiate that assumption, however, is lacking. The following section delineates the way in which LD are defined, in both the law and the research, in order to clarify the way in which the two groups are differentiated theoretically.
Students with Learning Disabilities

The Individuals with Disabilities Education Act (IDEA, 1990; 1997; 2004) defines a learning disability as:

a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia (United States Code 20 U.S.C. §1401 [30]).

While this definition has not changed in the 35 years since it was written into special education law, the way in which students with LD have been diagnosed has changed (Fletcher, Morris, & Lyon, 2003; Peterson & Shinn, 2002). There are two broad paths that can lead to an LD diagnosis: a) the intra-individual differences model, which has been the traditional form of identification; and b) the problem-solving model, which was included in the law with the reauthorization of IDEA (2004). The first model is generally based on the IQ-achievement discrepancy, where students are diagnosed with LD when there is a significant discrepancy (e.g., one standard deviation) between ability and achievement. In other words, students with LD demonstrate “unexpected underachievement” given their ability (Fletcher et al., 2003).

Though the federal law stipulates a severe discrepancy as necessary to the diagnosis of LD, it has left to the individual states the task of operationalizing “severe.” Therefore based on the way in which each state has defined it, a student diagnosed with LD in one state (or district, as some states have allowed their individual districts to determine the criteria) would not necessarily have that diagnosis in another, and vice versa (Peterson & Shinn, 2002; Gerber & Semmel, 1984). This arbitrariness of diagnostic
standards has had deleterious effects on research involving LD because, like states, researchers have used varying definitions for discrepancy. Thus broad generalizations based on findings in LD research cannot be confidently made from studies that have used different definitions to describe what should be the same population. Even with experienced researchers, inappropriate generalizations have occurred in both descriptive reviews and meta-analyses (e.g., Fuchs, Fuchs, Mathes, & Lipsey, 2000). This issue has, unfortunately, undermined high-quality research efforts.

Because of this and other problems in the field (e.g., the increasing number of LD identifications disproportional to the general population growth) the validity of the field itself has been brought into question (Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2002). Additionally, there have been several criticisms above and beyond definitional inconsistencies, including the IQ-discrepancy model’s “wait to fail” characteristic (Seigel, 2003), its instability over time (Francis et al., 2005), and its costliness of resources (MacMillan & Siperstein, 2002). The strength of the IQ-discrepancy model has always been in its theoretical foundation of psychometric methodology, but one of the most compelling arguments against this model has centered on this very strength; opposition criticizes the model for its inability to distinguish between the performance of IQ-discrepant students and students whose ability and achievement scores are consistently low (i.e., low achievers) (Fletcher et al., 2003). Thus in 2004 policymakers provided districts with the option of identifying students based on a lack of response to effective, high-quality instruction. Ultimately, the Response to Intervention (RTI) model is purported to be a “movement away from ‘test and treat’ models to ‘treat and test’ models” (Fletcher, Coulter, Reschly, & Vaughn, 2004, p. 309); struggling students are
provided tiered instruction based on responsiveness, where each tier provides increasingly intensive, individualized instruction until students either experience eventual success or are diagnosed with LD. This new model of identification of LD assumes that students ultimately diagnosed with LD will not respond to the tiered instruction to the same degree that low-achieving students would. In this way, RTI attempts to resolve the previously discussed limitations by providing early intervention, only identifying the most severe cases of LD (which will decrease the overall LD prevalence rate and thus decrease costs), and redefining the definition in a way that relates directly to classroom practices.

While there are clear differences between the IQ discrepancy model and RTI, for both models it remains that districts identifying students with LD need only diagnose its existence as a global disability. In research, however, a general consensus has supported the subdivision of students with LD into specific subgroups based on area of difficulty. In this study, students with LD with difficulties in math (as indicated by poor achievement on state assessments) will be identified as math LD as consistent with previous literature (Andersson, 2008; Fuchs & Fuchs, 2002; Hanich et al., 2001; Jordan & Montani, 1997).

**Math LD.** Since the NCTM recommendations there has been a moderate increase in mathematics research, with particular attention paid to students with LD (Andersson, 2008). Mathematics education is taught in a hierarchical way; that is, basic foundational skills are taught first, and then those skills are built upon with increasingly more complex concepts (e.g., fractions, problem solving, and scientific notation). Difficulties in an advanced skill could be attributable either to misunderstanding of that skill or to the various underlying skills necessary for its execution (Geary, Hoard, Byrd-Craven,
Nugent, & Numtee, 2007). Thus, young students with circumscribed math difficulties develop into older students with more generalized problems due to the confounding effects of a hierarchical curriculum. In an attempt to extract the specific math skills deficient in students with LD, several researchers have compared this population in the early elementary years to their typically-achieving peers on the main components that comprise mathematical proficiency. Though much of the research focuses on the younger years, abilities and deficits of students with LD in the mathematical domains of computation and problem solving at varying developmental stages are summarized below.

**Computation.** Computational skills are usually measured according to three factors: accuracy, speed, and strategy use, such as counting on or automatic retrieval (Mabbott & Bisanz, 2008). Students in early elementary grades at risk for math LD demonstrate a measurable weakness in their ability to automatically retrieve answers to number facts (Jordan & Hanich, 2000). Into the middle elementary grades, students with math LD are slower to produce answers, and the answers they do produce are more often wrong as compared to their typically-achieving peers. At this age, students with math LD maintain the use of immature counting procedures such as using their fingers or counting all the numbers in an addition equation instead of counting on from the greater number (Andersson, 2008; Gonzalez & Espinel, 2002). Even in the late elementary grades students rely on these immature strategies, and they continue to exhibit less efficiency in retrieving number facts (Geary, Hamson, & Hoard, 2000). Mabbott and Bisanz’s (2008) comparison among students with math LD, age-matched low-achieving students, and ability-matched younger students showed that 6th-grade students with math LD
performed most like ability-matched younger students on accuracy, speed, and strategy use during computation measures. This contrasts the development of typically-achieving middle school students, who should have most number facts memorized, be able to retrieve them quickly, and demonstrate mature strategies (e.g., derivation rules) when they are unable to automatically recall them. Hartman (2007) compared middle school students with LD to low-achieving students and found that the low-achieving students had more trouble with computational accuracy than students with LD. Though not readily apparent, Hartman’s study is congruent with previously discussed results because she did not impose time restrictions on the computation tasks. Thus, for students with LD, computational problems may be most evident when math fluency is measured (i.e., speed as well as accuracy), whereas LA students are consistently weak in computational skills.

**Problem solving.** Following the developmental progression in problem solving of students with math LD is not as straightforward as with computation because the characteristics of problem solving become increasingly more complex. Factors such as the inclusion of all four operations, multiple steps, the addition of irrelevant information, and the placement of the unknown affect the difficulty of the problems and complicate across-grade comparisons. Nonetheless, the literature has reported characteristics of students with math LD that distinguish them from their normally-achieving peers at different developmental levels, providing a snapshot of difficulties at each age. In the early elementary grades, research in math problem solving is focused on addition and subtraction of one-step problems. Jordan and Montani (1997) initially argued that students with math LD struggled with working memory and fact retrieval and not necessarily problem conceptualization because their research showed that they performed
similarly to typically-achieving students when solving problems in untimed conditions, but struggled in the timed condition. However, later research by Jordan and Hanich (2000) cited problem simplicity as a limitation of their prior research, a factor which masked existing deficits in students with math LD under both conditions. Using more complex problems (i.e., putting the unknown at the start of the problem), they found that students in the second grade at risk for math LD performed more poorly than average achievers in all facets measured. The literature consistently reports weaknesses in problem-solving performance by students with math LD through the upper elementary years. Andersson (2008) found significant weaknesses in the problem solving of students with math LD in the third and fourth grade, even after statistically controlling for calculation principles, arithmetic, and multi-digit computation. Upper grades show the same findings; in middle school, students with LD are less able to predict the appropriate operation to use and report using fewer representational strategies (Montague & Applegate, 1993). Metacognitive differences during problem solving have also been observed among students with LD as compared to their typically-achieving peers; specifically, students with LD are less likely to modify their metacognitive activity based on the difficulty of the word problem whereas typical achievers increase their metacognition commensurate with increasing problem difficulty (Sweeney, Krawec, & Montague, in press). In contrast, students successful in math problem solving have a developed repertoire of strategies, effective self-regulating behavior, greater attentional focus, and an ability to generalize their learning (Montague, 2009).
Summary

Students with LD have been diagnosed using essentially the same definition since 1975. Historically, the IQ-discrepancy model has provided evidence for LD through “severe” differences between IQ and achievement scores. However, in an attempt to circumvent the limitations of this model, in 2004 IDEA was revised to permit states to utilize RTI as an alternative method for LD identification. One major problem with the traditional model is the inability to distinguish practically between the performance of students with LD and LA students. Research in both reading and math has generally corroborated a lack of differences, but generalizations about LD characteristics must be made with caution due to varied operationalizations of the construct. Particularly in the younger grades, researchers have analyzed subtypes of LD based on difficulties in math, reading, or both. Students with math LD demonstrate weaknesses in automatic retrieval of number facts, are more likely to use immature strategies, struggle with determining the appropriate operations, and activate their metacognitive strategies ineffectively.

Statement of the Problem

Research has attempted to clarify the characteristics that affect successful math problem solving, including the presence of a disability (Fuchs et al., 2000; Geary, 2003; Jitendra et al., 1998), knowledge of strategies (Montague, 1991), the type of problem (Garcia, Jimenez, & Hess, 2006), irrelevant information (Censabella & Noel, 2008; Passolunghi et al., 2005), and the ability to visually represent the problem (Booth & Thomas, 2000; Hegarty & Kozhevnikov, 1999). However, despite the growing body of evidence in the math problem-solving literature, there are some key areas that have been underexplored. First, there is a dearth of literature explicating the problem-solving
performance of low-achieving students in relation to their peers, including those with
disabilities. More common are comparisons between typical achievers and high
achievers, and to a lesser extent, students with LD (e.g., van Garderen & Montague,
2003). Few studies directly compared students with LD and LA students; in fact, research
has justified grouping them together due to similar performance outcomes (e.g., Garcia et
al., 2006). This is problematic in that it assumes that because student performance
between the ability groups is the same, so too are the processes leading to that
performance. One major goal of this study is to investigate that assumption. Another
issue is that a substantial amount of the research in mathematics and LD focuses on
lower-level skills such as fact retrieval and counting strategies as opposed to higher-order
processing skills (Geary, 2003). Of those studies that do focus on problem solving, the
investigation of differences among abilities often concentrates on younger grades, where
the single-step problems consist solely of addition or subtraction (e.g., Andersson, 2008;
Garcia et al., 2006; Jordan & Hanich, 2000). While valuable information about the
development of early problem-solving skills is gleaned from this research, an abundance
of information is ignored by the exclusion of older students who should already have
developed the reasoning, processing, and metacognitive skills necessary to solve
complex, multi-step problems that include all four operations. One noteworthy exception
to the focus on younger students is the research of Montague and colleagues (e.g.,
Montague & Applegate, 1993; Montague & Bos, 1986; Montague et al., 2000) where
middle school students are the primary participants. This research has played a major role
in explaining the problem-solving characteristics of older students with LD, but more research is needed to fully understand the multiple aspects of LD across ages (Deshler, 2002).

This study attempts to build on this body of research by including low-achieving students and measuring specific phases within the problem-solving process. The decision to use eighth-grade students is based on developmental research that shows that the average 12- to 13-year-old student should be cognitively able to visualize, represent, and internalize information (Kass & Maddux, 2005). Consequently, the information gained should not be confounded by undeveloped or still developing cognitive processes. As previously discussed, research by Montague and her colleagues has established the processes necessary for successful problem solving, but there is an absence of literature on students’ proficiency with these processes. Thus, the purpose of this study is twofold. First, it attempts to add to the literature on ability differences in math problem solving by directly comparing students with LD, LA, and AA students. This study measures not only student accuracy in solving math word problems but also measures the ability to represent math problems by measuring in detail students’ ability to represent the word problems through paraphrasing and visual representation. At present, there is little information on students’ ability to paraphrase math word problems and how that correlates with successful problem completion. Second, the study extends prior research in the area of visual representation (Hegarty & Kozhevnikov, 1999; van Garderen & Montague, 2003) by assessing the accuracy of visual representations along a continuum, and by focusing on older students.
Chapter 2

Review of the Literature

Student performance in mathematical problem solving has received more attention in the past several years due to a shift in the curricular focus of instruction. Though higher order, concept-rich teaching has begun to replace traditional didactic instruction in the U.S., research continues to document the difficulties students encounter while solving math word problems. This chapter will first examine the literature on direct comparisons between the academic performance of students with learning disabilities (LD) and low-achieving (LA) students both in reading and in mathematics. Next, the definition of math LD and the problem-solving characteristics associated with students with math LD throughout their developmental years are described. Third, research on paraphrasing will be reviewed. This section will review the history of paraphrasing as it relates to reading comprehension before reviewing the brief literature on paraphrasing in mathematics. Finally, what is known about student abilities in visual representation of word problems will be presented. Specifically, this section will focus on the definition of visual representation and the two types identified in the literature. It should become evident that there is a decided lack of research in paraphrasing in math problem solving as well as limitations in the visual representation literature that make the need for more research apparent.

Students with LD and Low-Achieving Students

The definition of LD has never been free of controversy nor contention. One of the longstanding criticisms has centered on the distinctions (or lack thereof) between students with diagnosed LD and students who are LA but do not qualify for services
based on federal and/or state stipulations. The overwhelming majority of research studies have found no valid differences between the academic performance of students with LD and LA students (Fletcher, 1995). In a seminal study, Ysseldyke, Algozzine, Shinn, and McGue (1982) administered a full battery of psychoeducational assessments to a group of fourth-grade students with district-identified LD (n=50) and a group of fourth-grade LA students who scored at or below the 25th percentile on the statewide assessment (ITBS) but were not identified as having LD (n=49). No student in the LD group had been diagnosed more than six months prior to the study, a factor intended to eliminate any treatment effects associated with special education placement. The test battery included achievement, IQ, self-concept, and behavioral measures; altogether, eight different instruments were used in part or in whole. The purpose of the study was: a) to determine if consistent differences existed between groups to justify the mode of identification then in place, and b) to evaluate the consistency with which districts were adhering to the federal mandate for the diagnosis of LD. Results showed that up to 40% of the students were diagnosed incorrectly when adhering strictly to the federal definition of ability-achievement discrepancy and underachievement. Further, the authors found that 96% of student scores across groups were within a common range on the 49 different measures administered, showing no support for differences in performance. The authors acknowledged that their findings could be interpreted in two ways: one, that students with LD were being “missed” during identification and thus being denied needed services, or two, that students without LD who were low achievers were being inappropriately labeled as having learning problems. Either way, the results confirm that across the broad range of psychometric measures used, these two groups were more similar than different.
Due to the emerging evidence in the field that students with LD and LA students demonstrate similar performance, in the 1990s learning disability centers were funded by the National Institute of Child Health and Development (NICHD) in order to clarify definitional and operational issues with the LD construct. Four major research studies have been published examining the validity of what they have termed the “two-group hypothesis”: two were published based on data from the Connecticut Longitudinal Study (Fletcher et al., 1994; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996), and the other two used researcher-collected data (Foorman, Francis, & Fletcher, 1995; Stanovich & Siegel, 1994). All four studies failed to show differences between LA students and students with LD in reading; thus all authors concluded that students with an IQ-achievement discrepancy and LA students could not be differentiated with respect to variables related to reading proficiency, reading development over time, phonological processing growth, and reading subskills, respectively.

**Reading.** Research in reading has generally mirrored the findings of Ysseldyke et al. (1982) and the investigations linked to NICHD. O’Malley, Francis, Foorman, Fletcher, and Swank (2002) looked at performance on reading-related skills such as letter-sound knowledge, phonemic awareness, rapid serial naming, and word reading, among others. The authors were not interested in conducting cross-sectional comparisons that only examine one time point; they described the literature as having established the no-differences phenomenon, but still lacking evidence that this similarity persisted developmentally. In order to more fully support this view, the authors conducted a two-year longitudinal study, after which they concluded that there were no differential improvements in five of the eight skill areas between students with LD based on an IQ-
achievement discrepancy and low-achieving students. Similarly, in his dissertation,
Vining (2008) examined reading rate over time and came to the conclusion that there
were no uniform and meaningful differences between students with LD and LA students
when the former group was identified using the IQ-achievement discrepancy.

Fuchs et al. (2002) gathered all the published and unpublished studies comparing
LD and LA groups on their reading achievement between 1975 and 1996 and conducted a
meta-analysis to determine whether the two groups were distinguishable. Their findings
contradicted the majority of research on the topic; they found that the reading
achievement of students with LD differed “dramatically” (p. 755) from their LA peers
with a mean effect size of 0.61 standard deviation units across all studies. Yet while the
results appear to support a distinction between the two groups, methodological and
construct-related errors limit the generalizability of the study’s findings; in fact, this
study is representative of how research can actually impede the advancement of the LD
field. First, the authors failed to set explicit LD requirements for studies to be included.
Here, any studies that had a “learning disabilities” group, a “low-achieving” group, and
separate data for each were acceptable. As previously discussed, the various ways in
which LD has been defined (e.g., severe underachievement versus discrepancy between
expected and actual achievement) influences the results obtained. Fuchs et al.’s (2002)
undefined “LD” population makes the findings unable to be interpreted with confidence.
Further, methodological flaws have drawn questions about the accuracy of the reported
effect sizes. Deshler’s (2002) critique of the study points to the authors’ results as
contingent upon correlations between the two groups, which were estimated most of the
time. Also, different formulas were used to calculate the effect size of studies, raising the
question of comparability across studies. Finally, the authors’ (understandable) restriction to one academic domain should limit the results to that domain; thus Fuchs et al.’s generalization to broad LD-LA differences should be interpreted with caution.

**Mathematics.** Though the literature on differences in LD and LA groups in the area of mathematics is much less substantial, there are enough studies to generally support a no-differences argument. General math ability between the two groups was investigated by Mabbott and Bisanz (2008), who compared middle school students with diagnosed LD in math to typically-achieving students, low-achieving age-matched students, and younger ability-matched students on different measures of mathematical ability. The LD group was district-diagnosed as having a disability as reflected by average intelligence and low achievement, and further identified as having a math disability by a score below the 35th percentile on a standardized achievement test (WRAT-III; Wilkinson, 1993). The same measure was used to classify the other students as LA (below the 35th percentile) and AA (at or above the 35th percentile). Students were compared on measures of computational skills, working memory, and conceptual knowledge. It is important to note, however, that the computational measure used time as a factor; thus, it is more accurate to say that math fluency was measured as opposed to computational ability. Results showed that while distinctions were found between the typically-achieving students and the other groups, no statistically significant differences were found on the three measures between the age-matched LA and LD groups, though the working memory measure approached significance.

