Masters Theses 1911 - February 2014

1997

Mental arithmetic skill and its relation to complex mathematical problem solving ability.

Loel N. Tronsky
University of Massachusetts Amherst

MENTAL ARITHMETIC SKILL AND ITS RELATION TO COMPLEX MATHEMATICAL PROBLEM SOLVING ABILITY

A Thesis Presented

by

LOEL N. TRONSKY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 1997

Psychology

MENTAL ARITHMETIC SKILL AND ITS RELATION

TO COMPLEX MATHEMATICAL PROBLEM SOLVING ABILITY

A Thesis Presented

by

LOEL N. TRONSKY

Approved as to style and content by:

James M. Royer, Chair

Carole Beal, Member

Hariharan Swaminathan, Member

Melinda Novak

Melinda Novak, Department Head
Department of Psychology

I would like to acknowledge my wife Monica M. Tronsky. Through her continued love and support all of this was made possible.

ABSTRACT

MENTAL ARITHMETIC SKILL AND ITS RELATION
TO COMPLEX MATHEMATICAL PROBLEM SOLVING ABILITY

SEPTEMBER 1997

LOEL N. TRONSKY, B.A., DARTMOUTH COLLEGE
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST
Directed by: Professor James M. Royer

This study had two main objectives. The first was to provide the reader with a comprehensive account of how it is that a child is able to develop skills to become mathematically competent. The second objective was to conduct research that furthers the current understanding of how mental arithmetic proficiency is related to higher problem solving abilities in mathematics. Several studies have been conducted to analyze this relationship (e.g., Zentall, 1990; Muth, 1985, Balow, 1964) but the studies often suffer from a number of methodological flaws, the most serious being the assessment of mental arithmetic ability using imprecise paper and pencil measures. In this study, students in grades 5 through 8 from a local middle school were given a mental calculation test via computer and completed a more complex paper and pencil math computation and word problem test. It was found that 1.) basic mental calculation speed was a significant predictor of complex computational and word problem solving ability and 2.) these relationships changed from 5th to 8th grade.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE DEVELOPMENT OF MATHEMATICAL SKILL

Introduction

How is it that children become proficient in mathematics? In order to properly address this question it is necessary to review theory and research in several areas of mathematics, beginning with children's concept of number. One of the first things children must learn to do is count, and to do that they must know that number names and their graphemic representation are used to represent quantities. For instance, children must be able to answer questions like, "What does it mean to have five pieces of candy, and how does one arrive at that conclusion?" The first area that will be examined, therefore, is children's concept of number, followed by an analysis of how this affects children's acquisition of the ability to count. Once the ability to count has been attained, a child can then begin to develop strategies that enable him or her to mentally add, subtract, multiply, and divide numbers. Research in the area of children's strategy use, problem difficulty in arithmetic, and arithmetic error patterns has led several researchers to formulate models that attempt to explain how basic arithmetic facts are stored in long term memory and how these facts are subsequently accessed. Currently three models dominate the research literature. I will explain these models in detail and delineate their strengths and weaknesses.

Once children have the basic building blocks of mathematics skill, namely the arithmetic skills mentioned above, it is important to examine how these fundamental skills in turn affect other more complex problem solving abilities in mathematics. Several researchers (e.g., Zentall, 1994; Muth, 1984; Geary, 1994) have studied this phenomenon and have concluded that mental arithmetic abilities most likely do affect the acquisition of more complex skills in mathematics. More specifically, these researchers state that the ability to "automatically" access arithmetic facts frees up working memory resources that allows a child to focus on more complex aspects of word problem solution. Before reviewing the research on the relationship between mental arithmetic and word problem solving ability, however, it is necessary to engage in a discussion about the processes that are involved in word problem solution, automaticity, and working memory and mental arithmetic. Finally, upon reviewing the relevant research that establishes the relationships among working memory, mental arithmetic abilities, and word problem solving abilities, the methodology and results of a new study conducted by the author concerning mental arithmetic and complex mathematical abilities is reported. Several methodological problems exist in the study of arithmetic proficiency and its relation to word problem solving ability to date, and the study reported in this thesis has addressed

some of those shortcomings. Let me first, however, describe how it is that a child becomes arithmetically competent.

Concepts of Number

Innate Abilities

Several researchers have examined the possibility that the concept of number is pre-linguistic and have determined that humans may have the ability to abstractly represent small numbers immediately after birth (e. g., Starkey & Cooper, 1980; Antell & Keating, 1983; Starkey, Spelke, & Gelman, 1983). Still other investigators claim that infants as young as 5 months old have an understanding of very simple addition and subtraction (Wynn, 1992a) and that these innate abilities are the structure on which later developing number skills are built (Dehaene & Mehler, 1992; Gelman, 1990). While research centered around the understanding of a child's innate mathematical knowledge is interesting, much of it is still controversial and I have mentioned the above studies mostly as an aside. I will instead concentrate on older children's developing concept of number.

Number Words

Before learning to count and hence how to mentally calculate, a child first must learn the number words used in his or her language and must learn to what these number words correspond. A child learns, usually by rote, the number words in his or her language. By 4 years of age a child in the U.S. typically has memorized the number words from one to ten in their correct order (Fuson, 1988). Learning the

number words for numbers above ten proves to be very difficult for children in the U.S. and other countries with European based languages (e.g., Fuson & Kwon, 1992b). The difficulty lies in the fact that the number words (in European based languages) used to represent quantities between 10 and 100 do not correspond to the underlying base-10 number system. In contrast, most Asian languages use number words that do reflect the underlying base-10 system (e.g., Fuson & Kwon, 1991). For instance, compare the number word for 18 in Chinese, translated as ten eight, with the number word in English, eighteen. The number word in Chinese directly flows out of the base-10 system while the English equivalent does not.

After learning the number names and sequence of numbers used in counting, a child must learn the correspondence between number words and the quantities that they represent--the mapping of number words onto number concepts. It is then that a child may begin to learn how to count.

Number Concepts

Children at about three years of age start to use number names when they count (Gelman & Gallistel, 1978). Their first attempts at counting are characterized by a number of things. First of all, it may be noted that they have trouble counting even when the quantity to be counted is minimal. They may even use other labels for quantities such as letters instead of numbers. Two things do seem to be immediately apparent in children's initial mapping of number names to

quantity: they seem to implicitly know that each number word refers to a specific quantity, and that sequence is significant when counting. Before the age of 3, however, children still may not know which number refers to which quantity. Wynn (1992b) has noted that it might take up to a year of counting experience (from ages 2 to 3) for children to be able to consistently map number words on to representations of quantity, and even then are probably only able to do so for small quantities. Gallistel and Gelman (1992) agree, stating that this mapping may still be incomplete for four year-olds for numbers less than ten. It is even more difficult for children in Western countries to learn the number words for quantities greater than 10. As mentioned before, many number words in European based languages provide no clue as to the underlying base-10 system of numbering. As Fuson and Kwon (1991) have noted, this fact also leads to difficulty for children in Western cultures in conceptually understanding the base-10 number system .

From Number Concepts to Counting

Processes in Counting

To be able to count objects successfully a child must know more than numbers, what they represent, and their sequence in counting; they must also be able to perform the processes of tagging and partitioning of groups of objects. Tagging refers to assigning an object one (and only one) number and partitioning refers to the ability to keep track of those objects that have been counted and those items yet

to be counted. Tagging and partitioning are activities that need to be coordinated when counting, and young children usually perform these two processes by either physically moving an object from one location to another or (more advanced) by pointing to the objects as they are counted. Errors in counting most often are a result of an error in the tagging/partitioning coordination (Fuson, 1988) as when a child counts "2, 3, 4," saying "3" without pointing to an object, or pointing to one object while saying the first syllable of the word seven and pointing to another while saying the second syllable of seven. According to Geary (1994), even though children are prone to errors in these two processes, they usually arrive in kindergarten with the ability to count sets of objects, especially smaller sets. The next two critical abilities in counting are the learning of cardinality, the concept that the last number tag assigned is to be used to represent the total quantity of objects present, and the learning of ordinality, that subsequent number names refer to larger and larger quantities.

Cardinality

Children at about 3 or 4 years of age usually do not have an understanding of the concept of cardinality. When a child of this age is asked to count out a set of objects, he or she does so but when subsequently asked how many objects are present he or she will recount them, not understanding the significance of the last word tag used in the first count. Some 4 year-olds can answer the second question

without recounting and seem to show an understanding of cardinality. Often times these children are using the "say the last word rule" and do not really understand the concept of cardinality. If you ask these children to retrieve a certain number of objects from a box for you they will typically display no knowledge of cardinality; they will just grab a handful of items without counting (Wynn, 1990).

A complete understanding of cardinality typically does not occur until about the second grade (Piaget, 1965). Children as young as 4 or 5 can display advanced understanding of cardinality but are often convinced to override this knowledge in light of persuasive perceptual cues such as spreading out the to-be-counted objects (density cue) and comparing them to the same number of objects with less space between them (Beck, 1993). It is only at about age 7 or so that children reliably reject perceptual cues in determining set size and use their knowledge of cardinality exclusively.

Ordinality

There is some evidence (e.g., Bullock & Gelman, 1977) that children as young as 2 1/2 display knowledge of ordinality in experiments where the memory demands of the task are minimal and the sets of objects that are compared are small. The ability to determine the relative quantities of sets of larger numbers happens much later which begs the question, is understanding the ordinality of small sets of

numbers (2 is less than 3) the same as understanding the ordinality of larger sets (13 is less than 14)?

Another point to be made is the interconnectedness of cardinality and ordinality. When older children are given sets of numbers to make ordinal judgments about they typically compare the sets by counting the number of objects in each set and then comparing the counts. A child's ability to do this is tied to his or her understanding of cardinality. To count and compare sets of objects a child must know the quantities of the two sets are represented by their cardinal values and since this ability does not reliably emerge until about 7 years of age, ordinality judgments are also subject to perceptual influences as well. So the ability to make ordinal comparisons of larger numbers develops gradually as does the ability to understand cardinality (Fuson, 1988).

Now that I have touched on all the concepts a child needs to become a proficient counter, I would like to provide an overview of how a typical child's counting ability develops. I would also like to begin to show how a child's emerging counting ability has implications in the acquisition of mental arithmetic skills.

Development of Counting and Its Relation to Arithmetic

One of the most extensive and coherent developmental outlines of counting has been delineated by Fuson (1992). Much of her model pieces together what has been reviewed in

the preceding pages involving children's growing knowledge of number, number concepts, and counting.

In Fuson's model, the first counting sequence level is what is called the String Level and is characterized by children counting by saying a string of often undifferentiated numbers such as "onetwothreefour." The next level is called the Unbreakable List Level during which children begin to differentiate the number words in a string. Following this is the children's pairing of objects to words in a one-to-one correspondence without any understanding of the cardinality of a number. If you ask a child in this stage to count a number of objects, he or she will count from one and pair each number with an object until there are no more objects to be paired. If you ask the child again, he or she will again start from one and pair objects to numbers arriving at the number of objects. Youngsters are soon able to understand the cardinality of numbers by relating the last number word said in a sequence to the cardinal meaning of the group of counted objects (Fuson, 1988). At this sequence level, children are able to add by counting out a number of objects, counting out another set of objects, and then finally counting all of the objects, which is called a counting all adding procedure that will be described in detail later.

The next level is called the Breakable Chain Level in which children are able to count from an arbitrary number word. For example, children can begin to use a "counting all

starting with the first addend method" for arriving at a sum (still tied to concrete objects). This method involves counting out the first addend of a problem and then counting on the second addend from the end of the first addend to find the sum.

The Numerable Chain Level is characterized by children being able to break the association of number names from objects and work with the words alone to solve addition and subtraction problems. During this stage the counting on procedure is used but children can keep a "running tab" to solve the problem instead of using concrete objects. For example, if a child is asked to add 2 + 3, the child can use his or her knowledge of cardinality to immediately start with the value 2 and add on by one until the value of the second addend, 3, is reached, all the while keeping track of the running sum: "2 plus one is three (one), plus one is four (two), plus one is five (three). The answer is five." Keeping track of this running sum can be accomplished in a number of ways including using fingers or auditory patterns in addition to the aforementioned double counting method (Fuson, 1982).