In problem solving, performance-based differences between LA and LD groups have also been unsubstantiated. Gonzalez and Espinel (2002) looked specifically at the
differences between early elementary students with LD and low-achieving students, including typically-achieving students as a reference group. They looked at the counting, modeling, and number fact strategy choices of seven- to nine-year-olds while solving one-step addition and subtraction word problems. Groups were formed according to the federal definition; that is, group membership was based on performance below the 25th percentile on a standardized mathematics test and an IQ-discrepancy (15+ points; LD), performance below the 25th percentile on a standardized math test but no discrepancy (<15 point difference; LA), and performance above the 30th percentile on a standardized mathematics test (typical achievers). All students had an IQ score greater than 80. Results showed significant differences between the typically-achieving students and the other two groups on performance, counting strategies, and number fact strategies. No differences were found among groups on modeling strategies. The authors thus concluded that working memory, specifically needed for the computational processes of problem solving, appears to be a deficit in both LA students and students with LD and acts to hinder their ability to correctly solve mathematical word problems. Further, the authors suggest that the IQ-discrepancy model may be inappropriate for discriminating between students with LD and LA students due to this lack of difference. However, the authors’ conclusions about working memory were speculative, as the study had no measure that addressed the construct.

In an older population of students with LD and LA students, similar non-differences in math problem solving have been found. These comparable findings were reported by Montague and her colleagues (Montague, Enders, & Dietz, 2009) in an investigation of the effects of the cognitive strategy instructional intervention Solve It!
(Montague, 2003) on middle school students. On the pretest, students with LD and LA students showed no statistically significant differences on problem-solving performance on one-step, two-step, and three-step problems, as assessed by a curriculum-based measure (Montague et al., 2009). Both groups, as expected, performed less well than average-achieving students.

Only one study in math was found that questions the consensus in the literature on the nature of LD-LA differences. Documenting the struggle of students with LD on mathematical problem-solving tasks, Hartman’s (2007) study analyzed not only performance measures but also measures of processes. Her sample of students with LD was selected using the Learning Disabilities Diagnostic Inventory (LDDI; Hammill & Bryant, 1998). This scale uses general education teacher ratings to determine whether students display characteristics indicative of LD. Her sample of low-achieving students included those who displayed similar poor achievement in math as measured by the Iowa Test of Basic Skills, but were not likely to have LD as indicated by the LDDI. Through error analysis, Hartman was able to isolate differences between students with math LD and students with low achievement in mathematics. Scores for each group were generated from the Mathematical Word Problem Error Analysis adapted from Muoneke (2001), which classifies student problem-solving behavior (e.g., defective algorithm, grouping error, random choice, key words). Student problem-solving behavior was measured in four ways: think alouds during math problem solving where students were “asked to describe what they were thinking and how they were deriving their answers as they solved the mathematical word problems” (p. 105), written student work, interviews where students explained their written work, and researcher observations. Results showed that
students with LD had conceptual deficiencies in analysis, reasoning, and abstract thinking, while low-achieving non-LD students struggled more with the computational components of problem solving.

To summarize, the body of research across both reading and math appears to sustain the argument that there are no measurable differences between students with diagnosed LD and LA students, although that position is not unanimously supported. Studies such as Fuchs et al. (2002) which have found differences have received criticism for methodological and construct-related flaws. Hartman’s (2007) research also found differences both in performance and in processes used, but her use of a nontraditional identification measure for LD participants makes it difficult to situate her findings in the field. At present, the advancement of RTI has brought with it the no-differences assumption between IQ-discrepant students and low achievers. However, before this position is established as the consensus in the LD field, its veracity must be tested. At present, there is a common thread that binds together all the research espousing no differences between groups: the outcome measures are performance-based. There remains a shortage of math research examining the processes that students use to arrive at their problem-solving solutions, and whether those processes differ as a function of disability. The available research on the first two phases of Mayer’s (1985) problem-solving model, that is paraphrasing and visual representation, will be discussed below.

Paraphrasing

Before paraphrasing was ever considered in math research, it was a prominent issue in the study of reading comprehension. Schumaker et al. (1984) reported improved text comprehension through paraphrasing when students were required to follow three
steps: read the paragraph, identify the main idea and key details, and put it in your own words. These steps are, in essence, the definition of paraphrasing; in fact, the dictionary defines paraphrasing as “a restatement of a text, passage, or work, giving the meaning in another form” (Merriam-Webster, 2000). When paraphrasing is linked to comprehension, it serves two distinct purposes in research. Its first role, as described above, is as a strategy utilized to improve comprehension. Paraphrasing thus aids the reader in gaining understanding by rewording the text into a more familiar form. Consequently, much of the research studies that center on paraphrasing as a strategy provide students with interventions that teach the strategy and measure comprehension as the outcome of interest, often in the form of question-response (e.g., Ellis & Graves, 1990; Schumaker et al., 1984). The second form of paraphrasing that appears in the reading literature is as a measure of comprehension. This literature base uses paraphrasing as a way of assessing students’ understanding of text (e.g., Carver, 1975; Gillam et al., 2009). When used in this way, paraphrasing is more often referred to as retelling or story recall. For example, Gillam and colleagues (2009) used a passage recall measure to determine whether differences existed in the comprehension of text between students with and without language impairments in the fourth grade. The authors read passages of about 150 words in length and 15 essential elements were selected a priori. Students were asked to recall the passage and their recall verbalizations were coded as paraphrases, exact repetitions, and inferences. Each code was further divided into accurate versus inaccurate recalls. Results showed no differences between groups on the number of exact repetitions recalled accurately, but that significant differences existed favoring typically-achieving students in the number of accurately paraphrased components of the passage. In this
study, paraphrasing was used as a measure of the degree to which students processed what was read to them, and was shown to be an accurate indicator of differences between groups.

While some studies focus on paraphrasing as a strategy and others use it as a measure, still other studies have combined the two purposes and incorporated paraphrasing into the research as both a strategy to improve and a measure to assess comprehension. In this way, students were taught a method of paraphrasing as a comprehension strategy, and then assessed on their ability to comprehend by measuring the extent to which they could paraphrase all the information (Donahue & Pidek, 1993; Hagaman & Reid, 2008; Laing, 2000). Thus, whether it is used as a strategy, a measure, or a combination of both, the reading literature has demonstrated the value of incorporating paraphrasing into comprehension research.

In math research, only a few studies have considered paraphrasing either as a strategy or as a measure of comprehension. Montague (2003) included paraphrasing as a step in her cognitive strategy instruction of math problem solving after previous research demonstrated that it was a strategy used by successful problem solvers (Montague & Bos, 1986). This use of paraphrasing aligns with the first purpose based on the reading literature: as a comprehension strategy. As noted previously, however, paraphrasing in math problem solving is fundamentally different from the paraphrasing in reading. That is, students solving math word problems must not only paraphrase the given content, but also sometimes critically analyze it for relevance. Math research focusing on irrelevant
information has confirmed the difficulty students of varying abilities have with this added complexity in most word problems (Kouba et al., 1988; Littlefield & Rieser, 1993; Parmar et al., 1996; Passolunghi et al., 2005).

Passolunghi and colleagues (2005) found that the type of information, either linguistic (i.e., content information unnecessary for the solution path) or numerical, differentially complicated the problem-solving process for students depending on their disability. The authors’ focus was to determine how fourth-grade students with either ADHD or math LD were affected by the two types of irrelevant information. It should be noted that the authors’ definition of ADHD was consistent with the medical diagnosis as reflected in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 1994), while their definition of LD would more appropriately be termed severe underachievement. That is, students with math LD performed 2+ standard deviations below the Italian mean on the Wide Range Achievement Test (WRAT3; Jastak & Wilkinson, 1993) Arithmetic subtest and were identified by their math teachers as struggling with computation and word problems. In the context of the other studies reviewed, the results of Passolunghi et al. (2005) will be generalized to both students with LD and low-achieving students, as the author’s differentiation between the two groups was not clear. The authors divided irrelevant information into two discrete categories based on the hypothesis that the type of information causes problems for different reasons; that is, irrelevant linguistic information increases the load on working memory and processing of information, whereas irrelevant numerical information confuses the solver and may elicit extra or incorrect calculations. To test their research question, the authors examined three scores
across the eight total word problems: a) the amount of relevant information recalled, b) the amount of irrelevant linguistic information recalled, and c) the amount of irrelevant numerical information recalled. Four problems contained four pieces each of irrelevant linguistic information added by the researchers to the original problems, and four problems contained four pieces each of irrelevant numerical information, also added by the researchers. Thus, they were able to separate relevant from irrelevant by counting all original components as relevant and any added information as irrelevant. Results showed that students with math LD were found to be significantly more affected by irrelevant numerical information than either students with ADHD or typical achievers. It is important to note that students asked to recall the important information in each problem were doing so from memory; that is, their ability to paraphrase the important parts of the problem was limited by their working memory. The authors intended this, and in fact measured working memory with the Digit Span task (Wechsler Adult Intelligence Scale–Revised; Wechsler, 1994). However, none of their analyses built working memory in as a covariate or as an independent variable to be measured.

Littlefield and Rieser (1993) attempted to qualify the difficulty associated with irrelevant information in math word problems by manipulating the level of similarity the irrelevant information shared with the problem’s relevant information. Features of similarity included the agent (e.g., Michelle), the action (e.g., purchasing), and the unit (e.g., earrings), and problems were formatted in one of four conditions: no irrelevant information, no similarity (i.e., irrelevant information where none of the three features matched), low similarity (i.e., one of three features matched), or high similarity (i.e., two of three features matched). Their sample consisted of students who were identified as
either successful problem solvers or less successful problem solvers based on performance on a standardized math assessment and teacher ratings of student problem-solving ability. The authors’ predicted model of discrimination performance was validated in the study; the results showed that less successful problem solvers demonstrated a significant positive linear trend between error percentage and level of similarity, both for one- and two-step problems. Successful problem solvers only showed this trend for two-step problems. As expected, significant differences existed between the two ability groups across problem difficulty. One major limitation to this study that must be acknowledged is that the authors did not have students solve the word problems. They merely read each problem and highlighted the numbers they deemed relevant. Though Littlefield and Rieser assert that this strengthens the study by focusing solely on the discrimination phase of problem solving, their design fails to consider that students might modify their selections as they proceed through the problem-solving process. In essence, this study provides an important glimpse into the initial discrimination patterns of students of varying abilities, but lacks a more comprehensive understanding.

Similar to the findings reported by Littlefield and Rieser (1993) using more and less successful problem solvers, research focusing specifically on the discrimination abilities of students with LD has found that these students struggle more so than their non-disabled peers at differentiating between the relevant and irrelevant information (Passolunghi et al., 2005; Parmar et al., 1996). It is clear that the ability to accurately paraphrase the components of a problem necessary to the solution “does not simply involve the maintenance of given information, but it requires its control, that is, this information is examined for relevance, selected or inhibited (depending on its relevance),
integrated and used” (Passolunghi et al., 2005, p. 732). The second phase of the problem-solving process, integration of the problem into a visual representation, is greatly dependent upon the accuracy with which the problem is understood.

**Visual Representation**

Understanding a mathematical word problem is essential to solving it, and it is known that effective problem solvers use visualization in order to comprehend the problem (van Garderen & Montague, 2003). However, research on the benefits of creating visual representations to solve mathematical word problems has been equivocal, most likely because of the wide range of definitions used to define the construct. Early in the literature, the definitions were vague and operationally problematic. For example, Lean and Clements (1981) defined visual imagery as “imagery which occurs as a picture in 'the mind's eye'” (p. 268). They asked university students, after solving a mathematical word problem, to describe how they solved it. Researchers then categorized their methods as either primarily visual or primarily verbal-logical. The findings indicated that visual students performed less well than verbal-logical students on measures of mathematical problem solving. However, issues beyond the definition of visualization have limited the internal validity of these findings. The primary issue lies in the problem-solving measure used. Suwarsono’s (1982) Mathematical Processing Instrument was developed for middle school students in Australia. Therefore, it is likely that most of the university students in Lean and Clements’s study did not require the cognitive support of visualization for word problems five years below their level. Those who did report using visualization to help them solve the word problems probably had poor mathematical skills in general; thus, the
correlation between visual learners and poor performance was likely a function of the measure’s low ceiling and only served to mask more basic differences in ability.

Presmeg’s (1986) slightly more detailed definition characterized a visual image as “a mental scheme depicting visual or spatial information” (p. 297) and seemed to acknowledge the shortcomings of the definition by Lean and Clements (1981) by making clear that the definition “allows for the possibility that verbal, numerical or mathematical symbols may be arranged spatially” (p. 297). Presmeg qualitatively analyzed the problem-solving styles of seven 12th grade students identified as ‘star pupils’ by their mathematics teachers. Findings showed that almost none of the participants were visualizers. However, the authors listed several factors that may have influenced the findings, such as an instructional focus on non-visual proofs and numerical explanations, the additional time required to form visual representations, and the lack of value some teachers place on the use of visualization. Placing this research in the context of other studies, however, points to an alternative consideration; that is, students who are ‘stars’ – at the top of their mathematics classes – may not require the assistance of a visual representation to the same degree that less able students may.

**Categories of visual representation.** Hegarty and Kozhevnikov (1999) also examined discrete categorical divisions of student methods for solving word problems, but from an alternative perspective. Their unit of analysis was not the student but the representation; that is, rather than identifying students as using a certain cognitive style (e.g., verbal-logical), the authors analyzed student drawings generated while solving mathematical word problems. Those drawings were then categorized as either *pictorial,* “constructing vivid and detailed visual images,” or *schematic,* “representing the spatial
relationships between objects and imagining spatial transformations” (p. 685). The authors hypothesized that schematic representations would be positively correlated with successful problem solving, and pictorial representations would be negatively linked to successful problem solving. The results demonstrated just that. This study ties together previous literature by categorizing the general strategy of visual representation into distinct subtypes. The definition used by Lean and Clements (1981) encompasses what Hegarty and Kozhevnikov refer to as pictorial representations, and the definition put forth by Presmeg (1986) describes the schematic subtype of visual representation.

The distinction Hegarty and Kozhevnikov (1999) made between pictorial and schematic representations has shaped the way in which problem representation has since been studied. Van Garderen and Montague (2003) compared sixth-grade students with LD, average-achieving students, and gifted students on their use of visual representations. Because students were not prompted to create a representation, two scores based on visual representations were generated: present versus not present, and pictorial versus schematic. The authors adhered to Hegarty and Kozhevnikov’s operational definitions of each type of representation. Results showed differences both in the use of visual representations, where gifted students used them significantly more than the other two groups, and in the type of representations, where students with LD produced significantly more pictorial representations. Success in math problem solving was found to be positively correlated with the use of schematic representations because this type of representation attends to the relationships among problem parts, whereas pictorial representations only address the visual appearance of objects in the problem (Hegarty & Kozhevnikov, 1999). Both Hegarty and Kozhevnikov (1999) and van Garderen and
Montague (2003) were interested in whether students voluntarily produced visual representations, justifying why students were not prompted to draw a diagram. This method allows for information on whether students view visualization as an effective strategy, but for both studies it greatly reduced the amount of representation data generated among groups. That is, Hegarty and Kozhevnikov found that across all students, a mean of 5.75 pictorial representations and 2.00 schematic representations were produced on the 14-problem measure. This means that students only visualized on roughly half the problems. Van Garderen and Montague (2003) show similar results, where students drew a mean of 4.23 pictorial representations and 3.88 schematic representations. Developmental research has indicated that students develop abstract thinking and visual manipulation skills between the ages of 11 and 14 (Kass & Maddux, 2005). Therefore, the sixth-grade students used in these studies were likely still developing their visual representation abilities, a factor that may help account for the relatively low frequency of visual representations. Additionally, students in both studies only answered about half the problems correctly: “less than half” in Hegarty and Kozhevnikov’s study (p. 687), and 36.9% in van Garderen and Montague’s study. Thus, difficulty level of problems appears to have factored in to students’ proficiency both in the production of a representation and in forming an accurate solution.

In summary, results from the literature have shown that the way in which a researcher defines the construct of visualization greatly affects the results obtained. The research on visual representation in relation to problem solving has clearly evolved over the past three decades, progressing from a vague, operationally inadequate definition to a theoretically based, research-supported conceptualization. This has enabled the
characteristics of effective problem solvers to be described more clearly. However, the
current research on visual representation has to this point neglected a major component of
analysis: accuracy of visualizations. Both Hegarty and Kozhevnikov (1999) and van
Garderen and Montague (2003) reported high levels of agreement in identifying
representations as either pictorial or schematic (.85 and .83, respectively). However this
begs the question, what can be interpreted about a fully schematic representation that uses
the wrong numerical information or misrepresents the relationship among problem parts?
The research has not considered accuracy of representation. Two decades ago, Lewis
(1989) reported that “the majority of students' errors on word problems are due to
misrepresentation of problem structure rather than due to computational errors” (p. 521).
Accurate representation of problem structure must clearly include the schematic
component as well as the correct numerical information and relationships. Thus the
present study, which will consider the accuracy of representations along a schematic
continuum, should provide a stronger predictive model from which to determine
relationships than those previously offered.

Overview

In general, the problem-solving literature has described the differing
performances of students with and without LD on problem-solving performance (e.g.,
Jitendra et al., 1998; Jordan & Montani, 1997; Montague & Applegate, 1993).
Additionally, some studies have examined the differences between these groups on visual
representation of mathematical word problems (Hegarty & Kozhevnikov, 1999; van
Garderen & Montague, 2003; Willis & Fuson, 1988). However, there is a paucity of
literature on differences in specific strategy use among groups during mathematical
problem solving when students with LD and low-achieving students are directly compared. In fact, only one study (Gonzalez & Espinel, 2002) looked specifically at these two groups, but with a population of second and third graders who were still developmentally immature in their strategy repertoires and therefore unlikely to fully demonstrate the potential areas of contrast. Interestingly, an abundance of research has analyzed differences between these two groups in the field of early reading (e.g., Fuchs et al., 2000). Findings suggest that no consistent and meaningful differences exist between students identified as LD and students who are low achieving when performance is the measured outcome. Studies in mathematics that have looked directly at differences between these groups have generally found no clear distinctions (e.g., Gonzalez & Espinel, 2002; Mabbott & Bisanz, 2008; McNear, 1990), although some research has reported differences when qualitatively analyzing the data (e.g., Hartman, 2007). After an exhaustive search, no studies were found that focused specifically on visual representation differences in accuracy between students with LD and low achievers. Further, only one study was found that considered the ability of students with LD/LA students to paraphrase as a measure of problem understanding (i.e., Passolunghi et al., 2005). According to Riley, Greeno, and Heller (1983), “the acquisition of problem-solving skills is primarily an improvement in children’s ability to represent the relationships among quantities described in the problem situation” (p. 161). We know that students have varying levels of problem-solving abilities, so an analysis of their ability to represent the relationships should provide an insight into exactly where their misunderstandings – and understandings – lie.
**Statement of Purpose**

The purpose of this study was to investigate the assumption that no valid differences exist between students with LD and LA students by analyzing their ability to solve math word problems. The majority of research to date has found no valid differences between students with LD and LA students based on performance outcomes. Therefore, this study took into consideration not only solution accuracy, but also the math problem-solving processes (i.e., paraphrasing and visual representation) used to arrive at that solution. Research on paraphrasing is almost nonexistent in the math literature, and research on visual representation has not yet considered accuracy of representation on a schematic continuum. No study has examined the relationship between these two processes and how they relate to overall problem-solving performance. Analyzing differences among students with LD, low-achieving students, and average-achieving students should provide the heretofore lacking information.

**Research Questions**

The following research questions were addressed:

1) What are the relationships among paraphrasing accuracy (i.e., relevant information, irrelevant linguistic information, and irrelevant numerical information), visual representation accuracy (i.e., relevant information, irrelevant linguistic information, and irrelevant numerical information), and problem solving accuracy?

2) What are the differences in paraphrasing, visual representation, and problem solving accuracy among students with LD, LA students, and AA students?
3) Are the effects of paraphrasing accuracy and visual representation accuracy on students’ problem-solving accuracy different among students with LD, low-achieving students, and average-achieving students?

**Hypotheses**

Based on the review of the problem-solving literature as well as the research on LD-LA differences, several hypotheses were proposed. First, it was expected that students’ ability to identify (through paraphrasing) and represent (through visual representation) the relevant information in a word problem would be positively correlated with problem-solving accuracy (Hegarty & Kozhevnikov, 1999; Passolunghi et al., 2005). Conversely, irrelevant information was expected to be negatively correlated with problem-solving accuracy based on the paraphrasing research (e.g., Littlefield & Rieser, 1993; Passolunghi et al., 2005). Though no studies were found that studied visual representation with irrelevant information, the expectation was that students’ inclusion of the irrelevant information in diagram form would also be negatively correlated to problem-solving accuracy. Derived from the findings of Passolunghi and colleagues (2005), it was hypothesized that irrelevant numerical information would produce a stronger negative correlation to problem-solving accuracy than irrelevant linguistic information because the inclusion of irrelevant numbers would likely lead to the wrong solution whereas the inclusion of irrelevant content would only increase the context (which for some would increase the cognitive load, but for others may provide a helpful contextualization of the problem).

It was hypothesized that there would be differences in problem-solving accuracy between average-achieving students and both students with LD and low-achieving
students, where average achievers would outperform the other two groups. It was also expected that no differences would exist between the latter two groups (i.e., students with LD and LA students). Average-achieving students were expected to be more accurate (i.e., have higher scores) in identifying and representing the relevant information and more discriminant (i.e., have lower scores) in excluding the irrelevant information, both in their paraphrasing and in their visual representations. Though no research currently exists that compares the problem representation abilities of students with LD and LA students, it was hypothesized that students with LD would have more difficulty discriminating between the relevant and irrelevant information, in both their paraphrasing and their visual representations, due to deficits in conceptualization and representation of information (Montague & Applegate, 1993; van Garderen & Montague, 2003).

In addition to the expected differences among groups on paraphrasing, visual representation, and problem-solving accuracy, it was hypothesized that there would be a relationship of varying strength between paraphrasing and visual representation and problem-solving accuracy depending on ability. However, in explaining the nature of the differences, two contradictory predictions seem viable. On the one hand, it could be predicted that AA students would demonstrate a stronger relationship between the problem-solving processes (i.e., paraphrasing relevant information and visual representing relevant information) and problem-solving accuracy since they can better comprehend the problem, create a representative schematic, and utilize that schematic to solve the problem. For students with LD and LA students, even with the inclusion of relevant information in a schematic representation, there is the possibility that comprehensive understanding of the problem is still limited, which is further exacerbated
by their poor ability to choose the correct operation (Montague & Applegate, 1993). Thus this hypothesis would support the finding that AA students would demonstrate a stronger relationship among the problem-solving processes and problem-solving accuracy than students with LD and LA students. On the other hand, AA students have, by definition, a stronger understanding of mathematical concepts and procedures, and therefore may not require as detailed a schematic representation in order to get the answer as those LA students or students with LD would require. Also, AA students may be more recursive in their problem-solving process, and so able to adjust their initial decision-making as they progress through each problem (a characteristic not measured in this study). This would lead to a weaker relationship between the problem-solving processes and solution accuracy as well as weaker relationships between the two processes for the AA group. In contrast, students with LD and LA students may lack this recursivity when problem-solving. Further, previous research has identified problem representation as a weakness of students with LD (van Garderen & Montague, 2003); thus, for students with LD who were able to create a schematic representation, the proverbial hurdle would be overcome, thus increasing the likelihood of problem-solving accuracy. This conjecture would support the finding of a stronger relationship among the DVs and problem-solving accuracy for students with LD (and possibly LA students) than for AA students.
Chapter 3

Methods

Participants

Eighty-four eighth-grade students in Miami-Dade County Public Schools (M-DCPS) participated in this study. The mean age of the participants was 14.8 years. The rationale for using middle school students was based on Kass and Maddux’s (2005) developmental theory of learning, which suggests that between the ages of eight and 11 students develop the ability to visualize abstract information and internalize concepts. Thus, by the eighth grade, students should be able to represent math word problems by translating the linguistic information and developing schematic visuals of the problem. Participating middle schools were drawn from a pool of 36 schools involved in a larger study funded by the U.S. Department of Education, Institute for Education Services, eProst ID 20070008-03. The title of the larger study was Improving Mathematics Performance of At Risk Students in Urban Middle Schools (Middle School Math: MSM Project), supervised by Marjorie Montague, Ph.D. (Principal Investigator). From the larger pool, four schools agreed to participate in the study.

Students represented three groups: students with learning disabilities (LD; \(n = 25\)), low-achieving (LA; \(n = 30\)) students, and average-achieving (AA; \(n = 29\)) students. Criteria for the LD group are district-identified classification of LD and a low level of 1 or 2 (range 1 to 5) on the previous year’s math FCAT. District criteria for LD classification includes: a) a disorder in one or more of the basic psychological processes involved in understanding or in using language; b) academic achievement that differs significantly from the student’s measured aptitude; c) a learning problem not primarily
ineffectiveness of research-based teaching strategies in the general education setting. Participants included in the LA group had no diagnosed disability but had a level 1 or 2 on the previous year’s math FCAT. Participants in the AA group had no diagnosed disability and a level 3 or 4 on the previous year’s math FCAT.

All students in the eighth grade at participating schools who were placed in regular, inclusive, intensive, and self-contained special education instructional settings were eligible to take part in this study. Of these students, the first 30 students in each group who returned a signed parent consent form, assented to participate, and met the group eligibility requirements were selected. The school district from which the sample was selected is located in an area of rich cultural and linguistic diversity. Research has shown that students are more successful solving math word problems in their native language (Bernardo, 1999); therefore in order to be eligible to participate, students had to speak English as their primary language or demonstrate exit status from English for Speakers of Other Languages (ESOL) programs (i.e., level 5).

M-DCPS is the fourth-largest school district in the country, serving about 366,000 students. Sixty four percent of the students in the district are Hispanic, 25% are Black, 9% are White, and 2% self-report as Other (Florida Department of Education, 2010). Additionally, 60% of the students qualify for the free- or reduced-lunch program, a measure frequently used in school-based research as a proxy for student socioeconomic status (SES). The participating middle schools represented students from both medium and low SES backgrounds. Additionally, both high and low performing schools (based on
the school FCAT “grade” scores) participated. As is common across the district, student ethnicity was clustered by school. Table 3.1 reports the demographic information for participants.

<table>
<thead>
<tr>
<th>Table 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant Information by Ability Group</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>AA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>LA&lt;sup&gt;b&lt;/sup&gt;</th>
<th>LD&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>20.7</td>
<td>9</td>
</tr>
<tr>
<td>Female</td>
<td>23</td>
<td>79.3</td>
<td>21</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>--</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>25</td>
<td>86.2</td>
<td>18</td>
</tr>
<tr>
<td>Haitian American</td>
<td>2</td>
<td>6.9</td>
<td>1</td>
</tr>
<tr>
<td>White</td>
<td>--</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>ESOL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes/Exited</td>
<td>13</td>
<td>44.8</td>
<td>11</td>
</tr>
<tr>
<td>No</td>
<td>16</td>
<td>55.2</td>
<td>19</td>
</tr>
<tr>
<td>Free/Reduced Lunch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25</td>
<td>86.2</td>
<td>26</td>
</tr>
<tr>
<td>No</td>
<td>4</td>
<td>13.8</td>
<td>4</td>
</tr>
</tbody>
</table>

*Note.* AA = average achieving; LA = low achieving; LD = learning disabilities
<sup>a</sup>n = 29. <sup>b</sup>n = 30. <sup>c</sup>n = 25.

**Measures**

Two measures were administered to students. The researcher-revised form of the *Mathematical Processing Instrument* (Suwarsono, 1982) was used to measure students’
paraphrasing accuracy, visual representation accuracy, and problem-solving accuracy. This measure was administered individually during the first of two sessions; on average, testing took between 30 and 40 minutes. During the second testing session, the Test of Mathematical Ability (TOMA-2; Brown, Cronin, & McEntire, 1994) Computation subtest was given to assess students’ computational ability in an untimed setting. This measure was administered in small groups in unused rooms in the schools and took about 15 minutes.

Mathematical Processing Instrument. The researcher-revised Mathematical Processing Instrument (MPI; Suwarsono, 1982) was used to derive the following variables of interest in this study: the accuracy of paraphrasing (as measured by paraphrasing relevant information, irrelevant linguistic information, and irrelevant numerical information), the accuracy of students’ visual representations (as measured by visual representation of relevant information, irrelevant linguistic information, and irrelevant numerical information), and students’ problem-solving accuracy. The original MPI (Suwarsono, 1982), composed of 15 arithmetic word problems, was used in previous research on visualization in mathematics (e.g., Hegarty & Kozhevnikov, 1999; Lean and Clements, 1981; van Garderen, 2007). An internal reliability estimate using the Kuder-Richardson Formula 20 (KR-20) was reported as 0.85 (van Garderen & Montague, 2003). The questions in the MPI were designed to be conducive to visual representation in that each question lends itself to the use of a drawing to assist in problem solving (see Appendix A). For example, the third item says, “A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50
meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?"

A pilot study was conducted with nine middle school students. Based on pilot data, the MPI was modified in three important ways. First, only nine of the original 15 problems were used. Administration of the original MPI with both the paraphrasing and visual representation components took too long for students to solve in one session. Based on previous research (Hegarty & Kozhevnikov, 1999; van Garderen & Montague, 2003) as well as pilot data, questions 5, 7, 8, 9, 14, and 15 were eliminated leaving the modified MPI with nine word problems. The second way in which the MPI was changed was that for the original items 2, 11, and 13, two units of irrelevant linguistic information were added. For original items 4, 6, and 12, two units of irrelevant numerical information were added. No changes were made to original items 1, 3, and 10. Finally, six versions of the test were created so that item order was counterbalanced to prevent order effects. The revised MPI was presented with one question per page in order to provide ample room for representations and written work. See Figure 3.1 for a visual of the changes.
An internal reliability coefficient for the modified MPI was determined using Cronbach’s alpha (Cronbach, 1951), $\alpha = .65$. Though .70 is the traditional standard for an acceptable alpha level (Nunnally, 1978), it is known that a low number of items on a measure negatively affects its reliability; because the MPI was intentionally shortened, it follows to reason that it would be somewhat less reliable.

**Scoring.** Three scores were generated from the administration of the revised MPI. Each of the scores (i.e., paraphrasing accuracy, visual representation accuracy, and problem-solving accuracy) represent a summative total across the nine problems.

**Accuracy of paraphrasing.** Scores for students’ paraphrasing were determined using the Paraphrasing Accuracy (PA) Scale (see Appendix B). It is a researcher-developed measure that provides three scores across the nine problems: relevant information, irrelevant linguistic information, and irrelevant numerical information. All the components of each of the original MPI questions are counted as relevant.
information. For example, the revised item 5 reads: “From a long stick of wood, a man
cut six short sticks, each two feet long. He then found he had a piece of one foot long left
over. The original stick was two inches thick at one end, but only one inch thick at the
other. Find the length of the original stick.” There are four pieces of relevant information
in this problem (i.e., information that was in the original problem): a) “From a long stick
of wood, a man cut six short sticks”, b) “each two feet long.”, c) “He then found he had a
piece of one foot long left over.”, and d) “Find the length of the original stick.” The total
score for relevant information across the nine problems ranges from zero to 36. It is
important to note that each item on the revised MPI has a unique number of units of
relevant information. Though the average number of items across all nine problems is
four, some problems (i.e., problems 4 and 8) have three units of relevant information and
problem 3 has six units of relevant information. The remaining problems (i.e., 1, 2, 5, 6,
7, and 9) have four units of relevant information each.

All researcher-added components in the problems are counted as either irrelevant
linguistic information or irrelevant numerical information. As per the example above, the
irrelevant numerical information would be: “The original stick was two inches thick at
one end” and “but only one inch thick at the other.” See Figure 3.2 for one student’s
paraphrasing and visual representation, and the scoring for each process.
Student’s paraphrase:
Important Info: I'm going to need the journey of 60 miles, the first 5 miles he walked, and when he got the lift from the lorry driver who stopped for gas partway, he also, I also need the information for when he still had half his journey to go.

Student’s visual representation:

Problem 7 scoring:

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (✓/✗)</th>
<th>Represented? (✓/✗)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>A hitchhiker set out on a journey of 60 miles.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Irrelevant</td>
<td>He used a map to calculate the distance.</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Relevant</td>
<td>He walked the first 5 miles and then got a lift from a lorry driver.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Irrelevant</td>
<td>Who had to stop for gas partway.</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Relevant</td>
<td>When the driver dropped him he still had half of his journey to travel.</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Relevant</td>
<td>How far had he travelled in the lorry?</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

Total Relevant: 3 3
Total Irrelevant: 1 1

Overall MPI scoring:

<table>
<thead>
<tr>
<th>Across Problems</th>
<th>Paraphrasing Score</th>
<th>Visual Representation Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Relevant</td>
<td>25/36</td>
<td>25/36</td>
</tr>
<tr>
<td>Total Irrelevant</td>
<td>4/12</td>
<td>3/12</td>
</tr>
<tr>
<td>Total Irrelevant Linguistic</td>
<td>1/6</td>
<td>1/6</td>
</tr>
<tr>
<td>Total Irrelevant Numerical</td>
<td>3/6</td>
<td>2/6</td>
</tr>
</tbody>
</table>

Figure 3.2. Paraphrasing and visual representation scoring example: problem 7.
Accuracy of visual representation. For each of the students, the Visual Representation Accuracy (VRA) Scale was used to assess the visual representations produced during problem solving. This scale mirrors the PA Scale. The VRA Scale provides three scores across the nine problems: inclusion of relevant information, irrelevant linguistic information, and irrelevant numerical information. The first score is calculated by total number of relevant units of information used; again, the score is out of a possible 36. The irrelevant numerical information score, as in the PA Scale, is based on the sum of total irrelevant numerical units of information recalled across the nine problems; the score is out of a possible six points. Again, Figure 3.2 shows the scoring for MPI item 5. Identification of the components of each representation along a schematic continuum was determined using the guidelines set forth by Hegarty and Kozhevnikov (1999); that is, schematic components must represent “the spatial relationships between objects” (p. 685) and not simply be depicted in picture form. Each piece of relevant information was either represented schematically or not in the diagram (1 or 0). To illustrate the schematic continuum, three different participants’ representations of MPI item 1 are presented in Figure 3.3.

Problem-solving accuracy. The revised MPI is scored for accuracy, where each answer is scored as correct or incorrect. This results in a total score ranging from zero to nine for each participant.

Test of Mathematical Ability (TOMA-2). The TOMA-2 (Brown et al., 1994) is a standardized assessment with five subtests that combine to provide a composite score of overall ability. Subtests include Vocabulary, Computation, General Information, Story Problems, and Attitude to Math. For the purposes of this study, students completed only
the Computation subtest. The Computation subtest consists of 30 increasingly difficult equations covering an array of skills, including the four operations, exponents, fractions, degrees, ratios, percentages, and scientific notation. For example, the tenth question is $742 + 388 = \_\_\_\_\_\_\_\_$; the twenty-fifth question is $3:2$ as $15: \_\_\_\_\_\_\_\_$ . Directions are standardized, where students are told, “Write the answer to each problem. Write on the paper if you need to. Be sure to include any decimal points, dollar signs, or other marks if they are used in the problem” (TOMA-2 profile/record form, p. 2).

**Scoring.** There is no time limit on any of the TOMA-2 subtests; students continue until they reach the ceiling (i.e., three consecutive errors) or until they complete the subtest. Each of the 30 problems receives one point for being correct or zero points for being incorrect. The raw score is the total correct items out of 30.
Item 1: At each of the two ends of a straight path, a man planted a tree. Then every five meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?

Figure 3.3. The first representation (a) is completely pictorial in nature. None of the relevant pieces of information are represented in relation to the others. The second representation (b) is partially schematic but incomplete in representing the relevant information. The third representation (c) is fully schematic, depicting all relevant components of the problem.
Procedures

Consent. The study was approved through the Internal Review Board (IRB) at the University of Miami as well as by the M-DCPS Research Review Committee (RRC). Consent was obtained following the procedures mandated by the IRB and RRC. That is, students who fit the study criteria were assembled in a classroom by the school liaison (i.e., assistant principal, guidance counselor, or math teacher), and the study was explained to each class by the researcher. Students were informed that their participation was voluntary through all phases of the study, and that their withdrawal from participation would not result in any negative outcomes for them. The researcher handed out parent consent forms (Appendix C) to all the eligible students. The students who returned signed parent consents were individually assented by the researcher, during which the entire student assent form (Appendix D) was read and discussed. The first 30 students in each ability group who submitted approved assents/consents and met the inclusionary criteria participated.

Background information. As explained in the consent and assent forms, permission to participate in the study provided the researcher access to: a) prior achievement data, such as state assessment scores and report cards; b) student cumulative records including health information and psycho-educational testing results; and c) demographic and student information, such as SES, ethnic background, age and birth date, and class schedules.

Assessment. The researcher and a colleague trained in the protocol administered the measures to students both individually and in a group setting over a two-month period during the spring semester of the 2010 school year. Testing took approximately 40 to 60
minutes over the course of two sessions. All sessions followed the administration protocol below to ensure consistency across testing.

**Administration.** First, the revised MPI was individually administered to each student. Before the test was presented, the researcher said: “I am interested in learning about how middle school students solve math word problems. I will ask you to solve nine math word problems.” Each student was permitted to use a calculator, which was provided by the researcher. The student had the first problem placed in front of him or her, and all student paraphrasing was recorded.

1. The researcher read the problem to the student and then said, “Now read it to yourself and tell me when you understand it.”

2. After the student indicated that he or she was done reading, the researcher said, “Put the problem in your own words. In other words, if you were explaining this problem to a friend who hadn’t read it, what would you say?”

3. After the student responded, the researcher said, “Now tell me the important information that you would need to use to solve the problem.”

4. After the student responded, the researcher said, “Now I want you to draw a picture that will help you solve the problem. Then solve it.”

The researcher then repeated the procedure for the remaining eight questions. During the second session, the researcher administered the TOMA-2 Computation subtest in a group setting. No calculators were permitted for this test. The testing protocol was read to the students as per the TOMA-2 examiner’s manual (Brown et al., 1994).