When children reach the final level, what is termed the Bidirectional Chain/Truly Numerical Counting Level, numbers finally take on a sequential and a cardinal meaning at the same time. Children can see all of the pairs of addends that make up the sum of a number and many children see the relationship between addition and subtraction, namely that

they are reciprocal operations (Fuson, 1992). At this level children are able to use the derived fact strategy to solve addition problems. In the following section I will elaborate on the derived fact and other counting strategies alluded to in this section that are used to solve arithmetic problems and will also discuss how these counting strategies develop over time.

Counting Strategies and Mental Arithmetic

Mental Addition

As mentioned earlier, children employ many different addition and subtraction counting strategies, and these strategies can be arranged into a three level developmental framework (Fuson, 1992). A summary of these strategies appears in Table 1 on pages 65 and 66. At the first level, the strategies of concrete counting all (CCA) and counting all starting with the first addend (CAF) are ordinarily used (Fuson, 1992; also Baroody, 1987a). CCA involves counting all of one set of objects beginning with 1, counting all of a second set of objects beginning with 1, and then counting all of the objects together starting with 1. Children as young as 3 years of age may be able to use this counting strategy to add small sets of objects in every day contexts (Geary, 1994). The CAF strategy involves counting out the first addend from 1, and then counting the second addend beginning with the last number stated while counting the first addend. Often children use these strategies with the aid of fingers, but eventually mentally keep track of the operations they are

performing (Baroody, 1984a). In fact, about half of the simple arithmetic problems that kindergartners encounter are solved using a verbal, as opposed to finger, counting procedure (in Baroody 1987b; Geary & Burlingham-Dubree, 1989; Siegler & Shrager, 1984).

At Level II, children invent newer, more efficient ways to add and subtract numbers. The CAF strategy is abandoned for a counting on from the first addend procedure (COF), but this procedure is employed for only a short period of time (Baroody & Gannon, 1984; Baroody, 1987a). During this stage, children have learned about cardinality and can simply take the first addend without counting up to it and directly count the second addend on top of it. While using this strategy does save some mental effort, it involves the same amount of "mental bookkeeping" as the CAF strategy and therefore is short lived, if it is used at all (Baroody, 1987a).

The next strategy that is used may initially appear to be a bit of a regression. It is termed a counting all starting with the larger addend (CAL) (Baroody, 1984b; Baroody & Ginsburg, 1986). It appears to be a regression because children ignore the cardinality of the larger addend, but it turns out to be a more efficient strategy than CAF because the mental bookkeeping needed is minimized as a child must keep track of a double count for a minimum number of steps. This can easily be seen by comparing the mental bookkeeping involved in the COF procedure with that needed for the CAL procedure for a problem such as 1 + 6. In this

problem, the COF procedure requires six double counting steps while the CAL procedure only requires one. The last shortcut strategy developed in Level II of Solution Procedures is the counting on from the larger addend (COL) strategy where children use the cardinal value of the larger addend and simply add on the smaller addend in increments of one (Baroody, 1987a).

Level III is the final level that children reach entitled "Derived Facts and Known Procedures." At this level youngsters may use information from problems they already know how to solve to help them solve somewhat more difficult problems, or may use a memory retrieval strategy to find an answer. To illustrate the first point, let me use an example of a slightly more difficult problem such as 8 + 6. A child may see the tie problem 6 + 6 (it is a tie problem because both addends are the same) embedded in the larger problem 8 + 6 and will be able to solve the problem easily by retrieving the answer to the problem 6 + 6 and then complete the problem by adding the left over 2 to the retrieved sum, 6 + 6 = 12, 12 + 2 = 14.

In this section I have focused mainly on the development of strategies in the realm of addition. I would again like to refer you to Table 1 if a summary or review of these strategies is needed by the reader before moving on to the operations of subtraction, multiplication, and division.

Mental Subtraction

It is interesting to note that strategies much like those used in addition are also implemented in subtraction; for a summary of these strategies you should refer to Table 2 (pages 67-69) keeping in mind that Fuson's three levels of addition described above also apply to subtraction. Children start out using concrete objects for subtraction problems and, according to Carpenter and Moser (1984), use three different types of concrete manipulative procedures. The first is called "separating from" which involves counting out (using objects) the larger number (minuend), then removing objects from that set until the smaller number (subtrahend) is reached, and finally counting out the number of objects that are left to arrive at the answer. The second type of concrete procedure is an adding on procedure where a child first counts the subtrahend of a problem using objects, then adds a number of objects until the minuend is reached, and the child either concurrently keeps track of the number of objects added on or counts the number of objects in the "added on set" at the end of the procedure. The last procedure using manipulatives is termed a matching procedure and it involves using objects to form a one-to-one correspondence between the set of objects representing the minuend, and the set of objects representing the subtrahend. The child will then count the objects in the minuend set that have no corresponding objects in the subtrahend set.

The manipulative strategies are then abandoned for more complex strategies. A counting fingers method is often employed by counting out and holding up a number of fingers to represent the minuend and then folding down fingers as the subtrahend's value is counted out; the remaining fingers represent the answer. This strategy is deserted for a mental procedure that involves the same double counting procedure explained in the discussion of certain addition strategies. This "counting down" technique entails repeatedly subtracting one until the subtrahend is reached, keeping track of the result after each step. Once the subtrahend is reached the child reports the answer (Baroody, 1987a). This procedure is very difficult for children because it involves counting backwards, a much more demanding cognitive process for children than normal counting (Baroody & Ginsburg, 1983). Problems with large subtrahends require children to develop another procedure to reduce the amount of mental bookkeeping needed. Children create a counting on subtraction procedure, or finding the missing addend approach (Carpenter & Moser, 1982). It is very helpful to use this strategy in a problem such as 15 - 12 because a child minimizes the number of times he or she must double count: 1 is 13, 2 is 14, and 3 is 15; the answer is 3.

Another strategy to solve a subtraction problem is to use a complementary addition fact. By retrieving the fact that 4 + 5 = 9, a child might be able to quickly solve 9 - 5 = ? by filling in the missing number in the subtraction

problem with the corresponding value in the known addition problem.

The results of cross cultural research (e.g., Fuson & Kwon, 1992a; Fuson & Kwon 1992b; and Hatano, 1982) has revealed two more subtraction strategies for problems involving minuends that are greater than 10. One strategy is called the down-over-the-ten method. This strategy involves subtracting 10 from the minuend, subtracting that difference from the subtrahend, and then subtracting that difference from 10 to arrive at an answer. For the problem 15 - 7 a child would first solve 15 - 10 = 5. The next step is to take 7, the minuend, and subtract 5 from it leaving the child with 2. Finally, by subtracting 2 from 10 a child would arrive at the correct answer, 8. The second decomposition strategy is called the take-from-the-ten method. For the above problem a child subtracts the subtrahend from 10 and notes the answer, 3. The child then subtracts 10 from the minuend and notes the answer, 5. The two intermediate answers are then added together, 5 + 3, to arrive at the correct answer, 8.

According to Fuson and Kwon (1992a), these two decomposition strategies are used much more frequently by Asian students than by American students. One reason for differential use of strategies in different cultures is, once again, language differences. As was mentioned before, Asian based languages use number names that reflect the underlying base-10 number system that is used. It is therefore

extremely easy for a native speaker of Chinese to use either of the ten's methods mentioned above because subtracting 10 from a teen number simply involves removing the word ten from the beginning of the teen number word in Chinese. It is also interesting to note that American children can be taught to use these strategies (Steinberg, 1985) and that in our educational past young children did use such strategies more often than they do now (Ilg & James, 1951). It appears to be the case, though, that it is a more naturally occurring strategy in Asian cultures than in European based cultures due to the difference in number words.

Derived facts and known procedures strategies are applicable in subtraction as well. For example, students may break a problem such as 13 - 6, into its constituent parts 13 - 3 = 10, and 10 - 3, to find the answer, 7. There is also evidence that children can make accurate judgments about which one of the aforementioned procedures is most prudent to use in a particular situation (Woods et al., 1975).

Mental Multiplication

Multiplication lends itself to a similar type of strategy analysis (see Table 3, pages 70-71), although using the types of counting procedures in multiplication that are used in addition and subtraction can be very time consuming and inefficient. Baroody (1987a) has noted a few of the strategies used in multiplication problem solving called rule governed, informal computing (which seems to be an extension of the counting on process), known combinations, skip

counting, or some mixture of the above. Rule governed solution involves the use of a well learned rule to solve a problem such as n x 0 = 0, n x 1 = n, or n x 10 = n0. Informal computing involves simply starting off with the cardinal value of the first operand and counting all up to the answer, making note of how many operand increments) have been made (i.e., noting how many groups of 4 have been incremented while solving the problem 3 x 4). The known combination, or repeated addition approach, involves a series of addition problems: 3 x 4 is 3 + 3 = 6, 6 + 3 = 9, 9 + 3 = 12. Skip counting, or counting by n, involves skipping over the in-between numbers, in effect going through the multiples of an operand to find an answer: 3 x 4 is 3, 6, 9, 12.

In other research, Siegler (1988b) identified four different types of strategies in his application of the distribution of associations model (described in the next section) to mental multiplication. He identified the following strategies: counting sets of objects which involves making groups of tally marks on a paper and then counting each tally mark (really analogous to a counting all strategy), repeated addition (analogous to Baroody's known combination approach), retrieval, and simply writing down the problem and after no other overt behavior producing an answer (probably also retrieval). Again, some students are able to use derived facts (e.g., Geary, 1994) to help them solve problems such as representing 3 x 4 as the addition of two easy tie problems, (3 + 3) + (3 + 3).

Mental Division

Most of the strategies used to solve problems involving the three arithmetic operations already mentioned are also used in mental division (see Table 4, page 73), although research in the area of division is much more scarce than for the other three operations. Children seem to rely heavily on addition and multiplication when first trying to solve division problems. When presented with a division problem a child may use addition to count up the number of divisors (smaller number) that make up the dividend (the larger number). A strategy called multiplication reference, similar to the addition reference strategy used in subtraction, might be used as well where a child retrieves a complementary multiplication fact to solve a division problem (using the knowledge that $7 \times 9 = 63$ to solve the problem $63 \div 7 = ?$). Obviously a derived fact or decomposition procedure can be used as well. If a child knows that $60 \div 12 = 5$ and that $24 \div 12 = 2$, he or she may be able to use these two facts in determining that the answer to $84 \div 12$ is simply the addition of the answers to the two known problems, or $5 + 2 = 7$.

In solving problems involving any of the four arithmetic operations, when counting procedures have been used many times to solve problems, children are eventually able to rely on the most advanced strategy for solving problems, direct retrieval of an answer from a network of facts. Currently, four models of how basic arithmetic facts are stored in, and are subsequently accessed from memory have been formulated.

It is important to describe these models and evaluate their strengths, weaknesses, and explanatory power. I will do so beginning with Ashcraft's Network Retrieval Model, but first an historic note--a description of the first model of mental arithmetic problem solution.

From Counting Strategies to Memory Retrieval Models

The first noteworthy model for the mental solution of basic arithmetic facts was formulated by Groen and Parkman (1972). Initially these researchers formulated three possible models to explain what caused the variability in response latencies to different basic addition problems of the form a + b. One model proposed that reaction times to solve such problems would be governed by both addends; reaction times would be a function of the time necessary to count from 0 to the first addend and then count the second addend on top of that. In equation form the reaction time would be a + b. A second model under consideration was one that was analogous to the assumption of a counting on from the first addend (COF) strategy. The reaction time would simply be governed by the magnitude of b in the above equation. A third possible model would be that the reaction time would be best predicted by a or b in the equation, whichever addend was smaller, which is essentially equivalent to assuming the use of a counting on from the larger addend (COL) strategy is being used. Groen and Parkman tested first graders, older students, and adults and invariably the data came out in support of the third model mentioned above (COL

model), with the curious exception that tie problems were often solved more quickly than other problems with smaller addends. What Groen and Parkman were basically claiming was that the difference in response times for adults as compared to children was merely due to a change in the adults' speed in using the COL counting procedure.

Ashcraft and Battaglia (1978) were skeptical that the COL model of Groen and Parkman was sufficiently able to account for the performance of older children and adults. To test this they used the same basic addition facts that Groen and Parkman used although they used a different type of task. Where Groen and Parkman used a production task in which subjects had to produce the answer to a problem, Ashcraft and Battaglia used a verification procedure in which subjects were given a problem and needed to identify a given answer as being correct or incorrect by pressing a button. The results of experiments using samples of various age groups (e.g., Ashcraft & Fierman, 1982; Ashcraft & Stazyk, 1982; Ashcraft, 1987; Koshmider & Ashcraft, 1991;) revealed that Groen and Parkmans' model did not predict the data well and that the best predictor was instead some sort of memory retrieval strategy. These findings directed Ashcraft to study how basic number fact representation changes developmentally and led to the formulation of the Network Retrieval Model of Arithmetic.