**Coding.** All paraphrasing recordings were transcribed. All 84 of the transcriptions were cross-checked for accuracy with the original recording; accuracy of the
transcriptions was 100%. The transcriptions of student paraphrasing were then coded by the researcher following the scoring described above with the PA Scale. The coding of the visual representations was based on the images generated by the students, also following the scoring procedure above. Separate scores, both on the PA Scale and the VRA Scale, were provided for relevant information, irrelevant linguistic information, and irrelevant numerical information. In order to increase the reliability of the coding, a guide separated by problem was created and used by the researcher as well as by the second rater (see Appendix E for the coding guide).

**Reliability.** For the paraphrasing data and the accuracy of visual representation data, the researcher and a trained colleague independently scored and then compared 5 of the protocols using the PA Scale and the VRA Scale to establish initial agreement; discussion resolved any disagreements and clarified potential ambiguity. All protocols were returned to the pool, and the researcher scored each independently. A trained colleague independently scored 15% of the protocols, which were then used to calculate inter-rater agreement. Inter-rater agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. All differences were recorded as disagreements and were then discussed and resolved to 100% agreement. Inter-rater agreement was 96.0% for the PA Scale and 91.3% for the VRA Scale. Traditionally, greater than 80% agreement is considered acceptable reliability (Kazdin, 1982).

**Design and Analysis**

Four preliminary analyses were used to describe the characteristics of the sample. Chi-square goodness-of-fit analyses were employed to determine whether the sample was
reflective of the district from which it was drawn; specifically, ethnicity and socio-economic status were the variables of interest. Additionally, three univariate ANOVAs were conducted in order to determine whether groups differed on scores from the previous year’s math and reading FCAT and on scores from the TOMA-2 Computation subtest. The post hoc analysis using Tukey’s adjustment was used to further examine the significant differences that were found from the univariate ANOVAs.

A natural group comparison design was used in which the three groups by ability (i.e., LD, LA, and AA) were compared. The dependent variables (DVs) in this study were the: a) paraphrasing relevant information score, c) paraphrasing irrelevant linguistic information score, d) paraphrasing irrelevant numerical information score, e) visual representation relevant information score, f) visual representation irrelevant linguistic information score, f) visual representation irrelevant numerical information score, and g) problem-solving accuracy score.

Corresponding to the first question, Pearson’s product-moment correlation was employed to determine the magnitude and direction of the linear relationships among the variables of interest. All seven DVs were used in the model; the main interest was in the relationships of variables to problem-solving accuracy. Correlation matrices were generated first across all students, and then disaggregated by ability group.

In order to answer the second research question, a one-way Multivariate Analysis of Variance (MANOVA) was employed in which three ability groups were compared on the linear combination of seven dependent variables of the MPI. Because the dependent variables that represent one underlying construct (i.e., problem solving) are interrelated, it
was conceptually appropriate to analyze them together as a “system” as opposed to conducting a series of univariate analyses to determine group differences (Huberty & Morris, 1989, p. 304). MANOVA also prevents the unacceptably high Type I error rates that would be problematic with unadjusted multiple univariate analyses (Tabachnick & Fidell, 2007). Three groups by students’ ability level were compared on the scores associated with the MPI (i.e., paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information, visual representation relevant information, visual representation irrelevant linguistic information, visual representation irrelevant numerical information, problem-solving accuracy).

MANOVA operates effectively when four main assumptions are met: the sampling distributions of means are normally distributed, the variance-covariance matrices are homogeneous, there are no outliers, and there are linear relationships among all DVs (Tabachnick & Fidell, 2007). For the first two assumptions, an adequately large sample size with equal groups is robust against their violation. Additionally, histograms of DVs were examined to ensure the normality of their distributions, and Box’s test of homogeneity was conducted. For the third assumption, variables were plotted so that potential outliers could be identified; the fourth assumption was checked by visually checking scatterplots of DVs by groups to ensure that relationships were linear.

After running the MANOVA, the nature of the significant differences was further examined using Tukey’s HSD post hoc procedure. Tukey’s test was chosen because it makes all possible comparisons and maintains adequate power under most conditions as opposed to other post hoc procedures such as Bonferroni and the Fisher LSD (Seaman,
Levin, & Serlin, 1991). Additionally, Bonferroni’s post hoc test has been criticized for its overcorrection of Type I error whereas the Fisher LSD test has been criticized for its lack of control over Type I error; Tukey’s procedure falls between the two and thus will be most effective for this study. The eta-squared value was used to quantify the magnitude of group differences based on Cohen’s (1988) standards for magnitude of effect size (i.e., $\eta^2 = .01$ is considered small, .06 is medium, and .14 is high).

The third research question was addressed through hierarchical multiple regression. This analysis was conducted to a) investigate the unique contribution of paraphrasing accuracy and visual representation accuracy to problem-solving accuracy when the variance attributed to other factors was partialed out from the analysis and b) determine whether ability was a moderating factor for the effect of predictors on problem-solving accuracy.

Like MANOVA, there are assumptions that underlie multiple regression. Specifically, regression assumes a) linearity, b) independence, c) homoscedasticity, and d) normality. Linearity was assessed by plotting the residuals against the variables of interest. Scatterplots were generated to provide evidence of the linearity of the DV (Keith, 2006). Independence of error was determined by study design. The assumption of homoscedasticity was checked by plotting the predicted scores against problem-solving accuracy. Normality was determined by generating a histogram of the unstandardized residuals.

The general goal in multiple regression is to predict scores on a dependent variable based on the scores of predictor variables (Cunningham & Wang, 2005). Hierarchical multiple regression differs from simultaneous regression in that variables are
entered in a pre-determined sequential order, permitting the discovery of the effects of
certain predictor variables independent of other variables. Accordingly, this analysis
permitted the problem-solving process to be analyzed in sequence, where the total effect
of paraphrasing accuracy could be measured after accounting for ability, and the total
effect of visual representation accuracy could be measured after accounting for ability
and paraphrasing accuracy. The order of entry of variables in a hierarchical regression is
crucial not only to the outcome, but also to the validity of the analysis itself. Because any
shared variance between predictor variables would be attributed to the variable entered
first in the sequence, it is imperative that there be a sound rationale for the proposed
order. Common justifications include presumed/actual time precedence or theoretical
underpinnings of the research (Keith, 2006).

The order of predictor variables for this study was determined based on both time
precedence (ability preceding all predictor variables) as well as the aforementioned
theories of the math problem-solving process (i.e., Mayer, 1985; Montague, 2003). Since
the MANOVA results showed significant differences among ability groups on the
problem-solving variables, ability was included in the model. Two dummy codes were
created (DummyLD and DummyLA) to transform ability from a three-level categorical
variable into two dichotomous variables; these were then entered into the first block.
Both paraphrasing relevant information, paraphrasing irrelevant linguistic information,
and paraphrasing irrelevant numerical information were centered, and those centered
terms were entered in the second block as paraphrasing accuracy. The third block
contained visual representation accuracy (which consisted of visual representation
relevant information, visual representation irrelevant linguistic information, and visual
representation irrelevant numerical information); these variables were also centered. All variables were regressed on problem-solving accuracy. The presence of interaction effects was explored by separately entering each interaction variable into the fourth block of the model. Thus the interaction of ability and each of the previous variables (i.e., paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information, visual representation of relevant information, visual representation of irrelevant linguistic information, and visual representation of irrelevant numerical information) was analyzed after controlling for the variables in blocks one, two, and three. Interactions that were significant were incorporated into the final model; significant interactions provided information about which variables differed by ability after controlling for the variables entered earlier. The magnitude of effects was interpreted based on the standardized regression coefficient (β), which provided the practical significance of the results. Keith’s (2006) rules of thumb were applied to determine effect size (i.e., $0.05 < \beta \leq 0.10$ is considered a small effect size, $0.10 < \beta \leq 0.25$ a moderate effect size, and $\beta > 0.25$ a large effect size).
Chapter 4

Results

The purpose of this study was twofold. First, it examined the relationship between paraphrasing and visual representation as well as how these two processes relate to overall problem-solving performance. Second, it investigated the assumption that no valid differences exist between students with learning disabilities (LD) and low achieving (LA) students by analyzing the process by which they solved math word problems (as compared to a reference group of average achievers). This chapter is organized in the order of the research questions. However, before the results of the first research question are discussed, the results of the preliminary descriptive analyses (i.e., chi-square analyses for goodness-of-fit, ANOVAs) are provided. The following section addresses the first research question by discussing the nature of the relationships among the seven dependent variables: a) paraphrasing relevant information, c) paraphrasing irrelevant linguistic information, d) paraphrasing irrelevant numerical information, e) visual representation of relevant information, f) visual representation of irrelevant linguistic information, f) visual representation of irrelevant numerical information, and g) problem-solving accuracy. The subsequent section focuses on the second research question, where the relationships between ability group and paraphrasing, visual representation, and problem-solving accuracy are analyzed. The fourth section discusses the relationship of paraphrasing accuracy and visual representation accuracy to problem-solving accuracy and the role of ability in that relationship. Following the results for each question is a summary of findings.
Preliminary Analyses

A chi-square test of goodness-of-fit was performed to determine whether the ethnicity of the sample was reflective of the district from which it was drawn. The ethnicity of the sample did reflect that of the district, $\chi^2 (3) = 6.78, p = .079$. However, the chi-square test of goodness-of-fit comparing SES revealed that the sample did not reflect the district from which it was drawn, $\chi^2 (1) = 27.63, p < .001$. See Table 4.1 for a summary of the sample-by-district comparison on the variables of interest.

Table 4.1.

Sample Characteristics: SES and Ethnicity as Compared to the District

<table>
<thead>
<tr>
<th>Variable</th>
<th>Samplea (%)</th>
<th>District (%)</th>
<th>Residualsb (%)</th>
<th>Chi-Square $\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free/Reduced Lunch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>88</td>
<td>60</td>
<td>28</td>
<td>27.63***</td>
<td>1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>40</td>
<td>-28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>24</td>
<td>25</td>
<td>-1</td>
<td>6.78</td>
<td>3</td>
<td>.079</td>
</tr>
<tr>
<td>Hispanic</td>
<td>73</td>
<td>64</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1</td>
<td>9</td>
<td>-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. a N = 84. b Residuals = Sample - District
* p < .05, ** p < .01, *** p < .001.

Because the four schools participating in this study (two in particular) served primarily low-income, minority students, it was expected that the sample characteristics would
have more minority students than the district as a whole. While Black students and students identified as “Other” were accurately represented in the sample population, White students were underrepresented in the sample population; consequently, Hispanic students were overrepresented. In contrast to ethnicity, SES did not match the district; this was expected since all four schools had free/reduced lunch rates that ranged from 72% to 95% ($M = 87$, $SD = 10.42$).

Because the students were categorized by ability level and one group consisted of students with disabilities, there was the possibility that by the eighth grade, poor-performing students had failed previous years creating a difference in age. Mean ages were thus calculated for each group. Table 4.2 displays the means and standard deviations by ability group.

<table>
<thead>
<tr>
<th>Group</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>14.49</td>
<td>0.73</td>
</tr>
<tr>
<td>LA</td>
<td>14.67</td>
<td>0.61</td>
</tr>
<tr>
<td>LD</td>
<td>15.41</td>
<td>0.68</td>
</tr>
</tbody>
</table>

*Note.* AA = Average Achieving, LA = Low Achieving, LD = Learning Disabilities

In order to determine whether these differences were significant, a one-way ANOVA was run comparing the three groups (i.e., LD, LA, AA) on age. Results showed that age was significantly different among groups, $F(2, 81) = 13.78$, $p < .001$. Tukey’s post hoc procedure revealed that students with LD were significantly older than both LA
students ($M_{\text{dif}} = 0.74, p < .001$) and AA students ($M_{\text{dif}} = 0.918, p < .001$), though no differences existed between LA and AA students ($M_{\text{dif}} = 0.184, p < .548$).

In order to determine whether the grouping technique reflected the intended purpose, three one-way ANOVAs were run comparing the three groups (i.e., LD, LA, AA) on a) the previous year’s math FCAT, b) the previous year’s reading FCAT, and c) the TOMA-2 Computation subtest. See Table 4.3 for the means and standard deviations by ability group for each measure.

<table>
<thead>
<tr>
<th>Measure</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math FCAT AA</td>
<td>323.66</td>
<td>13.70</td>
</tr>
<tr>
<td>LA</td>
<td>266.03</td>
<td>30.07</td>
</tr>
<tr>
<td>LD</td>
<td>265.20</td>
<td>27.71</td>
</tr>
<tr>
<td>Reading FCAT AA</td>
<td>331.78</td>
<td>24.45</td>
</tr>
<tr>
<td>LA</td>
<td>267.20</td>
<td>42.66</td>
</tr>
<tr>
<td>LD</td>
<td>279.56</td>
<td>41.88</td>
</tr>
<tr>
<td>Computation AA</td>
<td>15.48</td>
<td>2.85</td>
</tr>
<tr>
<td>LA</td>
<td>14.23</td>
<td>3.46</td>
</tr>
<tr>
<td>LD</td>
<td>12.40</td>
<td>3.37</td>
</tr>
</tbody>
</table>

*Note.* AA = Average Achieving, LA = Low Achieving, LD = Learning Disabilities

Univariate effects revealed significant differences on all three dependent variables. See Table 4.4 for summary statistics of the ANOVAs. For all pairwise comparisons on the math and reading FCAT scores, AA students’ scores were significantly higher than both students with LD and LA students, while no differences existed between the latter two groups. Students with LD performed significantly more
poorly on the TOMA-2 Computation subtest than AA students. There were no significant
mean differences on the Computation subtest between students with LD and LA students,
nor between AA students and LA students. Table 4.5 displays the pairwise comparisons
between ability groups across the three analyses.

Table 4.4
One-way ANOVAs Summary Statistics: Math FCAT, Reading FCAT, Computation

<table>
<thead>
<tr>
<th>Measure</th>
<th>F</th>
<th>df1, df2</th>
<th>$\eta^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math FCAT</td>
<td>51.85***</td>
<td>2, 81</td>
<td>.56</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Reading FCAT</td>
<td>20.30***</td>
<td>2, 75</td>
<td>.35</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Computation</td>
<td>6.13**</td>
<td>2, 81</td>
<td>.13</td>
<td>.003</td>
</tr>
</tbody>
</table>

*Note.*
*p* < .05, **p** < .01, ***p** < .001.

These results suggest that the grouping of students was accurate to the purposes intended
by this study; that is, AA students outperformed both groups, and students with LD and
LA students demonstrated similar results on these performance-based measures. The
results are in line with the literature on the performance of LA students, AA students, and
students with LD (e.g., Gonzalez & Espinel, 2002; Mabbott & Bisanz, 2008; Montague et
al., 2009).
Table 4.5

One-way ANOVAs Pairwise Comparisons: Math FCAT, Reading FCAT, Computation

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean Difference</th>
<th>SE</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math FCAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>57.62***</td>
<td>6.46</td>
<td>2.49</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>58.46***</td>
<td>6.77</td>
<td>2.79</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>0.83</td>
<td>6.72</td>
<td>0.03</td>
</tr>
<tr>
<td>Reading FCAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>64.58***</td>
<td>10.52</td>
<td>1.88</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>52.22***</td>
<td>11.00</td>
<td>1.58</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>-12.36</td>
<td>10.28</td>
<td>-0.30</td>
</tr>
<tr>
<td>Computation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>1.25</td>
<td>0.84</td>
<td>0.40</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>3.08**</td>
<td>0.88</td>
<td>1.01</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>1.83</td>
<td>0.88</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Note.  
* p < .05, ** p < .01, *** p < .001.

Question 1: Relationships among Problem-Solving Variables

The first research question addressed the relationships among the seven dependent variables. Because no other study addressed the phases of the problem-solving process and how they relate to each other, it was important that the overall magnitude and direction of relationships be established. Pearson product-moment correlations of seven variables were used to answer this question. See Table 4.4 for a summary of the results of the across-groups correlation. Since the relationships between problem-solving accuracy and the other DVs were of particular interest, those relationships are described first. As hypothesized, results showed that problem-solving accuracy was significantly and positively correlated with paraphrasing relevant information; according to Keith’s (2006) standards for educational research, the magnitude of effect was small. Problem-solving accuracy was negatively correlated with paraphrasing irrelevant numerical and linguistic
information, both with a medium effect size. As with paraphrasing relevant information, problem-solving accuracy was positively correlated with visual representation relevant information, but its effect size was very large. Finally, problem-solving accuracy was negatively correlated with visual representation irrelevant numerical and linguistic information, and both effect sizes were small. These results show that as students’ identification and representation of relevant information increased, so too did their problem-solving accuracy. Conversely, as students’ identification and representation of irrelevant information (both linguistic and numerical) increased, their problem-solving accuracy decreased. In summary, all problem-solving variables’ relationships with problem-solving accuracy were significant and all were in the hypothesized directions.

In summary, the correlation matrix showed expected relationships among variables; that is, relevant information in both the paraphrasing and the visual representation phases were significantly positively correlated with problem-solving accuracy, whereas numerical and linguistic irrelevant information for the two constructs were significantly negatively correlated with problem-solving accuracy. Significant positive relationships between paraphrasing and visual representation were demonstrated with all variables except irrelevant linguistic information. However, because a major component of this study centers on group differences on the problem-solving process, it was important to analyze relationships separated by ability group; in this way, the size and magnitude of correlations by ability could be examined and compared. Table 4.7 displays the results of the by-group correlation matrices.
Table 4.6

Correlation Matrix for Problem-Solving Variables: All Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>P-Rel</th>
<th>P-Irrel L</th>
<th>P-Irrel N</th>
<th>V-Rel</th>
<th>V-Irrel L</th>
<th>V-Irrel N</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-Rel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>1.0</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-Irrel L</td>
<td>.38***</td>
<td>1.0</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.144</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-Irrel N</td>
<td>.32**</td>
<td>.48***</td>
<td>1.0</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>.003</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.102</td>
<td>.230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-Rel</td>
<td>.36**</td>
<td>-.10</td>
<td>-.08</td>
<td>1.0</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>.001</td>
<td>.354</td>
<td>.475</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.130</td>
<td>.010</td>
<td>.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-Irrel L</td>
<td>-.31</td>
<td>.15</td>
<td>.28**</td>
<td>-.08</td>
<td>1.0</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>.737</td>
<td>.168</td>
<td>.010</td>
<td>.481</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.096</td>
<td>.023</td>
<td>.078</td>
<td>.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-Irrel N</td>
<td>.22*</td>
<td>.33**</td>
<td>.54***</td>
<td>.10</td>
<td>.20</td>
<td>1.0</td>
<td>--</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>.042</td>
<td>.002</td>
<td>&lt;.001</td>
<td>.349</td>
<td>.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.048</td>
<td>.109</td>
<td>.292</td>
<td>.010</td>
<td>.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>.27*</td>
<td>-.32**</td>
<td>-.33**</td>
<td>.66***</td>
<td>-.23*</td>
<td>-.30**</td>
<td>1.0</td>
</tr>
<tr>
<td>$r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>.013</td>
<td>.003</td>
<td>.002</td>
<td>&lt;.001</td>
<td>.035</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>.073</td>
<td>.102</td>
<td>.109</td>
<td>.436</td>
<td>.053</td>
<td>.090</td>
<td></td>
</tr>
</tbody>
</table>

Note. P = Paraphrasing; V = Visual Representation; Rel = Relevant Information; Irrel = Irrelevant Information; N = Numerical; L = Linguistic.

$n = 84$.

* $p < .05$, ** $p < .01$, *** $p < .001$. 
Average-achieving students. The relationships among variables differed between AA students and the overall sample on several of the bivariate correlations. Specifically, neither paraphrasing relevant information nor visual representation of relevant information was significantly correlated to problem-solving accuracy, though the trend was in the expected direction. In line with the overall results, the AA group showed a significant and negative relationship with problem-solving accuracy and paraphrasing irrelevant linguistic information \((p = .041)\), paraphrasing irrelevant numerical information \((p = .020)\), and visual representation of irrelevant numerical information \((p = .008)\).

Though correlational research is non-directional, by nature of time sequence of the problem-solving process, this result suggests that AA students’ inclusion of irrelevant information is more predictive of their problem-solving performance than inclusion of the relevant information needed to solve the problem. Other relationships of significance for the AA group include a positive correlation between paraphrasing irrelevant linguistic information and paraphrasing irrelevant numerical information \((p = .001)\) as well as paraphrasing irrelevant numerical information and visual representation of irrelevant numerical information \((p < .001)\).
Table 4.7

**Correlation Matrix for Problem-Solving Variables: AA, LA, LD**

<table>
<thead>
<tr>
<th>Variable</th>
<th>P-Rel</th>
<th>P-Irrel L</th>
<th>P-Irrel N</th>
<th>V-Rel</th>
<th>V-Irrel L</th>
<th>V-Irrel N</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>LA</td>
<td>LD</td>
<td>AA</td>
<td>LA</td>
<td>LD</td>
<td>AA</td>
</tr>
<tr>
<td>P-Rel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.45*</td>
<td>.46*</td>
<td>.35</td>
<td>.20</td>
<td>.21</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.58**</td>
<td>.32</td>
<td>.33</td>
<td>.34</td>
<td>.35</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>P-Irrel L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.33</td>
<td>.58**</td>
<td>.70**</td>
<td>.46*</td>
<td>.25</td>
<td>.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.11</td>
<td>.34</td>
<td>.36</td>
<td>.21</td>
<td>.27</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>P-Irrel N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.34</td>
<td>.32</td>
<td>.10</td>
<td>.11</td>
<td>.10</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.12</td>
<td>.07</td>
<td>.03</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>V-Rel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.09</td>
<td>-.01</td>
<td>-.03</td>
<td>.35</td>
<td>.28</td>
<td>.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.01</td>
<td>-.01</td>
<td>.01</td>
<td>.12</td>
<td>.08</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>V-Irrel L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.17</td>
<td>.29</td>
<td>.10</td>
<td>.36</td>
<td>.21</td>
<td>.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.03</td>
<td>.08</td>
<td>.01</td>
<td>.04</td>
<td>.07</td>
<td>.22</td>
<td></td>
</tr>
<tr>
<td>V-Irrel N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.19</td>
<td>.15</td>
<td>.19</td>
<td>.38*</td>
<td>.30</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04</td>
<td>.02</td>
<td>.04</td>
<td>.14</td>
<td>.16</td>
<td>.24</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.19</td>
<td>.15</td>
<td>.19</td>
<td>.38*</td>
<td>.30</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04</td>
<td>.02</td>
<td>.04</td>
<td>.14</td>
<td>.16</td>
<td>.24</td>
<td></td>
</tr>
</tbody>
</table>

*Note. P = Paraphrasing; V = Visual Representation; Rel = Relevant Info; Irrel = Irrelevant Info; N = Numerical; L = Linguistic
* p < .05, ** p < .01, *** p < .001.
Interestingly, AA students also showed a significant and positive relationship between paraphrasing relevant information and paraphrasing irrelevant linguistic information \( (p = .014) \). This seems in line with the previously stated hypothesis that some students may view irrelevant linguistic information as enhancing the context of the problem and thus beneficial to include. Overall, AA students’ results were similar to those found across groups.

**Low-achieving students.** Most of the significant correlations found in LA students’ correlation matrix were expected, reflecting those relationships found across groups. However, the LA group had only one significant correlation with problem-solving accuracy: this was with visual representation of relevant information \( (p < .001) \). No other problem-solving process (relevant or irrelevant) significantly correlated with problem-solving accuracy. Of the correlations within the processes, LA students’ results showed a significant and positive correlation between paraphrasing irrelevant linguistic information and paraphrasing irrelevant numerical information \( (p = .011) \). Additionally, paraphrasing irrelevant numerical information was significantly and positively correlated with visual representation of irrelevant numerical information \( (p = .002) \). These last two relationships mirror those found with the AA group as well as across groups. However, not all of the relationships of the LA group reflected the expected results; there was a significant and positive relationship between paraphrasing relevant information and both paraphrasing irrelevant linguistic information \( (p = .010) \) and paraphrasing irrelevant numerical information \( (p = .001) \). These results suggest that LA students have poor
discriminatory abilities with regard to identifying relevant versus irrelevant information. Essentially, LA students paraphrased the information with little regard for its usefulness in solving the problem.

**Students with LD.** In contrast to the AA and LA groups, students with LD show some trends divergent from the results of the across-groups correlation. Of the twenty-one bivariate correlations analyzed, only three of them were significant for the LD group. One interesting finding was the significant and positive correlation between paraphrasing irrelevant numerical information and visual representation of irrelevant linguistic information ($p = .046$). Both AA and LA students had significant relationships between the former variable and visual representation of irrelevant numerical information, which makes logical sense—it means that the irrelevant numerical information that they identified as important was incorporated into their visual representation. For students with LD, however, the irrelevant numerical information they identified (i.e., paraphrased) was not statistically related to that information being incorporated into their representation; neither was the linguistic information they represented in their diagram significantly related to their identification of that linguistic information in their paraphrase. Another unique finding that differed from the LA and AA groups was that students with LD had a significant and positive relationship between visual representation of relevant information and visual representation of irrelevant numerical information ($p = .017$). This suggests that their discrimination abilities between relevant and irrelevant numerical information are poor; interestingly, this was not reflected in the paraphrasing phase. Finally, there was only one significant relationship for students with LD that was also found with LA students; that is, results showed a significant and positive relationship
between visual representation of relevant information and problem-solving accuracy ($p < .001$). This suggests that for low-achieving students regardless of disability status, the more their visualization schematically represents the problem, the better their problem-solving accuracy is.

**Summary.** Overall, correlations within the AA group generally showed the relationships expected, both in direction and strength. To a lesser extent, LA students also demonstrated similar patterns in terms of significant relationships. Significant correlations for the LD group, however, showed a departure from the pattern of results, both from the other two groups and from the across-groups results. Figure 4.1 summarizes the similarities and differences in significant relationships. Among the AA, LA, and LD correlation matrices, 10 of the 21 unique cells had significant results. Three of the 10 were significant for both AA and LA groups. The significant relationships in AA and LD groups shared no common cells (i.e., no significant relationship for AA was shared with LD). LD and LA groups shared only one common cell (i.e., the relationship between visual representation of relevant information and problem-solving accuracy). Thus, zero-order correlations by ability group on the problem-solving processes showed different profiles for AA, LA, and LD groups based on visual analysis. The significant relationships between the problem-solving variables of students with LD stood apart as compared to those of LA and AA students.
Question 2: Differences in Problem-Solving Variables by Ability Group

In order to answer the second research question, a one-way between-subjects multivariate analysis of variance was performed on seven dependent variables: paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information, visual representation of relevant information, visual representation of irrelevant linguistic information, visual representation of irrelevant numerical information, and problem-solving accuracy. The independent variable was ability group (LD, LA, and AA).

Prior to analysis, the seven DVs were examined through various procedures to ensure that data entry as well as the assumptions of MANOVA were met. MANOVA
efficacy is dependent on the satisfaction of four main assumptions: linearity, normality, homogeneity of variance-covariance, and the absence of outliers. Linearity was checked by visually inspecting scatterplots of DVs by groups to ensure that relationships were linear. Normality was determined by generating histograms of the seven DVs. Histograms showed that paraphrasing relevant information and visual representation of relevant information were both normally distributed. Paraphrasing irrelevant linguistic information, visual representation of linguistic information and visual representation of irrelevant numerical information were positively skewed, as was problem-solving accuracy but to a lesser degree. Paraphrasing irrelevant numerical information showed a flat profile. Because MANOVA is fairly robust to violations to the normality assumption with adequate sample sizes and equal groups, this violation was considered within acceptable limits. Homogeneity of variance-covariance was measured using Box’s test of equality of covariance matrices. Box’s test was not significant, $M = 81.76, F(56, 17824) = 1.28, p > .05$, meaning that the assumption of homogeneity was satisfied. The absence of outliers was examined in both univariate and multivariate form. Univariate outliers were detected by saving the standardized values of all DVs as variables and then examining each score for deviation from the distribution. None met the criteria set out by Tabachnick and Fidell (2007), who have identified $z$-scores greater than 3.3 as outliers. Multivariate outliers were detected by computing Mahalanobis Distance and then screening the generated variables grouped by ability for potential abnormalities. With $\alpha$ set at .05, the critical value was 14.07; no outliers were thus identified within LD, LA, or AA groups. In summary, most of the assumptions were met, and those that fell short of the standards were deemed acceptable due to the robust nature of the MANOVA model.
In order to determine whether the results of the overall MANOVA were significant, Wilks’ Lambda was used to assess the omnibus test; it showed that the dependent variables significantly differed by ability, $F(14, 150) = 3.72, p < .001$. Based on Cohen’s (1988) standards for magnitude of effects, the results showed a large association between ability scores and the combined DVs, partial $\eta^2 = .26$; the observed power was .99. Table 4.8 reports the means and standard deviations of the ability levels for each of the seven dependent variables.

Significant univariate effects were found for the following variables: paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information, visual representation of relevant information, visual representation of irrelevant numerical information, and problem-solving accuracy. The observed power for all variables ranged from .60 to 1.00, except for visual representation of irrelevant linguistic information, which had an observed power of .27. As such, this was also the only nonsignificant univariate effect found. See Table 4.9 for the summary statistics of the univariate effects.

Post hoc analyses using Tukey’s procedure were conducted for all significant univariate effects. Results revealed that students with LD paraphrased significantly less relevant information than both AA students and LA students. Pairwise comparisons for paraphrasing irrelevant numerical information showed that LA students identified more irrelevant numerical information than did AA students. No significant difference on paraphrasing irrelevant numerical information was found between students with LD and the other two groups.
Table 4.8

Means and Standard Deviations: Problem-Solving Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>AA</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-Rel</td>
<td>AA</td>
<td>22.52</td>
<td>4.60</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>21.33</td>
<td>5.36</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>17.84</td>
<td>5.65</td>
</tr>
<tr>
<td>P-Irrel L</td>
<td>AA</td>
<td>0.45</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>1.03</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>0.60</td>
<td>0.82</td>
</tr>
<tr>
<td>P-Irrel N</td>
<td>AA</td>
<td>2.14</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>3.93</td>
<td>2.07</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>2.80</td>
<td>2.26</td>
</tr>
<tr>
<td>V-Rel</td>
<td>AA</td>
<td>18.48</td>
<td>4.30</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>12.83</td>
<td>7.21</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>9.16</td>
<td>5.75</td>
</tr>
<tr>
<td>V-Irrel L</td>
<td>AA</td>
<td>0.28</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>0.53</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>0.44</td>
<td>0.65</td>
</tr>
<tr>
<td>V-Irrel N</td>
<td>AA</td>
<td>0.97</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>1.83</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>0.80</td>
<td>1.04</td>
</tr>
<tr>
<td>Accuracy</td>
<td>AA</td>
<td>4.28</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>LA</td>
<td>2.23</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>2.08</td>
<td>1.89</td>
</tr>
</tbody>
</table>

Note. AA = Average Achieving, LA = Low Achieving, LD = Learning Disabilities.
P = Paraphrasing; V = Visual Representation; Rel = Relevant Information; Irrel = Irrelevant Information; N = Numerical; L = Linguistic.
Table 4.9

*A Univariate Effect from MANOVA*

<table>
<thead>
<tr>
<th>Variable</th>
<th>$F$</th>
<th>$df$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paraphrase Relevant</td>
<td>5.78**</td>
<td>2, 81</td>
<td>.005</td>
<td>.125</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Linguistic</td>
<td>3.22*</td>
<td>2, 81</td>
<td>.045</td>
<td>.074</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Numerical</td>
<td>5.67**</td>
<td>2, 81</td>
<td>.005</td>
<td>.123</td>
</tr>
<tr>
<td>Vis Rep Relevant</td>
<td>17.26***</td>
<td>2, 81</td>
<td>&lt;.001</td>
<td>.299</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Linguistic</td>
<td>1.29</td>
<td>2, 81</td>
<td>.282</td>
<td>.031</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Numerical</td>
<td>3.89*</td>
<td>2, 81</td>
<td>.024</td>
<td>.088</td>
</tr>
<tr>
<td>Accuracy</td>
<td>13.51***</td>
<td>2, 81</td>
<td>&lt;.001</td>
<td>.250</td>
</tr>
</tbody>
</table>

Note. Wilks’ Lambda: $F(14, 150) = 3.72, p < .001$.  
* $p < .05$, ** $p < .01$, *** $p < .001$.

Paraphrasing irrelevant linguistic information differences that were found mirrored those found with irrelevant numerical information. That is, LA students identified more irrelevant linguistic information than did AA students, while no significant differences were found between students with LD and the other two groups. AA students used significantly more relevant information in their visual representations than both LA students and students with LD. The difference between students with LD and LA students was insignificant, though the former used less relevant information than the latter group. Pairwise comparisons of irrelevant numerical information used in visual representations revealed the second significant distinction between LA students and students with LD. The former group used significantly more irrelevant numerical information in their visual representations than did the latter group. Finally, pairwise comparisons were performed on problem-solving accuracy. As expected, results showed that AA students significantly outperformed both LA students and students with LD.

Also in line with the performance hypothesis was the non-significant differences between
students with LD and LA students, the post hoc revealing just how similar the two groups were. See Table 4.10 for the pairwise comparisons of ability groups on the seven DVs.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean Difference</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-Rel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>1.18</td>
<td>1.35</td>
<td>.658</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>4.68**</td>
<td>1.41</td>
<td>.004</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>3.49*</td>
<td>1.41</td>
<td>.040</td>
</tr>
<tr>
<td>P-Irrel L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>-0.59*</td>
<td>0.24</td>
<td>.043</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>-0.15</td>
<td>0.25</td>
<td>.817</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>0.43</td>
<td>0.25</td>
<td>.195</td>
</tr>
<tr>
<td>P-Irrel N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>-1.80**</td>
<td>0.54</td>
<td>.004</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>-0.66</td>
<td>0.57</td>
<td>.474</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>-1.13</td>
<td>0.56</td>
<td>.114</td>
</tr>
<tr>
<td>V-Rel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>5.65**</td>
<td>1.54</td>
<td>.001</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>9.32***</td>
<td>1.61</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>3.67</td>
<td>1.60</td>
<td>.062</td>
</tr>
<tr>
<td>V-Irrel N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>-0.87</td>
<td>0.39</td>
<td>.074</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>0.17</td>
<td>0.41</td>
<td>.914</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>1.03*</td>
<td>0.41</td>
<td>.035</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA vs. LA</td>
<td>2.04***</td>
<td>0.46</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>AA vs. LD</td>
<td>2.20***</td>
<td>0.48</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>LA vs. LD</td>
<td>0.15</td>
<td>0.48</td>
<td>.945</td>
</tr>
</tbody>
</table>

Note. P = Paraphrasing; V = Visual Representation; Rel = Relevant Information; Irrel = Irrelevant Information; N = Numerical; L = Linguistic
* p < .05, ** p < .01, *** p < .001.

In summary, average-achieving students demonstrated stronger problem-solving processing abilities than both students with LD and LA students in that they paraphrased
and visually represented more relevant information and paraphrased less irrelevant linguistic and numerical information. Students with LD differed significantly from LA students in that they paraphrased less relevant information but also visually represented less irrelevant numerical information. Though mean differences did not achieve statistical significance, students with LD also visually represented less relevant information than their low-achieving peers.

**Question 3: Effects of Problem-Solving Processes on Accuracy**

A hierarchical multiple regression was employed to determine whether the proposed model of the problem-solving process significantly predicted problem-solving accuracy. Before the analysis, it was first necessary to address the assumptions of regression. Specifically, regression assumes: a) linearity, b) independence, c) homoscedasticity, and d) normality. Linearity was assessed by plotting the residuals against the variables of interest. Scatterplots were generated to provide evidence of the linearity of the DV (Keith, 2006). Independence of error was determined by study design. The assumption of homoscedasticity was met by plotting the predicted scores against problem-solving accuracy. Because the variance around the regression line was the same, heteroscedasticity was not a threat to the interpretation of results. Finally, in order to assess the normality of residuals, a histogram of the unstandardized residuals was generated and a normal curve superimposed. Visual analysis confirmed that the data matched the curve and the assumption was met.

To accomplish the purpose of the research question and answer whether the problem-solving process predicted problem-solving performance, students’ problem-
solving accuracy on the MPI was regressed sequentially on ability group, paraphrasing accuracy, and visual representation accuracy. The results of the analysis are shown in Table 4.11.

Table 4.11

Hierarchical Regression: Model Summary Statistics

<table>
<thead>
<tr>
<th>Step</th>
<th>F</th>
<th>ΔR^2</th>
<th>p</th>
<th>sr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Ability</td>
<td>13.51</td>
<td>.250</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Step 2</td>
<td>Paraphrase Relevant</td>
<td>7.54</td>
<td>.169</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Paraphrase Irrelevant Linguistic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paraphrase Irrelevant Numerical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td>Vis Rep Relevant</td>
<td>14.87</td>
<td>.217</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Vis Rep Irrelevant Linguistic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vis Rep Irrelevant Numerical</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Vis Rep = visual representation. sr = semi-partial correlation.

Because the MANOVA showed ability to be significantly related to problem-solving performance, ability was entered in the first block and accounted for 25.0% of the variance in problem-solving accuracy. The omnibus test was statistically significant, \( R^2 = .250, F(2, 81) = 13.51, p < .001 \). The addition of paraphrasing accuracy accounted for an additional 16.9% of the variance in problem-solving accuracy after controlling for the effects of ability. The additional variance accounted for was statistically significant, \( \Delta R^2 = .169, F(3, 78) = 7.54, p < .001 \). Examination of standardized regression coefficients showed that all entered variables were significant at the .05 level, and all values were in
the expected direction (see Table 4.9). Effect sizes were large for all variables except paraphrasing irrelevant numerical information, which had a medium to large effect. Finally, visual representation accuracy was entered third in the model and explained an additional 21.7% of the variance in problem-solving accuracy after controlling for the effects of ability and paraphrasing accuracy. Again, the omnibus test was statistically significant, $\Delta R^2 = .217, F(3, 75) = 14.87, p < .001$. Standardized coefficients revealed that significant positive increases were found with both paraphrasing and visually representing relevant information, and significant negative increases were found with paraphrasing irrelevant linguistic information and visually representing irrelevant numerical information. Effect sizes were medium for the paraphrasing variables and large for both the visual representation variables. Therefore, these findings suggest that students' accuracy in paraphrasing and visually representing information in word problems is an important indicator of problem-solving accuracy. Table 4.12 displays the total effects of ability, problem solving accuracy, and visual representation accuracy on problem-solving accuracy.