Common Findings in Mental Arithmetic Studies

Before moving on to Ashcraft's model, it is first necessary to define and explain some well documented effects in research conducted on simple arithmetic (for a good review, see Ashcraft, 1992): the problem size/difficulty effect, error effects, relatedness effects, and strategies of processing. These effects will provide a basis for evaluating the three arithmetic fact models to be described.

Problem Size/Difficulty Effect

This effect simply stated is that problems in addition and multiplication (and certain related problems in division and subtraction) that have larger addends and multipliers and in turn larger sums and products, are more difficult for people to solve as evidenced by longer response times and higher error rates. The robustness of this effect is well documented in research involving all four operations (e.g., Ashcraft & Battaglia, 1978; Campbell 1985; Campbell & Graham, 1985; Siegler, 1987b); when using either response time or error rates (e.g., Miller, Perlmutter, & Keating, 1984; Siegler, 1988b); it holds for both production and verification methods of research previously mentioned (e.g., Geary, Widaman, & Little, 1986; Miller et al., 1984); and it holds across the entire range of ages from kindergartners (e.g., Koshmider & Ashcraft, 1991) to the elderly (Geary & Wiley, 1991). One exception to the problem size effect has been noted and that is what is called the tie problem effect. Problems that have addends that are the same or multipliers

that are the same such as 7 + 7 and 7 x 7 are actually more quickly solved and are less error prone compared to other problems with similar sized components and even many non-tie problems with smaller addends or multipliers. This is why problem size effect was deemed a misnomer and was changed to the problem difficulty effect. These tie problems although large are not difficult so are not grouped with other problems that are difficult.

Error Effects

This effect involves the types of errors that people most often make when performing mental arithmetic. Campbell and Graham (1985) set out to test what proportion of errors on simple multiplication problems fell into each of three categories. One of the categories was named table related errors, errors that were answers to other multiplication problems of one or both of the operands in the original problem (saying 36 to 9 x 6 is a table related error as 36 is an answer to 9 x 4 and 6 x 6). A second category was named table unrelated errors, errors that were answers to other combination of operands (saying 49 to 9 x 6). The third category were miscellaneous errors that were not part of the times tables at all (saying 9 x 6 is 57). One would expect that if errors to multiplication problems were random that 14% would be table related, 19% would be table unrelated, and 67% would be miscellaneous (Campbell and Graham, 1985). In one study subjects in grades 3, 4, and 5 as well as adults were used to conduct error analyses (Graham, 1987). The

results were intriguing: table related errors dominated across grade levels starting at 43% errors for grade 3 and peaking at 79% for adults, table unrelated errors started at a 21% rate for grade 3 and dropped to 14% for adults, and miscellaneous errors started at a 36% rate for grade 3 and declined steadily, bottoming out at 7% for adults. So for these adults more than 90% of their errors were not miscellaneous; they were answers to other multiplication problems. Almost 80% of those table errors were table related errors, while a little over 10% of the errors were table unrelated.

Relatedness Effects

Researchers using the verification method in their studies of simple mental arithmetic have found that subjects have a more difficult time rejecting false problems that have answers that are correct if a different arithmetic operation is performed (e.g., Zbrodoff & Logan, 1986). In other words, people take longer to decide that 7 + 2 = 5 (a correct statement if the plus sign is changed to a minus sign) is incorrect than they do in deciding that 7 + 2 = 11 is incorrect. The relatedness effect is not, however, merely an operation confusion effect, it also can occur within an operation. In multiplication, people are slower and more error prone to judge that false problems are incorrect if the products that are given are table related products (e.g., Zbrodoff & Logan, 1986). For example, people are slower and more error prone in judging that 7 x 4 = 21 is an incorrect

problem than judging 7 x 4 = 18 is an incorrect problem. Both of these types of problems affect accuracy of responses in production tasks as well.

Strategies of Processing

Again we have touched on this notion already that at any given time children (and even many adults) may use a variety of strategies to solve arithmetic problems including several different types of counting strategies as well as memory retrieval. Therefore it is necessary in models of arithmetic processing to have a component devoted to fact retrieval (declarative knowledge) and another component dedicated to strategy use (procedural knowledge).

Ashcraft's Network Retrieval Model

The Model

In its most general sense, Ashcraft's model for basic addition and multiplication fact representation in memory is an organized network of information that can be accessed by a process of spreading activation (Ashcraft, 1992). In more explicit terms, basic math facts are stored in a network that relates parent nodes (e.g. addends or operands) to an "answer" node (my own term), and each of these problem to answer nodes has a strength or degree of accessibility associated with it. Also, problems and answers in near neighbor nodes are associated, with the degree of relatedness of near neighbor nodes being much stronger than that of more distant nodes. The spreading activation that leads to the selection of an answer is triggered by three sources:

addends (or operands), the answer stated in a problem (remember, Ashcraft's model is based on the verification paradigm mentioned above), and the nodes in the network that are activated during the retrieval stage. Spreading activation is a parallel process and leads to different nodes having different degrees of activation in response to a particular problem. Whatever answer node receives the highest level of activation is selected as an answer to the problem and the time to retrieve the answer depends on the accrued activation at the selected node (Ashcraft, 1992). The strength of association of problems with their answers depends on practice on those problems (Ashcraft, 1987). Let us now turn to a discussion of research that delineates the strengths and weaknesses of Ashcraft's model and its predictions.

Evidence/Support for the Model

According to what was mentioned above, the strength and interconnectedness of stored problems in Ashcraft's model are dependent on practice on those problems. Practice is in turn largely dictated by the frequency and order of occurrence of problems in (especially elementary) textbooks. In light of this, Hamann and Ashcraft (1986) hypothesized that response time to arithmetic problems should correlate highly with frequency of occurrence of arithmetic problems in elementary texts. They also hypothesized that the frequency of occurrence of different problems probably remains the same from grade to grade and therefore response time to text

frequency correlations should be larger than correlations of response times from grade to grade.

The results were in line with predictions. It was found (Hamann & Ashcraft, 1986) that arithmetic problems with smaller addends and multiplicands are presented earlier and more frequently than larger problems, with the exception of problems involving 0 and 1 which are presented about as often as larger problems (these problems have been shown to be solved using rules as opposed to retrieval procedures [Baroody, 1984]). It was also determined that problem RT's correlated highly with their frequency of occurrence for students in grades 1, 4, 7, 10, and college (range of correlations -.55 to almost -.70) and that these correlations were significantly greater than correlations of problem RT's between grades. It is not difficult to see that this is one explanation for the problem difficulty effect mentioned above. Problem difficulty is explained as a lack of strength of association between problem and answer nodes that results largely from the lower frequency with which more difficult problems are encountered over the span of a student's schooling.

The model also makes predictions about priming effects. According to the model, whenever nodes are activated the activation is spread out over the network and decays over a short period of time (Ashcraft, 1992). A problem that is presented during this activation period will have its solution reaction time altered by the already activated

nodes. Many studies have shown this (e.g., Lefevre, Bisanz, & Mrkonjic, 1988) and Ashcraft and Koshmider (1991) have shown how (excitatory as well as inhibitory) priming effects change as a function of problem difficulty and age.

Lastly, the model is also able to account for what are called confusion and split effects using an argument similar to the one above that involves spreading activation. In the confusion effect, a false problem such as 6 x 4 = 18 is difficult to reject, as the problem's answer is one multiple away from the correct answer. According to Ashcraft's model both answer nodes 18 and 24 will be highly activated; the node 24 because it is the parent node of 6 x 4 and 18 because of the high degree of spreading activation of a near neighbor node. The attenuation of response time to these problems results because the process of choosing one of the answers over the other is disrupted due to their activation being both high and similar in magnitude. A split effect, the quicker rejection of a false product or sum that is very distant from the correct product or sum (e.g., 6 x 4 = 48), can also be explained via spreading activation. Activation of near neighbor nodes (answers) will be high but a distant answer in a false problem will receive very little spreading activation, if any, and the decision process of choosing an answer will likely not be disrupted. Several studies to date have documented split and confusion effects (e.g., Stazyk et al., 1982; Zbrodoff & Logan, 1986).

Criticisms/Weaknesses of the Model

While Ashcraft's model does adequately explain the above effects including problem difficulty, relatedness, and priming, it fails in three main areas. The first is in its lack of description of a decision mechanism. According to the general description of the model, the overall magnitude of activation of an answer node determines what answer is selected from the many activated nodes. If that is the case how are incorrect answers ever selected? If incorrect answers may be selected how does such a process work? Ashcraft is not explicit about this. The second drawback to the model is its lack of a procedural (strategic) component. Ashcraft (1992) briefly mentioned that in young children two types of answer searches, retrieval and some sort of backup (counting) strategy may be activated simultaneously and may compete in a "race horse" fashion to arrive at a correct answer. Ashcraft's treatment of this procedural component is cursory at best and was part of the impetus for Siegler (1988b) to formulate the model that I will explain in the next section. A final drawback is the verification technique used in all of Ashcraft's studies. While it is helpful in relaying knowledge about things such as split effects and priming effects, to me it seems more important to emphasize the study of production of answers instead--especially when one would like to relate proficiency in the domain of simple mental arithmetic to proficiency in another domain such as word problem solving that almost always calls for a student

to produce, as opposed to verify, answers in order to solve a problem. Furthermore, an implicit assumption in Ashcraft's model is that verification of a correct sum or product involves both producing an answer to the problem and then comparing that to the given answer. Zbrodoff and Logan (1990) have challenged this assumption and have shown that people might process the verification problems as a whole most of the time instead of going through two separate phases.

Siegler's Distribution of Associations Model

The Model

This model was initially developed in 1984 by Siegler and Shrager to represent strategy choices used in children's subtraction, although descriptions of the model appear in many articles (e.g. Siegler, 1988a; Siegler, 1988b; Ashcraft, 1992) where it has been applied to the operations of addition and multiplication as well. In this model a distribution of answers is associated with each arithmetic problem; it is a distribution of the correct answer as well as incorrect answers previously generated. Each time a child encounters a problem and arrives at a solution, the associative strength of that answer to that problem gains in strength relative to the strength of associations of other answers to the problem. If a child uses a counting procedure to find the sum of 6 + 7 and generates an answer of 10, that association, albeit incorrect, will gain in strength relative to other answers associated with the problem. Errors can occur either when a

counting procedure fails or when retrieval accesses an incorrect answer. According to the model, different problems will have different shaped answer distributions. A problem such as 2 + 3 which is fairly simple to compute using a counting strategy will have what Siegler termed a "peaked" distribution (see Figure 1 on page 32 for an example of this distribution). Most of the time children solve this problem correctly and therefore the strength of the association with 5 will be relatively high compared to the strength of association with 4 or 6 or other incorrect solutions. The problem 6 + 7 probably will have a "flat" distribution (see Figure 2 on page 33) as children solve this problem incorrectly more often than problems with smaller addends, and the associative strength to the correct answer 13 will not be as high relative to other incorrect associations (Siegler 1988a; Ashcraft 1992).

The second aspect of the model involves which strategy a child uses to produce an answer. In the original model there were three different ways a child could arrive at an answer for a problem and any of the three could end the process. The three strategies were retrieval, elaboration of representations, and use of algorithms, and they occurred across all distributions of associations. Siegler noted that the retrieval phase of this model was similar to phases in many other memory models (Anderson, 1983; Gillund & Shiffrin, 1984). Retrieval was governed by two parameters, a confidence criterion and a search length, both of which were

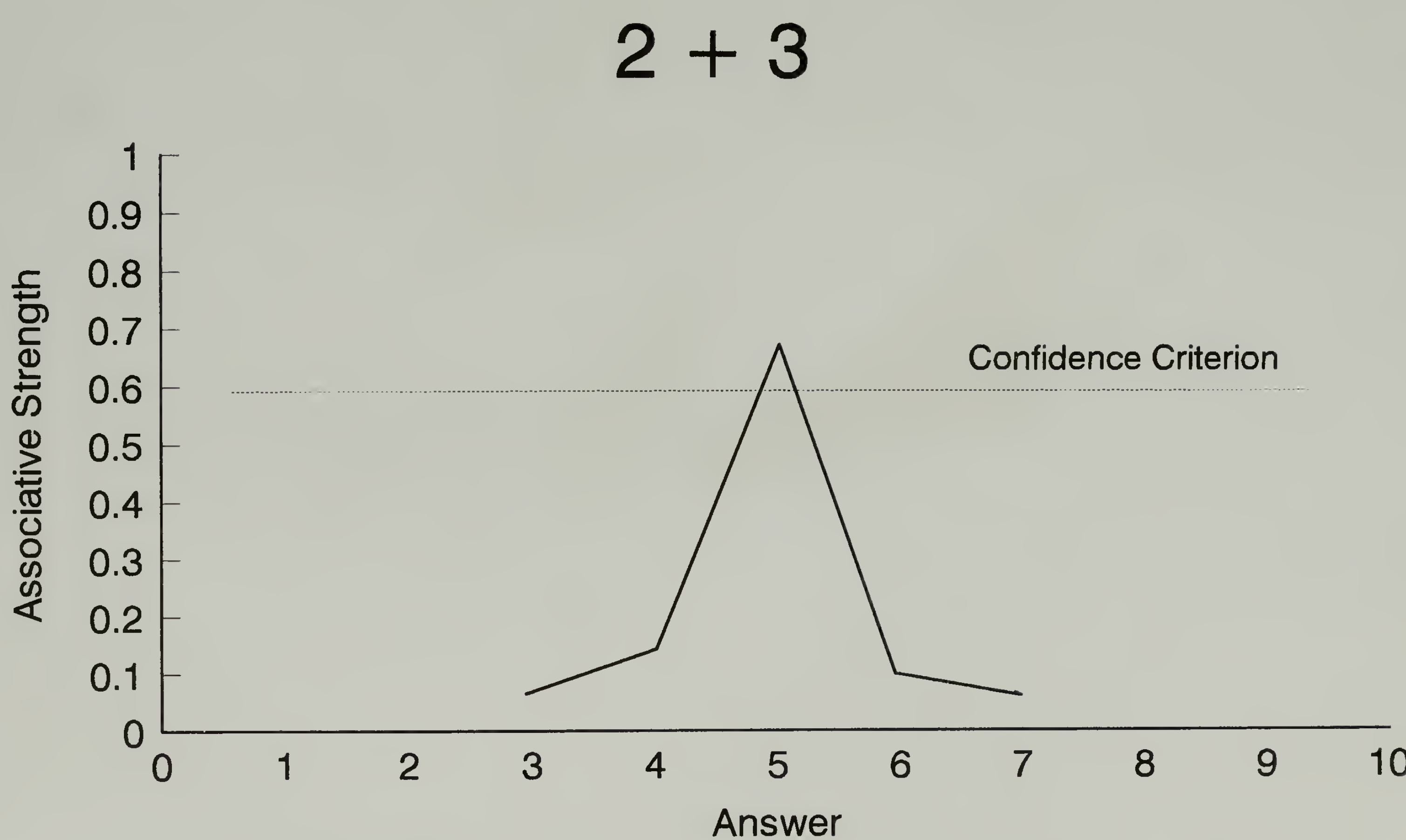

Figure 1. Hypothetical distribution of associative strength and confidence criterion for the simple addition problem 2 + 3.

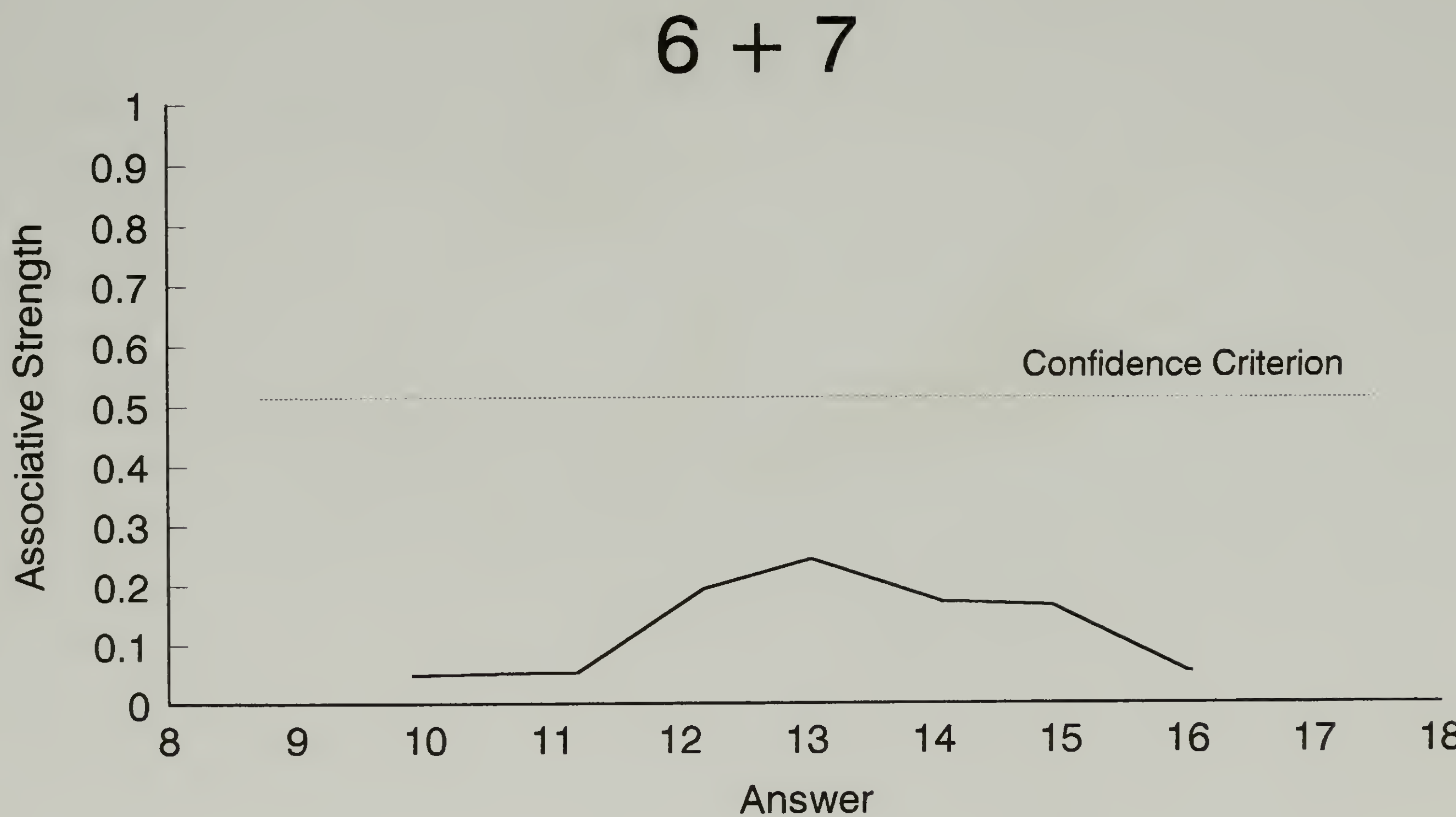

Figure 2. Hypothetical distribution of associative strength and confidence criterion for the simple addition problem 6 + 7.

assigned at the outset of a retrieval attempt. A child would give an answer only if the predetermined confidence criterion was exceeded by the associate strength of the answer retrieved. If the confidence criterion was not exceeded for a retrieved answer, the child would continue to search answers until the predetermined search length parameter had been equaled, in which case the search would be terminated and some sort of counting strategy would be used. Probability of retrieval of any specific answer was based on its relative associative strength compared to other answers. For example, for the problem 5 + 7 a child may have a distribution of associations that looks like this: associative strengths of .4 for 12, .2 for 13 and 14, and .1 for 11 and 10. To explain how the model works, let's assume for this problem a child is using a search length parameter of 3 and a confidence criterion of .35. If the child first retrieves the answer 12, that answer will be verbalized as it is above the level of the preset confidence criterion. If, however, a child retrieves the answer 11, it will not be uttered as its associative strength of .1, is less than .35 (confidence criterion) and another retrieval will take place. If the child then retrieves 14 and 13 on successive searches, no answer will be stated and the child will abandon retrieval for one of the other two strategies because the preset search length (3) has been equaled.

On a percentage of trials where the search length parameter is reached, Siegler (1988b) claimed that children

used something called "sophisticated guessing" where they would make one last retrieval effort and set their confidence criterion to 0 and would report whatever answer happened to be retrieved (although Siegler is not explicit about why this is "sophisticated" I would assume it is because the most probable answer is often going to be the correct answer as its strength and probability of being retrieved on any one trial will usually be greater than the probability of any other single answer). If retrieval did not produce an answer and children did not use the sophisticated guessing approach they could use an elaboration technique where they would write down the problem and retrieve an answer.

The third and final strategy a child could use is the algorithm strategy where one of the counting procedures mentioned previously in this paper would be employed. The specific type of algorithm used would be governed by the properties of the problem.

Updates of Siegler's Model

A more recent version of this distribution of associations model was recently presented by Siegler and Jenkins (1989). Siegler and Jenkins noted that a shortcoming of the earlier model was that it said nothing about how different strategies might be selected; the model merely stated that retrieval was always the first strategy attempted and that the backup strategies were just that, backups when retrieval was not successful. Under the new model any one of the three strategies could be selected at the outset and in

fact more than one strategy would be selected and would compete in a race horse fashion to arrive at the (correct) answer (this is basically the same race horse example explained above in Ashcraft's model, although Siegler's exposition is a bit more detailed. More than likely, both researchers incorporated this parallel search mechanism into their models in response to research [Compton & Logan, 1991] that provides evidence that retrieval and counting strategies are probably simultaneously accessed in trying to solve a problem and one method ends up "winning" and supplies an answer). In the new model not only are problems related to answers, but (counting) strategies are related to individual problems, specific types of problems, and classes of problems as a whole, and these relations become stronger or weaker based on the speed and accuracies that have been recorded on previous trials using these strategies. So, for example, for the problem 2 + 7 three strategies might be used to arrive at an answer: a counting on from larger addend strategy, a counting on from first addend strategy, or a retrieval strategy (I am only using three strategies to keep the example simple). Response time would be faster if the COL strategy were used; one only needs to increment twice while counting using the COL strategy while for COF one needs to increment 7 times. Accuracy also is greater using COL rather than COF because COF requires that a person double count seven times. Hence, the strategy of COL would be strongly associated with this particular problem and problems similar

to it (such as 2 + 8) and would likely compete in a race with retrieval to arrive at a correct answer to the problem.

In summary, according to the updated model the important determinants of strategy choice and arrival at the correct solution are threefold. One determinant is the strength of association between the problem and answer(s), another is a child's preset confidence criterion, and a third is the strength of association between a problem and strategies or procedures that have been used in the past for solving such a problem or similar problems.

Siegler and Jenkins also tackled the problem of new strategies and how they come to gain strength in light of other strategies that are, so to speak "tried and true." They offered an explanation using a term called "novelty points" (Siegler & Jenkins, 1989) that are awarded to new strategies that are used to solve problems. This notion was inspired by Piaget's view (1970) that people like to try out newly acquired cognitive abilities. The novelty points allow a person to try a new type of strategy out even though it has no previous record of accuracy or speed, and in turn no relative associative strength to any problems. Each time a person uses the new strategy, novelty points diminish. In exchange, however, valuable knowledge about speed and accuracy of the new strategy is gained and if the new strategy is fast and accurate, the strength of association (with a particular problem and/or types of problems) that this strategy gains, outstrips the amount of novelty points

that are lost. It is not difficult to see how this might explain how a child comes to use the novel strategy of retrieval after using a counting procedure that had previously been successful. Computer simulations of this new model although not perfect, seem to work very well (Siegler & Jenkins, 1989; Siegler & Shipley, 1995).