**Interactions.** Based on the findings in the by-ability correlation matrices, different bivariate relationships among problem-solving variables existed depending on ability group. Therefore, it was important to probe these differences in the sequential regression model. By adding interaction terms to the model, the variance in problem-solving accuracy of the interaction of ability and each of the six problem-solving process variables was examined after controlling the variance associated with ability, paraphrasing accuracy, and visual representation accuracy.
Table 4.12

Hierarchical Regression Coefficients, Standard Errors, and Beta Values

<table>
<thead>
<tr>
<th>Predictor</th>
<th>b</th>
<th>SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Block 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.28</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Dummy LD</td>
<td>-2.20</td>
<td>0.48</td>
<td>-.50***</td>
</tr>
<tr>
<td>Dummy LA</td>
<td>-2.04</td>
<td>0.46</td>
<td>-.49***</td>
</tr>
<tr>
<td><strong>Block 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.67</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Dummy LD</td>
<td>-1.30</td>
<td>0.49</td>
<td>-.29**</td>
</tr>
<tr>
<td>Dummy LA</td>
<td>-1.11</td>
<td>0.46</td>
<td>-.26*</td>
</tr>
<tr>
<td>Paraphrase Relevant</td>
<td>0.14</td>
<td>0.04</td>
<td>.38***</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Linguistic</td>
<td>-0.66</td>
<td>0.23</td>
<td>-.31**</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Numerical</td>
<td>-0.22</td>
<td>0.10</td>
<td>-.23*</td>
</tr>
<tr>
<td><strong>Block 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.08</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Dummy LD</td>
<td>-0.26</td>
<td>0.43</td>
<td>-.06</td>
</tr>
<tr>
<td>Dummy LA</td>
<td>-0.32</td>
<td>0.39</td>
<td>-.08</td>
</tr>
<tr>
<td>Paraphrase Relevant</td>
<td>0.08</td>
<td>0.03</td>
<td>.22*</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Linguistic</td>
<td>-0.42</td>
<td>0.19</td>
<td>-.20*</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Numerical</td>
<td>-0.07</td>
<td>0.09</td>
<td>-.07</td>
</tr>
<tr>
<td>Vis Rep Relevant</td>
<td>0.16</td>
<td>0.03</td>
<td>.55***</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Linguistic</td>
<td>-0.21</td>
<td>0.24</td>
<td>-.06</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Numerical</td>
<td>-0.35</td>
<td>0.11</td>
<td>-.27**</td>
</tr>
<tr>
<td><strong>Block 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.65</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Dummy LD</td>
<td>-0.43</td>
<td>0.48</td>
<td>-.10</td>
</tr>
<tr>
<td>Dummy LA</td>
<td>-0.86</td>
<td>0.43</td>
<td>-.21*</td>
</tr>
<tr>
<td>Paraphrase Relevant</td>
<td>0.09</td>
<td>0.03</td>
<td>.25**</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Linguistic</td>
<td>-0.43</td>
<td>0.18</td>
<td>-.20*</td>
</tr>
<tr>
<td>Paraphrase Irrelevant Numerical</td>
<td>-0.08</td>
<td>0.09</td>
<td>-.08</td>
</tr>
<tr>
<td>Vis Rep Relevant</td>
<td>0.03</td>
<td>0.06</td>
<td>.11</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Linguistic</td>
<td>-0.23</td>
<td>0.23</td>
<td>-.07</td>
</tr>
<tr>
<td>Vis Rep Irrelevant Numerical</td>
<td>-0.40</td>
<td>0.11</td>
<td>-.31***</td>
</tr>
<tr>
<td>DummyLD * Vis Rep Relevant</td>
<td>0.21</td>
<td>0.07</td>
<td>.39**</td>
</tr>
<tr>
<td>DummyLA * Vis Rep Relevant</td>
<td>0.13</td>
<td>0.06</td>
<td>.27</td>
</tr>
</tbody>
</table>

Note.

$R^2 = .250$ for Step 1; $\Delta R^2 = .176$ for Step 2; $\Delta R^2 = .217$ for Step 3 ($p < .001$).

* $p < .05$, ** $p < .01$, *** $p < .001$. 
After problem-solving accuracy was regressed on ability, paraphrasing accuracy, and visual representation accuracy to address the first part of the research question, each cross-product term (i.e., ability x paraphrasing relevant information, ability x paraphrasing irrelevant linguistic information, ability x paraphrasing irrelevant numerical information, ability x visual representation of relevant information, ability x visual representation of irrelevant linguistic information, and ability x visual representation of irrelevant numerical information) was added to the model separately in the fourth block. Though each interaction was added to the model to test the possible interaction between the problem-solving variable and ability in its effect on problem-solving accuracy, only those with significant interactions were included in the model. Table 4.13 shows a summary of the interactions and their contribution to the overall variance explained when entered as a fourth block in the previously-discussed model.

<table>
<thead>
<tr>
<th>Interaction</th>
<th>F</th>
<th>ΔR²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD * P-Rel</td>
<td>0.03</td>
<td>.000</td>
<td>.968</td>
</tr>
<tr>
<td>LA * P-Rel</td>
<td>0.61</td>
<td>.006</td>
<td>.544</td>
</tr>
<tr>
<td>LD * P-Irrel L</td>
<td>0.26</td>
<td>.003</td>
<td>.776</td>
</tr>
<tr>
<td>LA * P-Irrel L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD * P-Irrel N</td>
<td>4.47*</td>
<td>.040</td>
<td>.015</td>
</tr>
<tr>
<td>LA * P-Irrel N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD * V-Rel</td>
<td>0.68</td>
<td>.007</td>
<td>.509</td>
</tr>
<tr>
<td>LA * V-Rel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD * V-Irrel L</td>
<td>0.30</td>
<td>.003</td>
<td>.744</td>
</tr>
<tr>
<td>LA * V-Irrel L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD * V-Irrel N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA * V-Irrel N</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. R² = .250 for Step 1; ΔR² = .176 for Step 2; ΔR² = .217 for Step 3 (ps < .001). P = Paraphrasing; V = Visual Representation; Rel = Relevant Information; Irrel = Irrelevant Information; N = Numerical; L = Linguistic.
* p < .05.
Of the 12 interaction terms entered into the fourth block, only one significantly explained additional variance in problem-solving accuracy. Ability, paraphrasing accuracy, visual representation accuracy, and the interaction of visual representation of relevant information and DummyLD combined to account for 67.5% of the variance in problem-solving accuracy. The omnibus test was statistically significant, $F(10, 73) = 15.18, p < .001$. The addition of the interaction variable accounted for an additional 4% of the variance in problem-solving accuracy, after controlling for the effects of ability, paraphrasing accuracy, and visual representation accuracy, $\Delta R^2 = .040, F(2, 73) = 4.47, p = .015$. See Figure 4.2 for the scatterplot of the interaction. An examination of the standardized coefficients showed that for each $SD$ unit increase in the visual representation of relevant information for students with LD, expected problem-solving accuracy increased by 0.39 $SD$ units (i.e., $\beta = 0.39$), after controlling for the effects of ability, paraphrasing accuracy, and visual representation accuracy. This increase is statistically significant, $t (82) = 2.99, p = .004$; the effect size was large. Likewise, for each $SD$ unit increase in the visual representation of relevant information for LA students, expected problem-solving accuracy increased by 0.27 $SD$ units (i.e., $\beta = 0.27$), after controlling for the effects of ability, paraphrasing accuracy, and visual representation accuracy.

The interaction term for visual representation of relevant information and DummyLA did not reach statistical significance, $t (82) = 1.98, p = .052$; however, it should be noted that, according to Keith’s (2006) standards for educational research, the size of the effect was large. See Figure 4.3 for a visual of the hierarchical regression model of problem-solving accuracy.
Summary. The hierarchical multiple regression analysis was utilized to answer the third research question by a) determining whether the proposed model of the problem-solving process significantly predicted problem-solving accuracy and b) testing whether there were any differential effects of the problem-solving variables on problem-solving accuracy depending on ability. Ability (i.e., AA, LA, LD) was entered first and accounted for significant variance in problem-solving accuracy.
Figure 4.3. Hierarchical multiple regression model of the predictors of problem-solving accuracy. (VR Rel Info = visual representation of relevant information)

Paraphrasing accuracy (i.e., paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information) was entered in the second block and contributed an additional significant amount of variance after controlling for ability. In the third block, visual representation accuracy (i.e., visual representation of relevant information, visual representation of irrelevant linguistic information, visual representation of irrelevant numerical information) was entered, and contributed significantly to the variance added after controlling for ability and paraphrasing accuracy. Finally, the one significant interaction found (i.e., ability x visual representation of relevant information) was entered alone in the fourth block, and added a small but significant amount of variance in problem-solving accuracy after controlling for the previous three blocks. Standardized coefficients showed that LD (i.e., DummyLD), paraphrasing relevant information, paraphrasing irrelevant linguistic information, visual
representation of irrelevant numerical information, and the interaction of LD and visual representation of relevant information were all significantly different; effect sizes ranged from medium large to large (all $\beta$s > .20).
Chapter 5

Discussion

Research has consistently revealed that students of various ages and abilities struggle to solve math word problems (e.g., Fuchs & Fuchs, 2002; Gonzales et al., 2008; NCTM, 2000). Yet, little research has focused on the processes that precede problem-solving performance. Mayer’s (1985) model of problem solving posits that a problem must first be translated into a comprehensive understanding (i.e., paraphrasing), and then integrated into a conceptual schematic (i.e., visual representation) before the problem can be accurately solved. By focusing on these phases of the problem-solving process, this study addressed two previously uninvestigated ideas: how paraphrasing and visual representation are associated with problem-solving accuracy, and whether differences exist between students with LD and low-achieving students based on their problem-solving processes (i.e., paraphrasing and visual representation). As prior research focusing only on performance has frequently found no differences between the two groups, it was important to determine whether this conclusion applied only to performance or to students’ processing skills as well. The inclusion of an average-achieving group permitted the examination of the problem-solving process across a range of abilities as well as a reference to which the students with LD and LA students could be compared.

Conclusions drawn from the analyses are organized by research question. First, implications of the bivariate relationships among problem-solving processes and problem-solving accuracy are discussed, both across and within ability groups. Next, ability group differences on the seven DVs are considered. Finally, the effects of the
problem-solving processes on problem-solving accuracy are discussed, after which the role of ability as a moderating factor is examined. After the discussion of results of the three research questions, the subsequent section will address the limitations of the study. Next, the implications for instruction based on the results are discussed. Finally, directions for future research are outlined.

**Question 1: Relationships among Problem-Solving Variables**

The first research question examined the relationships among the seven dependent variables in order to make clear the magnitude and direction of the various relationships. Pearson product-moment correlations were used to answer this question; four correlation matrices were generated separately for all ability groups, the AA group, the LA group, and the LD group.

The correlation matrix that combined the three abilities generated the hypothesized outcomes for all relationships of interest. All processes in relation to problem-solving accuracy were significant; that is, when correlated to problem-solving accuracy, paraphrasing relevant information was positive, both paraphrasing irrelevant linguistic and numerical information were negative, visual representation of relevant information was positive, and visual representation of both irrelevant linguistic and numerical information were negative. These results partially confirm the hypothesis that there would be a stronger negative relationship between irrelevant numerical information than irrelevant linguistic information and problem-solving accuracy. That is, the paraphrasing phase did not support this hypothesis but the visual representation phase did. This disparity may be due to students’ notions that inclusion of irrelevant information in the paraphrase adds context but in the diagram adds clutter. Additionally,
students may become more accurate in their determinations of relevancy as they get closer to actually solving the problem. By the second phase of the problem-solving process, students are more familiar with the problem and may have sorted out previous confusions. Also, as they generate a visual representation students are forced to decide how each element fits into the overall concept. This activity may, in itself, help to discard unnecessary or erroneous information.

As expected, the relationship between paraphrasing and visual representation of relevant information was positive and significant with a medium effect size; the relationship between paraphrasing and visual representation of irrelevant numerical information was also positive and significant, with a large effect size. This is reflective of the connection between these two constructs; as paraphrasing is the first step in the problem-solving process, it would be expected that the information identified as relevant would be carried over to the subsequent step in the process, visual representation. In line with this theory, it was expected that the relationship between paraphrasing and visual representation of irrelevant linguistic information would also be significant and positive. It was in a positive direction but not significant; upon further inspection, this result may be explained by a) the low range within the two variables (paraphrasing irrelevant information, range = 4; visual representation of irrelevant linguistic information, range = 2), and b) the possibility that most students would recognize drawing a linguistic feature of the problem as unbenefficial in ultimately solving the problem.

Overall, the relationships of the problem-solving variables to problem-solving performance substantiate Mayer’s (1985) problem-solving model, where problem translation (i.e., paraphrasing) and problem integration (i.e., visual representation) are
essential precursors to accurately solving a problem. Additionally, the significant negative correlations between the irrelevant processes and problem-solving accuracy support the research on the detrimental effects of irrelevant information and the need to teach students to discriminate between what is important and what is not (Englert et al., 1987; Parmar et al., 1996). The significant positive correlation between paraphrasing and visually representing irrelevant numerical information validates the hypothesis put forth by Passolunghi and colleagues (2005) that irrelevant information would “interfere with the [accurate] construction of a mental model of the problem” (p. 735). In fact, even within individuals this pattern was evident. Figure 5.1 is an example of how a misinterpretation of the problem influenced the visual representation generated, and in the end affected the accuracy of the solution.

![Important Info: You must complete a 3 mile run, Jim is 4 meters ahead of Tom and Peter, they ... because Tom and Peter is 3 meters behind. Yea, Tom and Peter is 3 meters behind Jim, behind Jim, and Peter can run 10 miles in an hour.](image)

*Figure 5.1.* Effects of the misinterpretation of the problem information.
Overall, when ability groups are aggregated, there is a strong positive relationship between what students do in the paraphrasing phase and what they do in the visual representation phase. Mistakes or misinterpretations made initially in the problem-solving process appear to be carried through to affect subsequent phases, and ultimately, solution accuracy.

However, while the information garnered from an aggregated sample provides a broad scope of relationships in the problem-solving model, it is also important to understand the bivariate relationships within each ability group. When examined in this way, several distinctions were evident. First, neither AA students’ paraphrasing nor visual representation of relevant information was significantly correlated to problem-solving accuracy. This was surprising, as a general hypothesis was that these variables would be strongly related to accuracy. Instead, all significant correlations with problem-solving accuracy were negative correlations with irrelevant information. This result suggests that the ability to discriminate between important and unimportant information is, for average achievers in particular, a crucial factor in performance. Conversely, for both students with LD and LA students the opposite is true. That is, neither of the latter two groups showed any significant relationships between problem-solving accuracy and irrelevant information. Instead, visual representation of relevant information was significantly and positively correlated with accuracy. For these groups then, the ability to include all relevant information schematically in a visual representation is of utmost importance. For all three ability groups, the directions of relationships were consistent; however, magnitudes of relationships were not. Passolunghi et al.’s (2005) assertion that irrelevant information would be more detrimental for students with LD/LA students than
AA students may yet be correct, as indicated by the LD group’s poorer performance on problem-solving accuracy as well as their higher rate of identification of irrelevant information, but it may be more predictive of eventual success (or lack thereof) for AA students. This could be explained by AA students’ superior ability to differentiate between relevant and irrelevant information as well as to choose and carry out the appropriate operation; thus, when AA students do identify and depict irrelevant information as important, it would more likely result in an incorrect solution. For students with LD and LA students, their inferior skills at discrimination of information would “muddy” the correlation and create a less linear relationship between the two variables.

As the differences between students with LD and LA students were of great interest in this study, it is important to note the ways in which the two groups differed. First, students with LD showed a more unique profile of correlations than either LA or AA students. Of the 21 unique bivariate correlations made, only three were significant for this group. The positive correlation between paraphrasing irrelevant numerical information and visual representation of irrelevant linguistic information is nonsensical in many ways; no logical reasoning would justify why the irrelevant numbers identified would be positively related to the irrelevant linguistic information drawn. Further confusing the issue is that there were no significant relationships between paraphrasing irrelevant numerical and linguistic information, between visual representation of numerical and linguistic information, between paraphrasing and visual representation of linguistic information, nor between paraphrasing and visual representation of irrelevant numerical information. This outcome seems to best illustrate the non-linear processing of students with LD, where relationships expected were nonsignificant, and unexpected
relationships were present. Another interesting finding of the LD group was the significant positive relationship between visual representation of relevant information and irrelevant numerical information. Though it contradicts the across-groups results, this outcome seems reasonable if the findings of Passolunghi et al. (2005) are extended to visual representations. Specifically, the authors found that students with LD/LA students were more hampered by irrelevant numerical information than both AA students and students with ADHD in their paraphrasing. When generating their visual representations in this study, the more relevant information students with LD included, the more irrelevant numerical information was also included. This lack of discrimination between important and unimportant numerical information is an important indicator of one area of weakness particular to students with LD. Interestingly, LA students showed a similar relationship, but at the previous phase of the problem-solving process. That is, LA students showed a positive relationship between paraphrasing relevant information and irrelevant numerical information. Overall, students with LD not only had fewer significant relationships among problem-solving variables, but the ones they did have were out of sync with the aggregated group results.

Question 2: Differences in Problem-Solving Variables among Ability Groups

The second research question was answered using a one-way between-subjects MANOVA on seven dependent variables: paraphrasing relevant information, paraphrasing irrelevant linguistic information, paraphrasing irrelevant numerical information, visual representation of relevant information, visual representation of irrelevant linguistic information, visual representation of irrelevant numerical information, and problem-solving accuracy. Ability group (AA, LA, LD) was the
independent variable. It was hypothesized that AA students would significantly differ from both students with LD and LA students on problem-solving accuracy, while there would be no differences between the latter two groups. Additionally, it was hypothesized that AA students would have higher scores in paraphrasing and visual representation of relevant information, and lower scores on the irrelevant information variables.

Students with LD were expected to have less discrimination between relevant and irrelevant information, in both the paraphrasing and visual representation phase.