Evidence/Support for the Model

Evidence for support of Siegler's model comes from many different pieces of research. For example, Siegler's model does a good job of predicting response time, errors, and solution strategy distributions in research on children's addition (e.g., Siegler & Shrager, 1984), subtraction (Siegler, 1987b Sloboda), and multiplication (e.g., Siegler, 1988b) and also has been successfully applied to children's performance in the domain of spelling (Siegler, 1986). Geary and Burlingham-Dubree (1989) collected strategy choice data in an effort to assess the external validity of Siegler's strategy choice model for addition and found strong support for convergent validity of the model as well as modest discriminant validity. Probably the most impressive evidence that has been compiled for the model is the research that has been conducted largely by Siegler and associates and Geary and associates applying the model to examine performance of subpopulations of children. Siegler (1988a) was able to identify three subgroups of children in an experiment that involved addition, subtraction, and reading measures. Using a cluster analysis procedure he was able to identify good

students (fast and accurate on all tasks), not-so-good students (relatively slow and inaccurate on all tasks), and perfectionists (named so because they had high confidence criterion and even though as fast and as accurate as the good students, they used retrieval to solve problems significantly less often). Achievement tests were also administered and as one would predict good students on average scored significantly higher on the 3 math (computation, problem solving, and total) and the 3 reading (word recognition, reading comprehension, and total) measures. Perfectionists scored nonsignificantly lower on the math measures and significantly lower on reading measures than the good students and scored significantly higher on the math measures and nonsignificantly higher on the reading measures than the not-so-good students.

Other subpopulations have been examined as well. Geary, Brown, and Samaranayake (1991) used the strategy choice model to assess differences in strategy choice between normal and mathematically disabled children. First graders were given single digit addition problems to solve and their strategy use in solving the problems was recorded. The students were again tested one year later to document any changes in their arithmetic strategy choices. It turned out that the normal group increased their reliance on retrieval over time, decreased their reliance on counting to solve problems, and decreased their error rates. The math disabled group, however, showed no change over time in reliance on retrieval

and counting and showed only improvement in their error rate on counting trials. No difference in strategy choice was noted between the two groups at time 1, but the normal group at time 2 used the retrieval strategy more often with fewer errors and used a verbal counting strategy less often than the math disabled group.

A number of cross cultural studies have also been conducted using the strategy choice model to determine differences in mathematical abilities of U.S. and Chinese children (e. g., Geary, Fan, Bow-Thomas, 1992; Geary, Fan, Bow-Thomas, Siegler, 1993). In Geary et al., (1993) addition problems were presented to kindergarten aged students on a computer and accuracy, response time, and strategy data were collected for each student. A numerical memory span test and written addition test were administered as well. Results were that Chinese students used counting strategies more frequently than U.S. children and used a verbal counting procedure significantly more often than U.S. children. U.S. children in turn used retrieval significantly more often than Chinese children, however, Chinese children had a significantly lower error rate on retrieval trials (1%) than U.S. children (33%) did. The Chinese children also had an advantage in verbal counting and retrieval response times as well as a 3 to 1 advantage in the number of correct answers on the speeded written addition test. When taking into account the numerical span data in conjunction with strategy choices of the two groups it was concluded that the initial

advantage for Chinese children in math (arithmetic) is at least partially due to shorter number words in the Chinese spoken language. Shorter number words contribute to a greater memory span and accelerated speed of counting. This in turn allows Chinese children to count out solutions to arithmetic problems more quickly, increasing the probability that the problem and answer are associated in working memory, which in turn leads to a quicker development of the use of a retrieval strategy over subsequent problem presentations.

There are other studies that supply additional evidence to support Siegler's model that will not be considered here but should be noted (e.g., Geary, 1990; Geary & Brown, 1991; Goldman, Mertz, & Pellegrino, 1989; Goldman, Pellegrino, & Mertz, 1988). Furthermore, Siegler's model can explain many of the effects described in the previous section. According to the model, the problem size effect occurs because when counting strategies are being used initially to solve problems it is more likely that a counting error will be made in trying to solve a problem that is larger which requires more counting and mental bookkeeping steps. More errors will lead to a flatter distribution of associations for larger compared to smaller problems. Response times will be greater for larger problems either because no answer will be peaked enough to exceed a person's confidence criterion and therefore a time consuming backup strategy will have to be used, or a person may make several retrieval attempts if many different answers have associative strengths greater than his

or her confidence criterion. If we add to the mix the notion that small problems are encountered more by people in texts and everyday life as Hamann and Ashcraft (1986) would lead us to believe, according to the model the effect will be even more pronounced as larger problems will have even flatter distributions. It should be noted that Ashcraft (1992) is not convinced that such an argument can extend to the problem difficulty effect still existing for adults who largely use retrieval strategies and may have fairly peaked distributions for even difficult problems.

As for the relatedness effect, Ashcraft (1992) also concedes that the model predicts these effects even though no data exists as Siegler has conducted experiments using production instead of verification tasks. It is not difficult to see, however, that the many incorrect answers in a distribution of associations for a problem would be multiples of one or both of the operands and would act as interference in a verification task, just as a problem such as 3 x 6 = 24 would interfere by activation of near neighbor nodes in Ashcraft's model. The split effect (quicker rejection of false problems with answers much larger or smaller than the correct answer) would also be predicted because the strength of incorrect answers immediately surrounding the correct answer peak would be much stronger than incorrect answers far from the correct answer peak.

Siegler's model explains certain error effects well, too. Operand errors in multiplication (e.g., 7 x 8 = 49) can

be explained by the fact that in using a counting strategy a person has either added one too few, or one too many times. This will lead to fairly strong associations of the problem with incorrect answers that are one or two multiples away. Such an argument would also explain why operand errors are usually errors that are very close to correct answer. Another type of mistake when using a backup strategy in multiplication is to add incorrectly, such as adding 7 eight's in the above problem and arriving at an answer of 57. An error of this type is called a nontable error (Campbell, 1987) and although less common than other errors, it does occur and can be explained in Siegler's model.

Criticisms/Weaknesses of the Model

Other types of errors are more difficult for the model to account for, namely table and operation errors. According to Campbell and Graham (1985) table errors occur about twice as much as nontable errors. When an error is committed using a repeated addition strategy the result would most likely involve either adding too few or too many times or adding incorrectly within +/- 2 (Geary, 1994). Siegler (1988b) assumed that when using such a counting strategy, errors should be approximately normally distributed around the correct answer for a problem. Using this assumption one would expect people to make almost twice as many **nontable** errors as table errors (McCloskey, Harley, Sokol, 1991), the complete opposite of what actually happens. Operation confusion errors according to McCloskey et al. (1991) are

difficult for Siegler's model to explain as well, especially when considering a confusion effect where a multiplication answer is given to an addition problem. As these researchers state, how is it that "a child attempting to solve 8 + 5 through the use of a backup strategy arrives at 40?" (p. 389). Siegler has countered this notion by saying that these errors are not a result of backup strategy errors but rather a confusion of related operations. While this might be true, in postulating such an error mechanism it is important to determine why it might be so specific (why would this confusion effect only occur across operations and not within operations as well?).

Lastly, Siegler's model also has trouble explaining the priming effects found by, among others, Koshmider and Ashcraft (1991) and Campbell (1987b). Because each problem has a distribution of answers associated with it, inhibitory or facilitation effects should not be found. If a person has just processed the problem 5 x 9 arriving at the answer 45, according to Siegler's model it should not affect subsequent processing of the problem 6 x 9 that has its own distribution of answers associated with it. While 45 may be one of the incorrect answer associations with 6 x 9, the model does not say anything about how immediately preceding processing of such an answer may affect response latency to a related problem.

Campbell and Graham's Network Interference Model

The Model

The network interference model shares many properties with Ashcraft's and Siegler's models (Campbell, 1987a; Graham, 1987; Graham and Campbell, 1992; McCloskey, et al., 1991). Like Ashcraft's model, Campbell and Graham's model involves activation that is driven by the two operands given in a problem. Unlike Ashcraft's model, however, Campbell and Graham believe that the problem as a whole also drives activation. As in Siegler's model described above, problems can be associated with incorrect as well as correct answers. When counting strategies produce errors those incorrect solutions are associated in the network. There are also associations in Campbell's model that do not exist in the models of Ashcraft and Siegler, answer-answer associations and problem-general magnitude associations. According to Campbell and Graham (1985) answers that share digits (14 and 24 for example) are somewhat associated. The problem-general magnitude association involves the linking of problems to representations that specify the approximate magnitude of the correct answer (e.g., 8 x 4 might be labeled as a small-medium magnitude) and answer nodes may in turn be associated with this approximate magnitude as well.

The idea of a network stems from the assumption that problems that share an operand activate memory structures common to both problems. When a problem is presented, the appropriate operand and problem nodes are activated with the

activation spreading to both correct and incorrect associated answer nodes and general magnitude nodes. Activation then may in turn spread from answer to answer nodes and general magnitude to answer nodes. How an answer is actually selected is not entirely specified in the model. Supposedly the most highly activated answer node is selected, and the latency to make a selection is determined by the degree of activation of competing answer nodes. The interference part of the model occurs as a result of the activation of incorrect answers that compete with the activation of the correct answer. An incorrect answer may be selected or may simply increase response time in selecting a correct answer by reducing the magnitude of the difference between the activation of correct answer and competing answers.

Strength of association between the various nodes is a function of frequency and order of presentation. Problems that are presented more frequently will in turn have (most likely correct) answer associations that have greater strength relative to less frequently presented problems. Order of presentation is also important because problems that are studied first have less proactive interference with which they must contend. In learning answers to one's first multiplication problems there are not many competing interference associations established. When trying to learn problems that are introduced later (larger problems) there is a great deal of proactive interference from the smaller,

already practiced problems that already have many well established associations.

Evidence/Support for the Model

The information that Campbell and Graham (1985) used to formulate their theory was experimental data that many researchers overlooked--error patterns. Graham and Campbell found that errors to simple multiplication problems were systematic, as opposed to random in nature. Errors are largely table related, meaning that they are answers to other multiplication problems involving one of the operands (54 is a table related error to the problem 8 x 6). In addition to errors being table related, Campbell and Graham also discovered that the incorrect responses were clumped very close to the correct answer and found that this effect increased across grade level. Another interesting phenomenon discovered was the error priming effect (Campbell, 1987a). To review, the priming effect refers to a retrieved answer on one problem trial interfering with retrieval of an answer on a subsequent trial; when errors do occur, the probability that that incorrect answer had been retrieved on a recent prior trial is 10% - 20% higher than chance dictates (the priming effect can facilitate or disrupt subsequent trials depending on the time lag between trials, [see Campbell 1987a]). These two phenomena are evidence that retrieval is both problem and operand driven.

A number of interesting predictions are made from Campbell and Graham's model that are alternatives to what

Ashcraft and Siegler offer. As mentioned before, the problem difficulty effect is explained by the other models as occurring as a result of differential presentation, namely easier problems are encountered in elementary texts more often (e.g. Siegler 1988b; Hamann & Ashcraft, 1986). There is no evidence to contradict that this does not occur throughout school and beyond. Campbell and Graham (1985) do not doubt that this is possible but offer the explanation that teaching (learning) order is the culprit. In other words, the memorization of the smaller operand multiplication problems have a profound proactive effect on learning the larger problems that are invariably introduced at a later date (Campbell & Graham, 1985). An engaging experiment was conducted (Graham & Campbell, 1992) to test the proactive effect, in which subjects had to learn a system of alphaplication. The system is basically multiplication problems using letters. The letter combinations (A * I) = (I * A) = x, would be an example of an alphaplication problem. In this system once problems are learned only retrieval can occur--there is no such strategy as counting. Subjects in this experiment first learned a set of letter operand and product combinations and then learned a second set to see if there was a proactive interference effect as measured by increased errors and reaction time to the second set of problems. The results supported the hypothesis that there would be proactive interference and although the scope of the results were somewhat restricted because the experiment only

modeled real multiplication, it is evidence that Campbell and Graham may be correct about proactive interference in multiplication.

Another intriguing interpretation formulated by Graham and Campbell (1992) is their explanation of tie problems. In multiplication and addition, problems with repeat operands or addends (6 x 6, 7 + 7) are processed significantly faster than nonties by children and adults, even ties that have relatively large digits. Again, the aforementioned models indicate this is due to greater practice on the problems as a result of the significantly greater presentation of these problems in text books (e.g. Siegler, 1988b; although contrary findings in Hamann & Ashcraft, 1986). This is feasible. The explanation offered by the network interference model, however, that there is differential interference in tie and nontie problems is plausible as well. Nontie problems have two different operands that activate two different interfering sets of answers while tie problems with only one true operand activate only one set of interfering answers. Less interference means faster response times.

Finally, another major strength of Campbell's model involves his extensive empirical demonstrations of priming effects. Campbell has shown that relatedness of primes can lead to enhancement or debilitation of performance. For short lags between trials, recently solved problems are inhibited while at longer intertrial intervals they are promoted as errors (Campbell, 1990, 1991). Ashcraft (1992)

has noted that the inter-trial error priming effect is analogous to similar effects in other domains of research such as the semantic memory literature.

Criticisms/Weaknesses of the Model

Criticisms of the model center around three main areas: the overdevelopment of the model, in some areas the underdevelopment of the model, and the lack of a procedural component in the model.

Ashcraft (1992) cites criticism that the different types of associations that are enumerated in the model are less parsimonious than is desirable. McCloskey et al. (1991) also argue that the network model is overdeveloped and at the same time argue that some aspects of the model are underdeveloped; I would have to agree. One example of underdevelopment of the model is the omission of the description of the mechanism that chooses among answer nodes with different levels of activation, a criticism that is justifiably leveled here as well as against Ashcraft's model. It is also not possible to ascertain the comparative importance of the different types of associative strengths at present; what is the relative strength of an operand-answer association compared to a problem-answer association, for instance? The overdevelopment of the model lies in the fact that there are so many different associations. Often times it is conceivable that a subset of these associations is adequate to explain a phenomenon and it is then unclear as to why all of the associations are necessary.

Lastly and maybe most importantly, very little attention if any is given to a procedural component, a criticism that was leveled against Ashcraft's model earlier. How is it that these associations change over the course of development? How is it that children know the most efficient strategies to solve problems? These questions cannot be answered within the network retrieval model as it currently exists.

Overall Conclusions About the Models

All three groups of researchers agree in general that the theories they have put forth are compatible. Campbell and Graham's model shares with Ashcraft's model a network framework and operand driven activation. The idea of interference is central to both the model of Campbell and Graham and Siegler's model. All three groups of researchers agree that strengths of answers are associated with problem presentation and practice.

There are also, however, significant differences among the models. While the network interference model and the distribution of associations model both incorporate interference effects, the distribution model claims that interference occurs within a specific problem, and that the distribution of associations for problems are completely independent of one another, something Campbell and Graham see as incorrect. And, of course, the two network models, in emphasizing the study of retrieval, largely ignore the fact that children use many different strategies, retrieval being only one of them. At present, although the models have

differences, there is no evidence that any one of the models is superior to the others. At this point it is beneficial to have these three different vantage points as they bring different hypotheses to light (e.g. the tie problem explanation that Graham and Campbell give compared to Siegler and Ashcraft's explanation). In addition, Ashcraft has made an attempt to assimilate some of the theoretical aspects of the other models into his model (Ashcraft, 1992), an indication that it may be possible to eventually weave the theories into a coherent whole.

My personal bias is to favor Siegler's distribution of associations model. The evidence mentioned in a previous section of this paper concerning the models external validity (Geary & Burlingham-Dubree, 1989) and the many studies that have been carried out that apply the model to specific subpopulations of children (e.g., Geary & Brown, 1991; Siegler, 1988a; Goldman, Pellegrino, & Mertz 1988; Geary, Fan, Bow-Thomas, Siegler, 1993) make starting with Siegler's model the most appealing option. I agree with McCloskey et al. (1991) that of the three models the distribution of association model is "more complete, more explicit, and more tightly constrained than other current models of arithmetic fact retrieval" (p. 389). The challenge for the model is to be able to make the necessary adjustments so that it can better explain certain phenomena (e.g., priming effect, operation confusion effect) without radically altering its composition so as to lose its present predictive power.

Now that the discussion of mental arithmetic is complete let me turn your attention to another important aspect of mathematical development--problem solving or mathematical reasoning. Soon after children are introduced to arithmetic problems, the problems are embedded in contextual situations. These new types of problems require children to form representations and reason as well as do computation. In the next section I will focus on a particular area of problem solving in mathematics, namely word problems.

Word Problem Solving

There are many different aspects of problem solving and it would take an enormous amount of time and space to cover all topics in this domain. I will therefore focus on one of the dominant areas of problem solving, word problem solving, which will be part of the focus of the study soon to be described. Two of the sub-areas of word problem solving that I will focus on are features of word problems, semantic features of word problems in particular, and on processes that children use to solve word problems. As the focus of Chapter 2, I will try to show that developing the arithmetic skills mentioned above, most importantly being able to rapidly and almost effortlessly retrieve arithmetic facts from memory, plays a profound role in the development of mathematics word problem solving skills and mathematics reasoning skills in general.

Semantic Structure of Basic Word Problems

Semantic structure, in its most general sense, refers to the meaning of sentences in word problems and how they are interrelated. Riley, Greeno, and Heller (1983) put together a comprehensive classification system based on the semantic structure of addition and subtraction word problems. Their work has also been extended to multiplication and division word problems by other researchers such as Lewis (1989) and Lewis and Mayer (1987), but since most of their research was conducted using addition and subtraction problems, that is what will be focused on here.

Riley, et al. (1983) identified four major semantic categories of word problems and labeled them change, combine, compare, and equalize problems. I will be discussing these types of problems extensively so to avoid confusion I have provided the reader with Table 5 (adapted from Riley, Greeno, & Heller, 1983) on page 73 that lists the four types of problems and gives examples of each; the letters in parentheses at the end of each problem refer to their relative difficulty--(E)asy, (I)ntermediate or (D)ifficult--judgments that have been determined by a number of studies using samples of children of many different ages (e.g., Carpenter, Hiebert, & Moser, 1981; Nesher, Greeno, & Riley, 1982; Riley, Greeno, & Heller, 1983).

The first type of problem that appears in Table 5, the change type of problem. implies that an action is taking place. In the first change problem, Jane has three toys to

begin with and Bill gives (the action or change part of the problem) her 4 more toys. This action leads to a change of inventory, so to speak, for both children. Bill has 4 fewer toys than he did before the action and Jane now has 4 more. This problem requires that a child be able to count on 4 units from Jane's original 3.

The first combine problem requires this same solution strategy of counting on 4 (or 3) to solve the problem, yet the conceptual nature of the problem is totally different. In this type of problem there is no action per se, neither Jane nor Bill is giving or receiving toys, but rather their inventory of objects are being combined. A child has to realize that neither of the two subsets are being changed but rather are being combined to form a new superordinate set that encompasses both of the subsets.

The next set of problems, compare problems, like the combine problems involve static relationships. Again Bill and Jane do not experience a change in their overall supply of toys but one unknown set of toys is being compared to another to establish the quantity in that unknown set of toys. Equalize problems are similar to compare problems in that an action is performed to change the quantities of one of the sets. The difference is that in the equalize problems there is a constraint on the change so that one set of quantities must be acted on until that constraint is reached; there is no such constraint with compare problems.

Research has been conducted on what types of counting strategies young children use to solve the four types of word problems described above. Results of research in this area indicate that the types of actions that are implied in a problem, which is determined by a problem's semantic structure, affects what type of arithmetic strategy elementary school children will select to try to solve a problem (Carpenter, et al., 1981; De Corte & Verschaffel, 1987). For example, De Corte and Verschaffel found using manipulatives that 75% of their elementary aged sample used a strategy called an adding method to solve the first change problem in Table 5. The adding method involves counting a set of objects that represents the augend of a problem (first number) and then counting out a set to represent the addend, and finally counting out all of the objects. Similarly, the researchers also found children used a manipulative strategy called the no move strategy 68% of the time when solving the first type of combine problem. Even more interestingly, De Corte and Verschaffel found that semantic features of problems also dictated verbal counting strategies used to solve problems; a COL strategy was used much more frequently for combine problems than change or compare problems. In summary, it seems that semantic features of word problems dictate the strategies children use to solve them. This, however, is only one of many processing steps a child must use to solve a problem as can be seen in the next section.

Problem Solving Processes

Mayer (1985) has argued that arithmetic word problem solving involves four separate steps: problem translation, problem integration, solution planning, and solution execution. The processes of translation and integration are used by a child to form a representation of the problem and solution planning involves transforming that representation into a strategy that can be used to solve the problem. Because in large part I have already gone over solution execution in going over the strategies used to solve simple arithmetic problems, the focus of this section will be on representation of problems and solution planning.

A child's ability to solve different problems depends on how well he or she is able to represent a problem which, in turn, largely depends on understanding the text that outlines the problem (e.g., Kintsch & Greeno, 1985). This comprehension process involves understanding the meanings and mathematical implication of certain words (such as more, increase, less, etc.) as well as the composition of the problem as a whole. The way a problem is presented has a strong influence on how difficult it is to form a representation. To illustrate this, let us examine two compare problems from Table 5, the 3rd and 4th problems. Problem #4 is more difficult largely due to the second sentence in the problem. Let's examine why.

One reason that problem #4 is more difficult has to do with the relational word "less" in the second sentence of the

problem. The mathematical implication of this word is that the operation of subtraction should be used to solve the problem. This is not the case. Several researchers have documented that even college students have trouble solving problems correctly that contain what is termed inconsistent relational statements such as the "less" statement in problem #4 (Bovenmyer-Lewis & Mayer, 1987; Bovenmyer-Lewis, 1989; Mayer, Lewis, & Hegarty, 1992).

Also important in the second sentence of #4 is how the problem is structured. When a child or adult reads the first sentence of a problem he or she will use what is called a schema, which is a "general format for extracting and representing or translating the basic meaning of a problem" (Geary, 1994, p. 105). The organizing feature of the first sentence of problems #3 and #4 are the same, quantity. This is further elaborated by the reference to the who, what, and how many stated in the sentence. The important features of the first sentence of problem #3 might be concretely represented by the diagram at the top of Figure 3 on page 59. Similarly the second sentence could also be concretely represented as shown by the second diagram, the only difference being the unknown quantity. As children solve more and more of these types of problems it is thought that they do not have to continually build these schemas but can simply call up the schema and fill in the bubbles for how many, who, and what (Stigler, Fuson, Ham, & Kim, 1986).

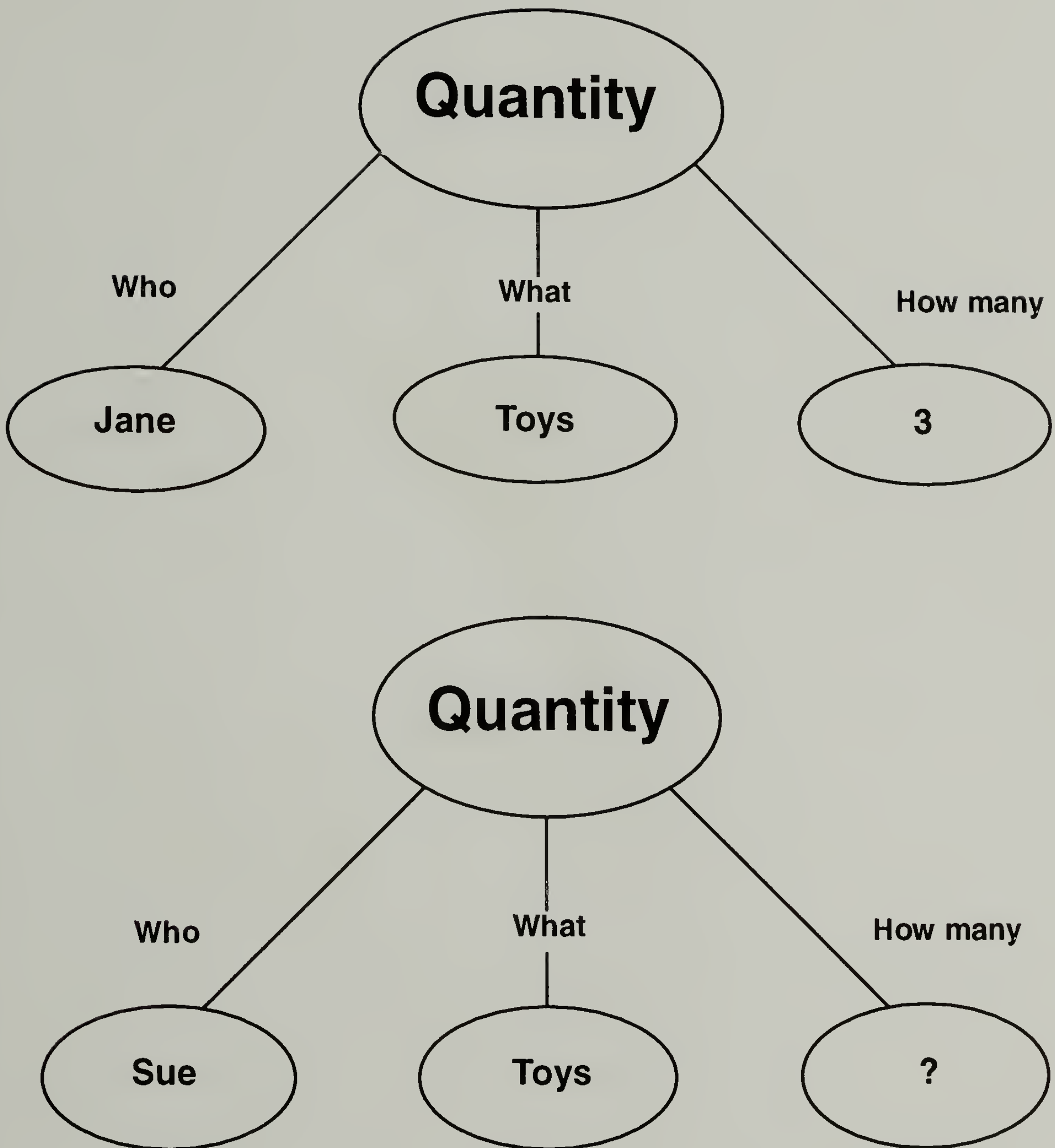

Figure 3. Possible schema formed to solve a compare word problem. Adapted from Geary (1994).