The findings showed that the ability groups differed significantly on the combination of the seven DVs. Univariate effects revealed that all the variables save visual representation of irrelevant linguistic information had significant mean differences by ability. Pairwise comparisons revealed an important distinction between students with LD and LA students that was not apparent in the correlation analysis. Whereas the first analysis showed that LA students and students with LD had poor discrimination abilities as demonstrated by the significant positive relationships between relevant and irrelevant information in paraphrasing (LA students) and visual representation (students with LD), the test of univariate effects showed that significant mean differences existed in the amount of relevant information paraphrased. While LA students were able to identify much of the relevant information in problems, they were unable to distinguish it from the irrelevant information; this lack of distinction separated them from their AA peers who more often suppressed the irrelevant information. An example from one LA student’s paraphrase may clarify why students struggled to discriminate. The following is her response to provide the important information in question 4:
“I just don’t understand why they give informational background that Peter can run 10 miles an hour… *(Because you don’t think you need it?)* Well I’m sure I need it because it’s there, but I just don’t know how to comprehend it... *(Can you tell me the other information you need to use to solve it?)* How the meters and you know… Jim is 4 meters ahead of Tom and Peter is 3 meters behind Jim.”

That it was outside her frame of reference to consider that some information provided could actually be irrelevant to solving the problem may be reflective of a lack of exposure to varied problem types. Additionally, that she gave no merit to her confusion about the information’s use seems to illuminate the poor self-efficacy that many low-achieving students develop after years of failure (Huang, Montague, & Dietz, 2010; Klassen, 2002).

In contrast to the LA students, students with LD identified very little relevant information, but were still unable to discriminate it from the irrelevant information. These results illuminate an additional deficit in students with LD beyond discrimination; that is, the language-based comprehension necessary in math problem-solving may differentially affect these students. According to Plomin and Kovas (2005), the correlation between math and reading disabilities is .53. Therefore, many students with cognitive deficits in math would also demonstrate reading deficits; in fact, it is currently accepted that up to 80% of LD are language-based. The research of Vilenius-Tuohimaa, Aunola, and Nurmi (2008) identified “technical reading skills (i.e. flexible word recognition and decoding skills and the ability to adjust reading method and reading speed to the text at hand)” (p. 410) as a necessary reading skill that was deficient in poor problem solvers. Thus, many of the students with LD in this study may have had poor technical reading skills, which
prevented them from accurately identifying and thus representing the relevant information. Because reading ability was known to be connected to math problem-solving performance, reading ability was built into the design in two ways to ensure that unexpected differences between groups were not skewing the results. Students’ performance on the reading FCAT was analyzed by group and the results showed that, as expected, no differences existed between the LA and LD groups. Additionally, each question was first read to the students before they read it themselves, which also should have minimized language difficulties. That despite these preventive steps deficits were still apparent verifies the presence of some deficit at the language-based (i.e., paraphrasing) phase of problem solving.

A second consideration raised by these results is that students with LD may approach problems in an oversimplified way. For example, during testing, protocol dictated that if students provided an incomplete paraphrase they were asked if they had identified all the important information needed to solve the problem. Even with prompting, students with LD often provided noticeably incomplete paraphrases. For example, several students in the LD group, despite prompting, identified only one of the four units of relevant information in question 6:

- “The 10 meters apart from the, the trees are 10 meters apart, and that’s it. (Any other information you need to know to solve the problem?) No that’s all.”
- “How far would a gardener have to walk… (Anything else that would help you solve the problem? Ok.)”
- “The 10 meters apart which the trees are, and that one bucket is needed to water two trees. (Is that all?) Yes.”
Previous research has identified the tendency of students with LD to extract the numbers, regardless of their meaning, and then choose an operation to apply to them – oftentimes, addition (Montague, 1992; Montague & Bos, 1986). Counter to the research, this behavior was more often displayed by the LA students with their indiscriminant identification of all information (relevant and irrelevant), whereas students with LD often failed to identify the relevant information.

Accuracy in solving a word problem depends on one’s ability to suppress the irrelevant, identify the relevant information and then construct a representation based on that information to solve the problem. Though students with LD and LA students differed in the amount of paraphrasing relevant information, they did utilize the same amount of relevant information in their visual representations, though again this was significantly less than that of AA students. In this phase, it was the irrelevant numerical information that differed between LD and LA groups. The latter used significantly more irrelevant numerical information in their representations than did students with LD. The conjectures put forth in the discussion of paraphrasing differences may also be applicable to the differences found in visual representation. That is to say, in their diagrams, LA students included information without discerning its value to solving the problem. As the example of the LA student’s paraphrase demonstrated, the presence of the number in the problem was synonymous with its importance. In contrast, students with LD used less irrelevant information; however, based on the results of the paraphrasing, it is unclear whether they were able to more accurately discriminate between relevant and irrelevant information or whether they simply identified less information overall.
In interpreting these results, it is important to note that visual representation of relevant information was measured on a scale that focused on schematic representation. Especially in light of the previously discussed paraphrasing results, it is possible that LA students actually did include more relevant information than students with LD but it was not counted due to schematic inaccuracies. As this study did not measure incorrect representation of relevant information, it cannot be stated with certainty whether this was a factor; however, it is a possibility that would explain the varying results of the two groups on paraphrasing and visual representation.

**Question 3: Effects of Problem-Solving Processes on Accuracy**

A hierarchical multiple regression was employed to determine whether the proposed model of the problem-solving process significantly predicted problem-solving accuracy. Then to determine whether ability was a moderating variable, the interactions of ability and each of the problem-solving processes were regressed on problem-solving accuracy in the fourth block of the previously-discussed model. It was hypothesized that both paraphrasing and visual representation would significantly predict problem-solving performance, but that there would be a relationship of varying strength with predictor variables and problem-solving accuracy depending on ability. However, the nature of those differences was not predictable; two contradictory hypotheses were put forth. Results of the analysis showed that ability on its own accounted for 25.0% of the variance in problem-solving ability, that adding paraphrasing accuracy accounted for an additional 16.9% after controlling for ability, and visual representation accuracy added 21.7% to the variance after controlling for the previous two blocks. The omnibus tests were all significant, indicating that each set of variables significantly predicted problem-solving
performance. The standardized regression coefficients provided information on the relationship to the DV of each of the variables entered in the blocks. When paraphrasing accuracy variables were entered after controlling for ability, relevant, irrelevant linguistic, and irrelevant numerical information all significantly predicted problem-solving performance. With the paraphrasing variables, the ability level still significantly predicted performance. In general, the first hypothesis concerning the relationship of paraphrasing accuracy to problem-solving accuracy was upheld. Interestingly, when visual representation accuracy was entered into the model, both ability dummy variables became nonsignificant. Said another way, students’ ability to correctly represent a word problem explained more variance in problem-solving accuracy than did their ability level. This is an important finding for low-achieving students, regardless of disability status. Visual representation is, in essence, a tool, and as such its use can be taught; forming accurate representations demonstrates comprehension of the problem and sets students up to develop an accurate solution path (Silver, 1987). That students can succeed in math-problem solving based on utilizing this strategy and not based on their overall math functioning is an exciting prospect.

Examination of the fourth and final block of the hierarchical regression coefficients sheds more light on the relative importance of each variable when the interaction of ability and visual representation of relevant information was considered. It showed that both paraphrasing relevant and irrelevant linguistic information were still significant, as were visual representation of relevant information and irrelevant numerical information. It is worthy to note that the significance of ability disappeared in the third block, but LA reappeared as significant in the fourth. This suggests that membership in
the LA group negatively predicted problem-solving performance when paraphrasing, visual representation, and the interaction term were considered, but to a lesser degree than was predicted when just paraphrasing accuracy was considered. Thus, membership in the LA group was predictive of performance, but factors such as ability to represent the relevant information and exclude the irrelevant information were more predictive of performance, as indicated by the relative magnitudes of effects. The significance of the interaction term for the LD group had the largest effect size. This is particularly exciting for students with LD, in that even more so than for LA students, their ability to accurately represent the relevant information in math word problems highly predicts problem-solving accuracy. This result strongly supports previous research on the role of visual representation in math problem solving when there is a distinction in type of representation (e.g., Hegarty & Kozhevnikov, 1999; van Garderen & Montague, 2003). The caveat is that there must be a qualitative differentiation among students’ diagrams; this has been the persistent weakness in visual representation studies prior to Hegarty and Kozhevnikov (e.g., Lean & Clements, 1981).

Based on the analysis, neither of the contradictory hypotheses proposed was fully supported. That is, AA students did not demonstrate a stronger relationship between the problem-solving processes and problem-solving accuracy based on their superior skills. The alternative hypothesis stated that if students with LD could manage to generate a schematic representation and thus overcome their deficits in conceptualization of word problems (Englert et al., 1987), they would show a higher correlation between the processes and problem-solving accuracy. This conjecture held true for students with LD
and visual representation of relevant information, but did not extend to LA or AA students nor did it include any other problem-solving variables.

**Limitations**

Though the findings in this study are promising, several limitations should be noted. First, the identification of students with LD is based on district diagnoses. Ysseldyke et al. (1982) found that up to 40% of students were misdiagnosed as either having or not having LD based on strict adherence to the federal definition; thus, there is no guarantee that the district-identified students in this study necessarily reflected the intended definitional parameters. Likewise, students this study identified as LA may have been referred for LD testing in the past but not received services (e.g., their parents would not consent to testing). Therefore they may actually meet the LD criteria in this study but be misplaced in the LA group. This limitation requires cautionary generalizations to students with LD both in research and in practice, but reflects the reality of the current situation in schools across the country where students with and without identified LD are misplaced.

Other limitations to the study are related to the measures. The MPI was modified by adding irrelevant information to six of the nine problems. Though an older population of students was chosen to prevent a floor effect, the added difficulty of the irrelevant information increased the difficulty level beyond that expected, which ultimately reduced the range of the problem-solving accuracy variable. Future research might consider using an older population more competent in the necessary skills.

The study was designed so that students would be prompted to draw a diagram; this decision was made in order to attend to the limitations of previous research in visual
representation (i.e., when representation was voluntary, too few images were created to adequately analyze). However, this brought to light the issue of strategy use. Especially for AA students, if an answer can be obtained without a visual representation, then a required drawing may not include all the relevant elements since it was not actually used as an aid in solving the problem. This may have diluted the positive relationship between visual representation and problem-solving accuracy for some of the AA students.

Also built into the design was that the problems remain in front of students while they paraphrased. Though an effort was made to make eye contact with students and so dissuade them from reading off the page, it was not a preventative measure. However, the alternative was to remove the problems from in front of students and introduce working memory into the model. Because of feasibility issues with testing time and sample size, working memory was instead controlled for by the constant presence of the problem.

Although the aforementioned limitations should be considered, the impact of the study to the field of learning disabilities as well as math problem solving must be acknowledged. First, this study examines a previously-overlooked area in comparing students with LD to their LA peers: that is, processing skills. Too often studies have focused on performance-based differences (or lack thereof) without questioning whether differences may exist elsewhere. This study also contributed to the relatively small body of research on paraphrasing irrelevant information and visual representation. No study was found that examined students’ representations when irrelevant information was included in the problem. The investigation of the problem-solving process illuminates specific directions for differentiated problem-solving instruction in classrooms for students based on abilities.
Implications for Instruction

The results of this study emphasized the importance of paraphrasing and visual representation accuracy on students’ ability to successfully solve math word problems. Because students with LD were compared to LA and AA students, there are instructional implications for special education and general education teachers alike. First, this study has demonstrated that ability level need not define performance in math problem solving. Paraphrasing and visual representation are ultimately strategies that can be explicitly taught to students. The large effect sizes related to these two processes highlight the practical significance of their use in instruction. Paraphrasing and visual representation of relevant information variables were associated with, respectively, .38 and .55 standard deviation unit increases in problem-solving accuracy on the MPI; thus, explicit instruction in these processes should likewise improve student performance on high-stakes tests that include problem solving (e.g., the math FCAT). Though it is difficult to directly apply these effect sizes to FCAT performance due to the collection of skills assessed by the test, implications for improvement remain applicable.

There are currently research-based instructional programs in place that utilize both paraphrasing and visual representation processes; this study reinforces their effectiveness. For example, Solve It!, the cognitive strategy instruction program for solving math word problems (Montague, 2003) uses paraphrasing and visual representation to help students progress through problems. By utilizing the information gained from this study regarding characteristics of students with LD, LA, and AA students, the program can be modified based on students’ proficiency with each phase to allow specific instruction on students’
individual deficits. For example, below are the paraphrases of problem 4 by three students:

- “I would, I would cut out the 10 miles an hour from Peter ‘cause it doesn’t help me, and I would use the, I would also cut out the 3 mile run, and I would use Jim is 4 meters ahead of Tom and Peter is 3 meters behind Jim to find Peter’s location.”
- “Uh, um, Peter can run 10 miles an hour, and Jim, um, um… (Any other important information?) Uh uh.”
- “The runners must complete a 3 mile race. Jim is 4 meters ahead of Tom, and Peter is 3 meters behind Jim. Peter can run 10 miles in one hour.”

Based on their paraphrases, it is apparent that the first student has a clear grasp on identifying and discriminating between relevant and irrelevant information. The second student, maybe due to language deficits, could not identify the relevant information and seemed unable to determine that some information was irrelevant. The third student identified all numerical information in the problem, regardless of its use in solving the problem. Thus, while all students will benefit from instruction, differentiating that instruction based on their needs would increase its effectiveness and bring about greater improvement for each student.

Instruction in the creation of visual representations is not new to math problem solving (e.g., van Garderen, 2007; Willis & Fuson, 1988; Yancey, Thompson, & Yancey, 1989). However, based on the results of this study, that instruction can be tailored to help students differentiate between simply including the information from the problem and actually creating an accurate representation that clarifies the interrelationships among
problem parts. Some research has proposed providing a priori schematic “frameworks” from which students then decide which to use based on problem type (e.g., Jitendra & Hoff, 1996). While these types of studies have proven to be successful, they limit the students’ cognitive processing by having them look for “clues” in order to categorize the problem. Teaching visual representation should instead be taught in a more flexible way so that the strategy can be utilized outside of the constrained circumstances of problem type. The goal here should be to develop conceptual understanding of problem parts and how they relate in order to construct a representative whole. The risk in limiting students to “types” may result in student thinking similar to the previously-discussed student who said, “Well I’m sure I need it because it’s there.” The development of conceptual understanding will bring about the ability to piece together comprehension from the text and generate a representation that reflects that particular problem.

In line with the previous instructional implication, the results of this study show that for too many students, critical thinking is not a prerequisite for problem solving. Thus instruction that exposes students to problems outside the focus of the structured lesson plan will prevent students from absentmindedly applying a formula to a problem which may not require it. This circumstance was evident with an AA student who dutifully applied the proportional reasoning formula of cross-multiplication to any MPI problem that had three numbers. That the formula was completely unrelated to those problems did not occur to her as she followed the procedure without fail. Her problem-solving accuracy score was 3.

However, while results demonstrated that students’ ability to discriminate between relevant and irrelevant information was a contributing factor to problem-solving
accuracy, the practical relevancy of students’ discrimination abilities may be debated. Most classroom math textbooks focus each lesson on a specific skill and provide practice questions based on that skill (e.g., McGraw-Hill, 2006). Instruction thus amounts to procedural repetitions of the same process; problems rarely include math skills extraneous to the lesson, and even less often include information irrelevant to solving the problem (Massey, Montague, & Fults, 2009; van Garderen, 2008). Consequently, when students are confronted with irrelevant information, they may assume its relevancy based on past experience. This “duping” of students when irrelevant information does occur broaches the debate of traditional math instruction versus project-based instruction that incorporates real-world problems with naturally-occurring irrelevant information (Darling-Hammond et al., 2008). Instruction using real-world problems requires students to apply multiple math skills and be able to conceptualize and distinguish the purposes of each (Fuchs & Fuchs, 2002). While it is necessary to develop the math-based skills of problem solving, the ultimate goal for students is to learn critical thinking and analysis skills that can be applied in the real world regardless of context. Thus, a combination of instruction focusing on traditional textbook-based math skills with project-based problem-solving activities should improve both specific and generalized problem-solving ability.

Finally, the implications for instruction must also acknowledge the changing way in which special education services are provided. With the advent of RTI, both students with LD and LA students will be exposed to research-based accommodations in the different tiers of instruction. Because many of the process variables showed similarities between the two groups, this new approach may increase the support provided to LA
students who fail to demonstrate a discrepancy in IQ-achievement. At the same time, results from this study showed that some differences did exist. First, students with LD were less able to identify the relevant information in problems; this may illuminate deficits in language and comprehension that are not present in LA students. Further, the usefulness of representation for students with LD is differentially greater than for LA students; instruction in representational concepts would be beneficial to both but may compensate for other deficits particular to students with LD.

**Directions for Future Research**

Future research is needed to confirm the relationships found in this study, ideally using a sample more aligned with the population from which it was drawn. Additionally, a model that incorporates working memory may further elucidate the differences between students with LD and LA students. Working memory is a factor that has been included in several models of reading to identify students with LD; its use during the processes of math problem solving may clarify some of the ambiguities found in this study. Future research could also examine the third and fourth phases of the problem-solving process, *solution planning* and *solution execution*, where students determine how they will solve the problem and then actually do it. Relevant and irrelevant information could be analyzed in both these phases to determine whether the information identified in the translation and integration phases are carried through to the latter phases. Because irrelevant information added a higher level of complexity than anticipated, future research could utilize the same population but code irrelevant information based on similarity to features of the relevant information, as Littlefield and Rieser (1993) did. That is, instead of having the irrelevant information reflect the agent, the action, and the
unit, it could reflect only one or two similarities and thus reduce the complexity. Finally, instructional practices that incorporate the processes identified as important (i.e., identifying all relevant information, suppressing irrelevant, creating a schematic representation that shows the relationships among problem parts) should be developed and tested; current programs can incorporate these findings to further strengthen their effectiveness with students. Instructional practices that differentiate between student needs based on areas of deficit are necessary to improving student performance; research that validates effective, parsimonious instruction is needed.
Appendix A: Original Mathematical Processing Instrument (MPI)

The Mathematical Problems on the Mathematical Processing Instrument

1. At each of the two ends of a straight path, a man planted a tree and then every 5 meters along the path he planted another tree. The length of the path is 15 meters. How many trees were planted?

2. On one side of a scale there is a 1kg weight and half a brick. On the other side there is one full brick. The scale is balanced. What is the weight of the brick?

3. A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?

4. In an athletics race, Jim is four meters ahead of Tom and Peter is three meters behind Jim. How far is Peter ahead of Tom?

5. A square (A) has an area of 1 square meter. Another square (B) has sides twice as long. What is the area of B?

6. From a long stick of wood, a man cut 6 short sticks, each 2 feet long. He then found he had a piece of 1 foot long left over. Find the length of the original stick.

7. The area of a rectangular field is 60 square meters. If its length is 10 meters, how far would you have traveled if you walked the whole way around the field?

8. Jack, Paul and Brian all have birthdays on the 1st of January, but Jack is one year older than Paul and Jack is three years younger than Brian. If Brian is 10 years old, how old is Paul?

9. The diameter of a tin of peaches is 10 cm. How many tins will fit in a box 30 cm by 40 cm (one layer only)?

10. Four young trees were set out in a row 10 meters apart. A well was situated beside the last tree. A bucket of water is needed to water two trees. How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?