For compare problem #4 (refer to Figure 4 on page 61), children use the same schema to represent the first sentence; it is identical to that of problem #3. The difficulty comes in representing the meaning of the second sentence. The word "she" refers to Jane and if we try to represent this sentence with the schema used for the first sentence of the problem, we once again come up with Jane in the who bubble, toys in the what bubble, and 2 in the how many bubble. There is no way to represent Bill in this sentence using the above schema and an answer to the question, "How many does Bill have?" cannot be answered.

One way that children might be able to solve this problem is to set up a third representation that makes Bill the subject rather than the object of the second sentence (Mayer & Lewis, 1987). This would involve setting up a third representational schema, shown as the third relationship in Figure 4, that switches from having Bill as the object to the subject of the sentence as well as changing the relational statement from less to more. Because a child may have to set up a third representation for this problem it is no wonder children, and even adults, find this problem more difficult to solve than other compare problems. First, a child may not be able to set up and use this reversal schema and second, working memory demands are increased in performing this representational/reversal operation which would also increase the possibility of error.

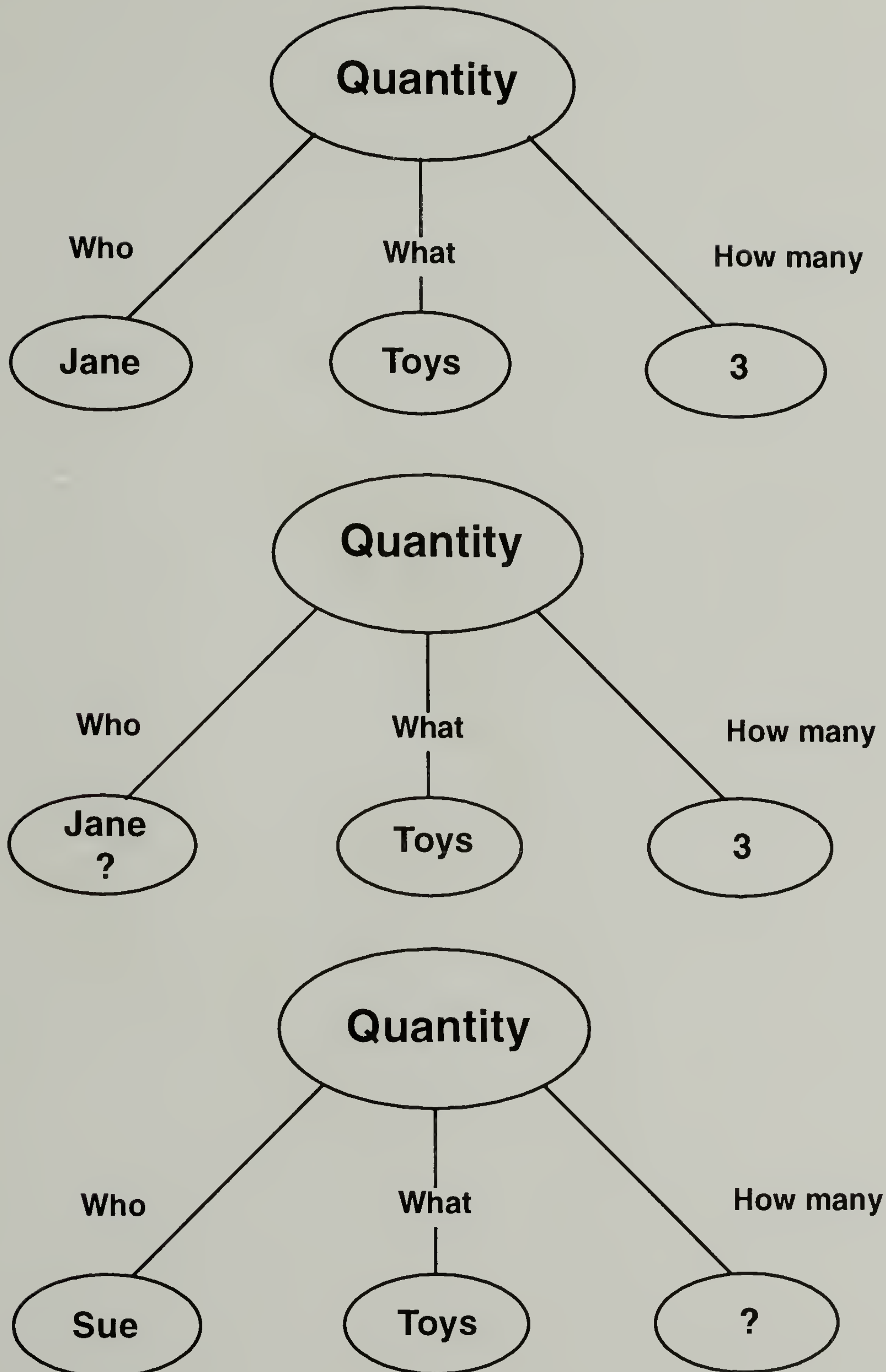

Figure 4. Possible reformulation of a schema to solve an inconsistent language compare word problem. Adapted from Geary (1994).

After a child has represented the basic meaning of a problem, a second type of schema needs to be composed that represents the relationship between important **features** of the problem instead of the meanings of each sentence. This process is part of the problem integration process. One way to understand relational schemas is shown in Figure 5 on page 63. The individual schemas are defined by the quantity dimension. This dimension depicts the most important feature of each individual sentence and the features are depicted as the horizontal line in the figure. For problem #3 Jane's position on the number line is the first quantity represented as the quantity associated with her is given in the first sentence of the problem. The relationship between how much candy Bill has in comparison to Jane is shown by his relative position to her on the number line. His name is placed to the left and this placement is guided by the relational word "more" in the problem.

In problem #4 the same use of this relational schema would result in a child solving the problem incorrectly. Again, Jane would be placed on the number line first as the first sentence gives us the quantity associated with her. Trying to represent the relationship between Jane's and Bill's number of toys is difficult because the relationship needs to be extracted from the statement "She has one candy less than Bill." The word "less" is very striking in this problem and often leads people to incorrectly locate Bill to the left of Jane on the number line and eventually make an

Diagram for Compare Problem #3

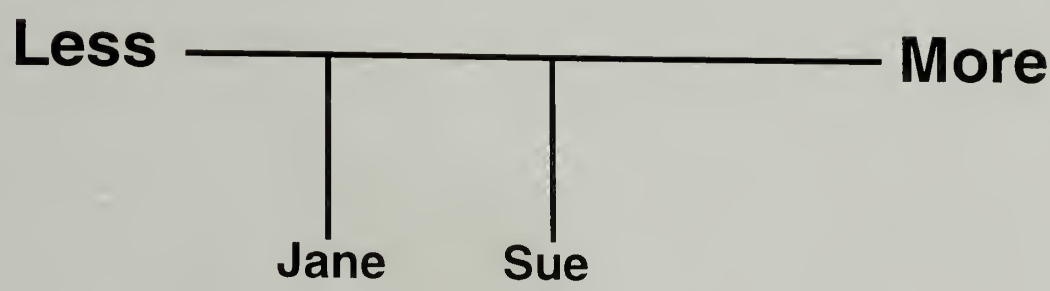

Diagram for Compare Problem #4

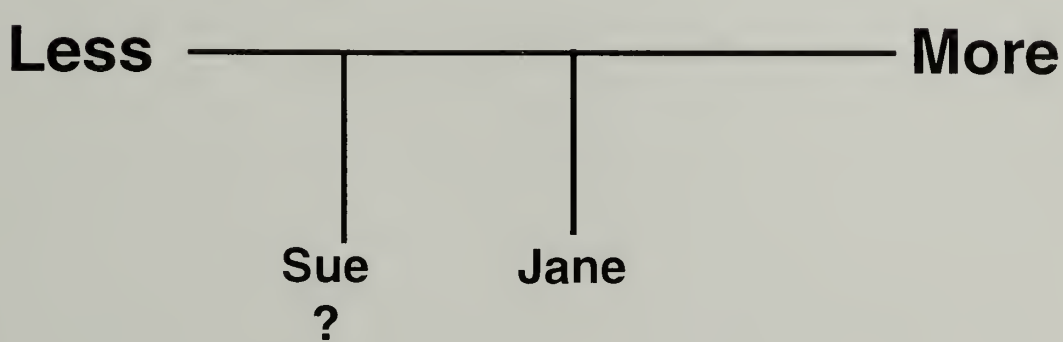

Figure 5. Possible relational schemas for solving two compare word problems. Adapted from Geary (1994).

error in trying to solve the problem. Lewis (1989) showed with a sample of college students that if trained in representing relationships on a graph in this way, students could significantly reduce the number of errors made on these types of problems.

The next step in the solution process is solution planning or the choosing of the best strategy to solve a problem. Riley et al. (1983) have argued that a third type of schema called an action schema bridges the gap between the relational schema and the actual selection of a strategy to solve the problem. Action schemas do this by providing implicit knowledge about the results that various arithmetic strategies produce, and the contexts in which they are most typically used. One example would be the adding method that was described earlier when reviewing the research of De Corte and Verschaffel (1987). The actual strategy used to solve a problem depends on "the best fit between the actions implied in the problem representation and the actions associated with the schemas that represent the outcome of each of the strategies available to the child" (Geary, 1994, p. 109). In more simple terminology, solution planning involves associating a child's representation of the important relationships of a problem with the child's best available (arithmetic) strategy for acting on that relationship.

Table 1. Strategies used to solve addition problems.

	Simple Mental Addition	
Strategy	**Description**	**Example**
Counting Manipulatives (CCA)	The problem's augend and addend are represented using objects. The objects are then counted starting from 1.	To solve 3 + 4, a child would count out 3 blocks, count out 4 more blocks, and then count out the entire set of 7 blocks.
Counting fingers	The problem's augend and addend are represented using fingers. At first, the fingers are counted from 1.	To solve 3 + 4, a child raises 3 fingers on one hand, then raises four fingers on the other, and finally moves each finger as he or she counts all 7 fingers starting from 1.
Verbal Counting		
Counting all starting with the first addend (CAF)	The problem's augend is verbally counted out and the addend is counted on top of it.	To Solve 3 + 4, a child counts "1, 2, 3, 4 is 1, 5 is 2, 6 is 3, and 7 is 4. The answer is 7."
Counting on starting with the first addend (COF)	The child counts on the second number from the cardinal value of the first.	To solve 3 + 4, a child counts, "3, 4 is 1, 5 is 2, 6 is 3, and 7 is 4. The answer is 7."
Counting all starting with the larger addend (CAL)	The larger of the augend and addend is counted out from one and the remaining number is counted on top of it (this strategy is quickly abandoned, if used at all).	To solve 3 + 4, a child counts, "1, 2, 3, 4, 5 is 1, 6 is 2, and 7 is 3. The answer is 7."
Counting on from larger (COL or min)	The cardinal value of the larger value in the problem is used as the starting point and the smaller value is counted on from there. This taxes working memory resources less because a child needs only to keep track of a minimum number of counts.	To solve 3 + 4, a child counts, "4, 5 is 1, 6 is 2, and 7 is 3. The answer is 7."

Continued, next page

Table 1 continued.