11. A hitchhiker set out on a journey of 60 miles. He walked the first 5 miles and then got a lift from a lorry driver. When the driver dropped him he still had half of his journey to travel. How far had he traveled in the lorry?

12. How many picture frames 6 cm long and 4 cm wide can be made from a piece of framing 200 cm long?

13. On one side of a scale there are three pots of jam and a 100 g weight. On the other side there are a 200 g and a 500 g weight. The scale is balanced. What is the weight of a pot of jam?

14. A ship was North-West. It made a turn of 90 degrees to the right. An hour later it made a turn through 45 degrees to the left. In what direction was it then traveling?

15. There are 8 animals on a farm. Some of them are hens and some are rabbits. Between them they have 22 legs. How many hens and how many rabbits are on the farm?
Appendix B: Paraphrasing /Visual Representation Accuracy Scale

Paraphrasing – Visual Representation Accuracy Scale

#1 (no change)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (✓ / -)</th>
<th>Represented? (✓ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>At two ends of a straight path a man planted a tree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>and then every 5 meters along the path he planted another tree.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>The length of the path is 15 meters.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How many trees were planted?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Relevant:  
Total Irrelevant:

#2 (irrelevant linguistic)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (✓ / -)</th>
<th>Represented? (✓ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrelevant</td>
<td>A man is building a brick wall.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>Before he begins building he must buy machinery that can lift the heavy bricks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>On one side of the scale there is a 1 kg weight and half a brick.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>On the other side there is one full brick.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>The scale is balanced.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>What is the weight of the brick?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Relevant:  
Total Irrelevant:

#3 (no change)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (✓ / -)</th>
<th>Represented? (✓ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>A balloon first rose 200 meters from the ground,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>then moved 100 meters to the east,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>then dropped 100 meters.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>It then traveled 50 meters to the east,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>and finally dropped straight to the ground.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How far was the balloon from its original starting place?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Relevant:  
Total Irrelevant:
### #4 (irrelevant numerical)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased?</th>
<th>Represented?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrelevant</td>
<td>In an athletics race, runners must complete a 3 mile run.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>Jim is 4 meters ahead of Tom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>and Peter is 3 meters behind Jim.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>Jim can run 10 miles in an hour.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How far is Peter ahead of Tom?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**

### #5 (irrelevant numerical)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased?</th>
<th>Represented?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>From a long stick of wood, a man cut 6 short sticks, each 2 feet long.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>He then found he had a piece of 1 foot long left over.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>The original stick was 2 inches thick at one end, but only 1 inch thick at the other.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>Find the length of the original stick.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**

### #6 (no change)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased?</th>
<th>Represented?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>Four young trees were set out in a row 10 meters apart.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>A well was situated beside the last tree.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>A bucket of water is needed to water two trees.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How far would a gardener have to walk altogether if he had to water the four trees using only one bucket?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**
#7 (irrelevant linguistic)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (√ / -)</th>
<th>Represented? (√ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>A hitchhiker set out on a journey of 60 miles.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>He used a map to calculate the distance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>He walked the first 5 miles and then got a lift from a lorry driver,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>who had to stop for gas part way.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>When the driver dropped him he still had half of his journey to travel.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How far had he traveled in the lorry?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**

#8 (irrelevant numerical)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (√ / -)</th>
<th>Represented? (√ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrelevant</td>
<td>Picture frames are sold for $10.00 each, but the framing they are made from only costs $2.00 per frame.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>How many picture frames 6 cm long and 4 cm wide can be made from a piece of framing 200 cm long?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**

#9 (irrelevant linguistic)

<table>
<thead>
<tr>
<th>Type of Information</th>
<th>Component</th>
<th>Paraphrased? (√ / -)</th>
<th>Represented? (√ / -)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrelevant</td>
<td>A company sells pots of jam in bulk, and needs to know how much each pot of jam weighs individually in order to estimate shipping costs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>On one side of a scale there are 3 pots of jam and a 100 g weight.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>On the other side there are a 200 g and a 500 g weight.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>The scale is balanced.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant</td>
<td>What is the weight of a pot of jam?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total Relevant:**

**Total Irrelevant:**

<table>
<thead>
<tr>
<th>Across Problems</th>
<th>Paraphrasing Score</th>
<th>Visual Representation Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Relevant</td>
<td>/36</td>
<td>/36</td>
</tr>
<tr>
<td>Total Irrelevant Linguistic</td>
<td>/6</td>
<td>/6</td>
</tr>
<tr>
<td>Total Irrelevant Numerical</td>
<td>/6</td>
<td>/6</td>
</tr>
</tbody>
</table>
Appendix C: Parent Consent Form

Problem Representation and Mathematical Problem Solving of Students of Varying Abilities
Investigator: Jennifer Krawec

PARENTAL LETTER OF CONSENT FOR MINORS

Dear Parent:

I am a graduate student conducting my dissertation research under the direction of Professor Marjorie Montague, Ph.D., in the Department of Teaching and Learning at University of Miami. I am conducting a research study to investigate the problem solving abilities of students with varying abilities in math.

Your son/daughter is invited to participate in a research study investigating middle school students’ mathematical problem solving. Your consent is necessary in order for your child to participate in the study, which will involve a total of about 45 minutes of his/her time. If you consent, your child will be asked for his/her permission to participate. Your child's participation in this study is voluntary. If you choose not to have your child participate or to withdraw your child from the study at any time, there will be no penalty (i.e., it will not affect your child's grade). Likewise, if your child chooses not to participate or to withdraw from the study at any time, there will be no penalty. The results of the research study may be published, but your child's name will not be used.

Procedure:

- Your child will be assessed to measure computational ability using a standardized measure of mathematics achievement. Students will be group-tested; this test is expected to take about 15 minutes.
- Your child will also complete 9 math word problems and be asked to paraphrase and visually represent the problem information. It is expected to take about 30 minutes. Paraphrasing will be audio-recorded for accuracy; your child must be audio-recorded as a condition of participation.
- As part of data collection, we will access your child’s records for both ability and achievement data, including previous psychoeducational assessments, individual education plans and FCAT scores.

Although there may be no direct benefit to your child, the possible benefit of your child's participation is that after testing, he/she will receive detailed feedback on his/her performance. We do not anticipate any risks to your child. All information relating to the proposed study will be provided to your child before he/she assents to participate. It will be explained that risk will be minimal, and every effort to protect against risk will be maintained.

Confidentiality will be protected. All records are locked in the Principal Investigator’s (Marjorie Montague) University of Miami designated research office. Only student subject codes will be used to connect audio data to student data files. Student names and other personal identifiers will not be recorded. Your child may request that the tape recorder be turned off at any time during test administration. Audio files will be stored on a password-protected computer in the locked office of the Principal Investigator.
Representación de problemas y resolución de problemas matemáticos por estudiantes con una variedad de habilidades matemáticas
Investigadora: Jennifer Krawec

CARTA DE CONSENTIMIENTO DE LOS PADRES DE HIJOS MENORES

Estimado Padre:

Soy estudiante de postgrado y estoy realizando mi tesis de grado bajo la dirección de la Profesora Marjorie Montague, Ph.D., en el Departamento de Enseñanza y Aprendizaje de la Universidad de Miami. Estoy llevando a cabo un estudio para investigar las habilidades de los estudiantes con una variedad de habilidades matemáticas.

Su hijo/hija ha sido invitado(a) ha participar en un estudio que investiga cómo los estudiantes de la escuela media resuelven los problemas de matemáticas. Necesitamos que usted nos dé su consentimiento para que su hijo/hija pueda participar en el estudio, que tomará cerca de 45 minutos de su tiempo. Si usted da su consentimiento su hijo/hija tendrá la oportunidad de firmar un formulario de asentimiento en el que acepta participar en el estudio. La participación de su hijo/hija en este estudio es voluntaria. Si usted decide no permitir que su hijo/hija participe o si desea retirar a su hijo/hija del estudio en cualquier momento, no habrá penalidades (por ejemplo, esto no afectará las calificaciones de su hijo/hija). Igualmente, si su hijo/hija decide no participar o si decide retirarse del estudio en cualquier momento, tampoco habrá penalidades. Los resultados del estudio de investigación pueden ser publicados, pero no usaremos el nombre de su hijo/hija.

Procedimiento:

- Su hijo/hija será evaluado para medir su habilidad aritmética utilizando una medida estandarizada de rendimiento en matemáticas. Los estudiantes serán evaluados en grupos; se espera que esta evaluación dure cerca de 15 minutos.
- Su hijo/hija también completará una medida de resolución de problemas matemáticos de 9 problemas en la que le pediremos que parafrasee y represente visualmente la información del problema; se espera que esta evaluación dure cerca de 30 minutos. Por cuestiones de precisión, la parafrasis será grabada en cintas de audio. Igualmente, como condición de participación en el estudio la parafrasis de su hijo/hija tendrá que ser grabada en cinta de audio.
- Como parte de la recopilación de datos, tendremos acceso a los registros de su hijo/hija para obtener información sobre su habilidad y rendimiento, incluyendo las evaluaciones psicoeducativas realizadas con anterioridad, planes de educación individual y los resultados de las pruebas del FCAT.

Aunque tal vez no haya un beneficio directo para su hijo/hija, el posible beneficio de la participación de su hijo/hija es que después de las pruebas, recibirán información detallada sobre su desempeño. No anticipamos ningún riesgo para su hijo/hija. Toda la información relacionada con el estudio propuesto se le entregara a su hijo/hija antes de que él/ella aceptara participar. También se le explicara que el riesgo es mínimo, y que realizaremos todos los esfuerzos necesarios para protegerlo contra este riesgo.
The information on the audiotapes will be transcribed, and all audio files will be destroyed three years after the study is complete. The results of this research study may be presented at conferences or in publications. Your child’s identity will not be disclosed in those presentations.

If you have any questions concerning the research study or your child's participation in this study, please call me (or Dr. Montague) at (305) 284-2891.

Sincerely,

Jennifer Krawec
Doctoral Student, University of Miami
School of Education
Department of Teaching and Learning

Name of Parent/Guardian (please print)  Signature of Parent/Guardian  Date

Name of Student (please print)

M-DCPS STUDENT ID #: ____________  DATE OF BIRTH: ____________

SCHOOL: ___________________________ GRADE: _________

MATH TEACHER: _____________________ PERIOD: ________

YES. I want my child to participate in this study, and I understand that my child will be audio-recorded in this study.  YES ____

NO. I DO NOT want my child to participate in this study.  NO ____

If you have any questions about your child's rights as a subject/participant in this research, or if you feel you or your child have been placed at risk, you can contact the Human Subjects Research Office, at (305) 243-3195.
La confidencialidad será protegida. Todos los registros se mantienen bajo llave en la oficina de la investigación ubicada en la Universidad de Miami y que ha sido asignada a la Investigadora Principal (Marjorie Montague). Los archivos de audio serán almacenados en una computadora bajo clave de seguridad que estará a su vez a resguardo en la oficina del Investigador Principal. La información de las cintas de audio será transcrita, y todos los archivos de audio serán destruidos tres años después de que el estudio haya finalizado. Los resultados de este estudio de investigación pueden ser presentados en conferencias o publicaciones. La identidad de su hijo/hija no será divulgada en dichas presentaciones.

Si tiene alguna pregunta en relación con este estudio de investigación o con la participación de su hijo/hija en este estudio, por favor llámeme (o a la Dra. Montague) al teléfono: (305) 284-2891.

Atentamente,

Jennifer Krawec
Estudiante de Doctorado, Universidad de Miami
Escuela de Educación
Departamento de Enseñanza y Aprendizaje

<table>
<thead>
<tr>
<th>Nombre del padre/Guardián (por favor escriba con letra de imprenta)</th>
<th>Firma del padre/Guardián</th>
<th>Fecha</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Nombre del estudiante (por favor escriba con letra de imprenta)</th>
</tr>
</thead>
</table>

ID del ESTUDIANTE en M-DCPS: ___________ FECHA DE NACIMIENTO: ________

ESCUELA: _________________ CURSO: ________

MAESTRO: _________________ PERIODO: ________

SÍ. DESEO que mi hijo/hija participe en este estudio, y comprendo que mi hijo/hija será registrado en audio para el estudio. SÍ ______

NO. NO DESEO que mi hijo/hija participe en este estudio. NO ______

Si tiene cualquier pregunta sobre los derechos de su hijo/hija en relación a su participación en esta investigación, o si piensa que su hijo/hija puede estar expuesto a un riesgo, puede contactar con la oficina Human Subjects Research Office, en el teléfono (305) 243-3195.
Appendix D: Youth Informed Assent

Problem Representation and Mathematical Problem Solving of Students of Varying Abilities
Investigator: Jennifer Krawec

STUDENT INFORMED ASSENT FORM

Purpose:
We are doing a study about how middle school students solve math word problems. A study is a way to learn more about people. If you decide to participate in this study, you will be asked to do two sets of math problems, one set on computation and one on math problem solving. This will take place over two sessions and is expected to take about 45 minutes for both sessions.

Procedure:
There are some things about this study you should know. As part of data collection, we will check your student records for both ability and achievement data, including previous psychoeducational assessments, individual education plan and FCAT scores. You will be given a computation measure that will show us how well you add, subtract, multiply, and divide. It usually takes about 15 minutes to complete. You will also meet one-on-one with us to solve nine math word problems. This session will take about 30 minutes. On it, you will be asked to paraphrase and visually represent the word problems. Your paraphrasing will be audio recorded for accuracy. After you finish, we will talk about how you did. Altogether, this study will take about 45 minutes of your time.

Risks/Benefits:
We think there will be little or no risk to you if you choose to be in this study. We will take every effort to protect you against any risk. Not everyone who takes part in this study will benefit. A benefit means that something good happens to you. We think you might get the opportunity to receive feedback on your problem solving performance.

Alternatives:
You do not have to be in this study if you do not want to be. If you decide to stop after we begin, that’s okay too. No one will be mad at you if you decide not to do this study and your standing at school will not be affected. You may ask questions about the study at any time.

Confidentiality:
All data collected are confidential which means your answers are kept private. Only student subject codes will be used to connect audio data to student data files. Your name and other personal identifiers will not be recorded. You may request that the tape recorder be turned off at any time during test administration. Audio files will be stored on a password-protected computer in the locked office of the Principal Investigator. The information on the audiotapes will be transcribed, and all audio files will be destroyed.
three years after the study ends. When we are finished with this study we will write a report about what was learned. This report will not include your name or that you were in the study.

<table>
<thead>
<tr>
<th>Name of Student (Please print)</th>
<th>Signature of Student</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Person Obtaining Assent</th>
<th>Signature of Person Obtaining Assent</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

M-DCPS STUDENT ID #:____________ DATE OF BIRTH:____________

SCHOOL: ________________ GRADE: ______

Principal Investigator: Marjorie Montague, Ph.D. Phone: 305-284-2891

<table>
<thead>
<tr>
<th>YES. I want to participate in this study, and I understand that I will be audio-recorded in this study.</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO. I DO NOT want to participate in this study.</td>
<td>NO</td>
</tr>
</tbody>
</table>
Appendix E: Coding Guide for Paraphrasing and Visual Representation

1 pt. (sc.2)

3 pts (sc.5)
(15 m path; every 5 m.)

"every 5 m on the path" 1 pt (sc.13)
"the 5 m along the path" 1 pt (sc.16)

1 pt for each end
- 1 pt for 15 path
- 0 pts for each 5 m.

"the 15 m long" 1 pt (sc.21)
"2 ends of a straight path" 1 pt (sc.39)

NOTE: - 0 pts. for saying wrong #s (eg. #9, "a 50g and a 20g weight")
- any mention of the irrelevant info is a pt, even if they don't have the entire concept (eg. #8, "they're 110") = 1 pt
- if #s aren't labeled, then 0 pts.
- incorrect units are correct (eg. #9, "4 mm x 6 mm") unless the incorrect unit is in the question (eg. #4, "Jim is 4 miles ahead of Tom") 4 pt.
"It is 1 kg, half a brick, and one full brick"  
- 0 pts (sec 2)

"and the other side is full" 0 pts (sec 4)

- 1 pt for balanced
- 1 pt for full brick
- 0 pts for left side of scale (sec 6)

- if no scale or illustration of balance, 0 pts (sec 10)

- 2 pts - balanced (sec 26)
- full brick

- 1 pt for balance (sec 29)
- 1 pt for full brick
- 0 pts if 1 kg weight not labeled

- "the 1 kg and half a brick, and the full brick" 0 pts (sec 40)
"3 miles and 40 miles an hour"
- 2 pts

- all 6 pts. (se 7)

100 m
100 m
50 m

9100 9200 0 pts. (se 10)

9100 950

"if student says only the 100 m east and 50 m east, they get the 200 m up and 100 m down as automatic 2 pts. They must still state the question and say the balloon dropped to the ground to get those points. (se 16)

100 m east
300 m

1 pt for 100 m east
1 pt for dropped 100
1 pt for 50 m east (se 31)

"and to solve where it started off" 0 pts (se 24)

#3
"I need the 4m ahead of Tom" 0 pts
(needs Jim's name)

- In any order, 0 pts unless numbers are included

Jim  Peter Tom  - 2 pts

4m
"The two feet, the six short sticks, (sec 31)
and the one foot long left over.

- 0 pts for 2 ft
- 1 pt for 6 sticks
- 1 pt for 1 ft long left over
"That the trees are in a row 10 m apart." 0 pts (SC.6)

"A well is all the way at the end." 1 pt (SC.16)

"How far would the gardener have to walk altogether?" 1 pt (SC.27)

"He has to water all 4 trees using only one bucket." 0 pts. (SC.24)

1 pt for trees 10 m apart (SC.84)
"he still had halfway to go" 1 pt (sc 4)
"how far did he travel?" 0 pts. (sc 4)
"he walked the first 5 miles" -1 pt. (sc 6)
"the 60 miles" Opt. (sc 8)
"How many miles-60" Opt. (sc 10)

5 miles ☐ ☐

60 miles

"journey of 60 miles" 1 pt.
"he stopped 1/2 way from 60 miles" 1 pt
Journey of 60 miles
5 miles

half = 30 miles

5 miles

1/2 of 60 miles before dropped off

Opt. for 60
Opt. for 5

1 pt for 5
1 pt for 1/2 of journey to travel Opt. for 60 m journey (sc 6)
the 6cm, the 4cm, the 200cm' - 0 pts (sec 4)

1 pt for 6x4
1 pt for =200

The 6cm 2 pts. (sec 21)

6cm = 200cm

6cm long and 4 inches wide and 300cm long' 
-1 pt. for 614
-1 pt. for 200cm long

6cm 0 pts (sec 30)
#9

**#7**

- 60 miles walked 5 miles = 55 left 0 pts. (sec 25)

- "how much each pot of jam weighs individually:" 1 pt. for relevant
  (as long as nothing re: shipping costs)

- "3 pots of jam and a 100 g. weight": somewhere in the paraphrase must distinguish sides of scale. If not, 0 pts. (sec 13)

  - if no scale is drawn or no way to show balance, 0 pts.
  - eg. 100g = 500g 200g → shows balance

  - 100g 100g 200g & 500g → no scale, 0 pts. (sec 26)

  - 20g 50g → 0 pts (sec 32)

#9
References