Strategy	Description	Example
Rule	An answer is based on a learned rule of addition that can be retrieved and applied to a problem such as n + 0 = n, or 0 + n = n.	To solve 5 + 0, a child simply states 5 on the basis of retrieving the rule that states any number plus 0 is itself.
Derived facts (decomposition)	A problem is broken down into two simpler problems, one (or both) of which may be retrieved from long term memory. Once the fact is retrieved, the value of the second problem is added on.	To solve 8 + 7, a child may break the problem into (8 + 2) + 5. The child may be able to easily retrieve the problem 8 + 2 = 10 and therefore may decompose 7 into 2 + 5. Once the child retrieves 8 + 2 = 10, the remaining 5 is most likely added on using one of the strategies above.
Fact Retrieval	The problem is solved by directly stating an answer retrieved from long term memory.	Retrieving 7 to solve 3 + 4.
	Complex Mental Addition	
Verbal counting on from larger	The same as COL strategy described above	To solve 21 + 5, a child counts, "21, 22 is 1, 23 is 2, 24 is 3, 25 is 4, and 26 is 5. The answer is 26."
Regrouping	The augend and addend are decomposed into their ten and unit values so that the tens units can be combined, the unit values can be combined, and these two sums combined to arrive at a solution.	To solve 43 + 29: Step 1: 43 = 40 + 3 Step 2: 29 = 20 + 9 Step 3: 40 + 20 = 60 Step 4: 9 + 3 = 12 Step 5: 60 + 12 = 72
Columnar addition using retrieval	The problem is solved by retrieving columnwise sums and trading or carrying when necessary.	To solve 43 + 29: Step 1: 3 + 9 = 12 Step 2: note the trade (carry) Step 3: 4 + 2 = 6 Step 4: 6 + 1 (trade) = 7 Step 5: combine 7 from tens column and 2 from ones column to state answer of 72

Table 2. Strategies used to solve subtraction problems.

	Simple Mental Subtraction	
Strategy	**Description**	**Example**
Manipulatives		
Separating from	The value of the minuend is counted out using objects. The subtrahend is subtracted by removing one object at a time and the answer is stated after the remaining objects have been counted from 1.	To solve 4 -2, a child counts out four blocks, then removes two from that set of four, and finally counts the blocks that are left, "1, 2. Two is the answer."
Adding on	The value of the subtrahend is counted out using objects. Objects are added one by one until the value of the minuend is reached. The child keeps track of the number of objects added on and states this number.	To solve 4 -2, a child counts out two blocks. The child then counts on, "3 is 1, 4 is 2. The answer is 2."
Matching	Two rows of objects are lined up in a one-to-one correspondence; one row represents the value of the minuend and one the value of the subtrahend. The answer is determined by counting the number of unmatched objects.	To solve 4 - 2, a child lines up a row of 4 blocks with a row of 2 blocks in one-to-one correspondence and counts the blocks that do not have a match, "1, 2. The answer is 2."
Counting fingers	The value of the minuend is represented by holding up the appropriate number of fingers. The subtrahend is represented by folding back down the appropriate number of fingers and the remaining fingers are counted and an answer is stated.	To solve 4 - 2, a child counts and holds up 4 fingers, "1, 2, 3, 4," and then counts and folds down 2 fingers, "1, 2," and counts the fingers that are still up, "1, 2; the answer is 2."

Continued, next page

Table 2 continued.

Strategy	Description	Example
Verbal counting		
Counting up	Involves counting up from the subtrahend until the minuend is reached. The child verbally keeps track of the number of counts completed.	To solve 4 - 2, a child counts, "2, 3 is 1, and 4 is 2. The answer is 2."
Counting down	Involves counting down from the minuend until the subtrahend is reached. Again, the child verbally keeps track of the number of counts completed.	To solve 4 - 2, a child counts, "4, 3 is 1, and 2 is 2. The answer is 2."
Addition reference	A problem is solved by retrieving a complementary addition fact.	To solve 9 - 3, a child may retrieve the answer to the complementary addition problem, 6 + 3 = 9, and use that information to state the answer, 6.
Rule	An answer is based on a learned rule of subtraction that can be retrieved and applied to a problem such as n - 0 = n.	To solve 5 - 0, a child simply states 5 on the basis of retrieving the rule that states that 0 subtracted from any number leaves the original number unchanged.
Fact retrieval	The problem is solved by directly stating an answer retrieved from long term memory.	Retrieving 3 to solve 7 - 4.
	Complex Mental Subtraction	
Verbal counting		
Counting down	Same as described above	To solve 17 - 4, a child counts, "17, 16 is 1, 15 is 2, 14 is 3, 13, is 4. The answer is 13."

Continued, next page

Table 2 continued.

Strategy	Description	Example
Decomposition		
Down over the ten	Involves first subtracting 10 from the minuend. This difference is subtracted from the subtrahend and this difference is then subtracted from 10.	To solve 16 - 8: Step 1: 16 -10 = 6 Step 2: 8 - 6 = 2 Step 3: 10 - 2 = 8
Take from the ten	Involves first subtracting the subtrahend from 10. Ten is then subtracted from the minuend. These two differences are then summed to arrive at an answer.	To solve 16 - 8: Step 1: 10 - 8 = 2 Step 2: 16 -10 = 6 Step 3: 6 + 2 = 8
Delete 10's rule	Involves increasing the value of the subtrahend to 10 and then subtracting 10 from the minuend. The difference between 10 and the subtrahend is then added to arrive at an answer	To solve 43 - 7: Step 1: 7 + 3 = 10 Step 2: 43 - 10 =33 Step 3: 33 + 3 = 36
Columnar retrieval	The problem is solved by retrieving columnwise differences and trading or borrowing when necessary.	To solve 32 - 9: Step 1: 30 -10 = 20 (trade) Step 2: represent 20 in working memory Step 3: 10 + 2 = 12 Step 4: 12 - 9 = 3 Step 5: 20 - 0 = 20 Step 6: 20 + 3 = 23

Table 3. Strategies used to solve multiplication problems.

	Simple Mental Multiplication	
Strategy	**Description**	**Example**
Repeated addition	The value of the multiplicand is added the number of times the multiplier dictates to get an answer.	To solve 6 x 4, a child represents it as 6 + 6 + 6 + 6 and adds, "6 is 1, 6 + 6 = 12 is 2, 12 + 6 = 18 is 3, 18 + 6 = 24 is 4. The answer is 24."
Counting by n	A problem is solved by counting up the number of multiples of the multiplicand that the multiplier specifies.	To solve 6 x 4, a child counts by multiples of 6, "6 is 1, 12 is 2, 18 is 3, 24 is 4. The answer is 24."
Rule	An answer is based on a learned rule of multiplication that can be retrieved and applied to a problem such as n x 0 = 0, n x 1 = n, n x 10 = n0, and n x 11 = nn.	To solve 5 x 1, a child simply states 5 on the basis of retrieving the rule that states any number times 1 is itself.
Derived facts (decomposition)	A problem is broken down into two simpler problems, one (or both) of which may be retrieved from long term memory. Once the fact is retrieved, the value of the second problem is added on.	To solve 6 x 7, a child may break the problem into (6 x 6) + (6 x 1). Both facts may be retrieved from long term memory as the tie problems (6 x 6) are quickly committed to memory and the rule for 6 x 1 is easily retrieved as well. These two simpler solutions are then added together.
Fact retrieval	A problem is solved by directly stating an answer retrieved from long term memory.	Retrieving 42 to solve 6 x 7.

Continued, next page

Table 3 continued.

	Complex Mental Multiplication	
Strategy	**Description**	**Example**
Decomposition	Involves solving a problem by breaking the multiplicand into its ten and unit values. The multiplier is then applied first to the tens value and then to the units value. These two products are then added .	To solve 43 x 6: Step 1: 43 = 40 + 3 Step 2: 40 x 6 = 240, hold in working memory Step 3: 6 x 3 = 18 Step 4: 240 + 18 = 258
Columnar retrieval	The problem is solved by retrieving basic facts from long term memory by column, holding in working memory, and carrying as needed.	To solve 43 x 6: Step 1: 3 x 6 = 18 Step 2: Note to trade value 10 to next column Step 3: 40 X 6 = 240 Step 4: 240 + 10 = 250 Step 5: 250 + 8 = 258

Table 4. Strategies used to solve division problems.

	Simple Mental Division	
Strategy	**Description**	**Example**
Repeated addition	The value of the divisor is added the number of times necessary to reach the value of the dividend.	To solve 24 ÷ 6, a child represents it as 6 + 6 + 6 + 6 and adds, "6 is 1, 6 + 6 = 12 is 2, 12 + 6 = 18 is 3, 18 + 6 = 24 is 4. The answer is 4."
Counting by n	A problem is solved by counting up the number of multiples of the divisor to reach the value of the dividend.	To solve 24÷ 6, a child counts by multiples of 6, "6 is 1, 12 is 2, 18 is 3, 24 is 4. The answer is 4."
Multiplication reference	A problem is solved by retrieving a complementary multiplication fact.	To solve 24 ÷ 6, a child may retrieve the answer to the complementary multiplication problem, 6 x 4 = 24, and use that to state the answer, 4.
Rule	An answer is based on a learned rule of division that can be retrieved and applied to a problem such as n ÷ n = 1.	To solve 7 ÷ 7, a child states 1 on the basis of retrieving the rule that states anything divided by itself is one.
Derived facts (decomposition)	A problem is broken down into two simpler problems, one (or both) of which may be retrieved from long term memory. Once the fact is retrieved, the value of the second problem is added on.	To solve 72 ÷ 8, a child may break the problem into (64 ÷ 8) + (8 ÷ 8). Both facts may be retrieved from long term memory as 64 ÷ 8 is the complement of a multiplication tie problem which are quickly committed to memory, and the rule for 8 ÷ 8 may be easily retrieved as well. The two simpler solutions are added together, "8 + 1 = 9. The answer is 9."
Fact retrieval	A problem is solved by directly stating an answer retrieved from long term memory.	Retrieving 9 to solve 72 ÷ 8.

Table 5. Semantic classification of word problems.

Change Problems

1. Jane had three toys. Bill gave her four more toys. How many toys does Jane have now? (E)

2. Jane had four toys. Then she gave three toys to Bill. How many toys does Jane have now? (E)

3. Jane had six toys. Bill gave her some more toys. Now Amy has nine toys. How many toys did Bill give her? (I)

4. Jane had some toys. Then she gave two toys to Bill. Now Jane has five toys. How many toys did Jane have in the beginning? (I)

Combine Problems

1. Jane has two toys. Bill has four toys. How many toys do they have altogether? (E)

2. Jane has five toys. Three are dolls and the rest are trucks. How many trucks does Jane have? (I)

Compare Problems

1. Bill has five toys. Jane has three toys. How many fewer toys does Jane have than Bill? (D)

2. Bill has seven toys. Jane has three toys. How many more toys does Bill have than Jane? (D)

3. Jane has three toys. Sue has one more toy than Jane. How many toys does Sue have? (D)

4. Jane has three toys. She has two less toys than Sue. How many toys does Sue have? (D)

Equalize

1. Bill has five toys. Jane has two toys. How many toys does Jane have to buy to have as many as Bill? (E)

2. Jane has three toys. If she buys three more toys, then she will have the same number of toys as Bill. How many toys does Bill have? (I)

3. Bill has six toys. If he loses three toys he will have as many as Jane. How many toys does Jane have? (I)

4. Jane has four toys. If she buys two more toys then she will have the same number of toys as Bill. How many toys does Bill have? (I)

5. Jane has three toys. If Jane loses one of her toys Bill will then have as many toys as Jane. How many toys does Bill have? (I)

CHAPTER 2

AUTOMATICITY AND WORKING MEMORY IN ARITHMETIC AND THEIR RELATION TO MATHEMATICAL PROBLEM SOLVING ABILITIES

Introduction

I would like to start off this section with a brief introduction to explain why I have included exposition of the topics automaticity, working memory, and complex mathematics in the same section. The availability of working memory resources is a very important factor in children's ability to solve more complex mathematical problems such as word problems. In order to study the effect that mental arithmetic abilities have on more complex abilities such as word problem solving, it is first necessary to determine if there is a savings of working memory resources when arithmetic problems are committed to memory. To ascertain whether working memory is saved, it is important to discuss studies involving automaticity and related constructs such as the autonomy and modularity of cognitive processes as they relate to mental arithmetic. In the following pages then, I will discuss mental arithmetic and automaticity followed by a related discussion of mental arithmetic and working memory resources. This introduction will conclude with an examination of issues surrounding word problem solving abilities in children, and how working memory/complex mathematical abilities might relate to each other. First, let me turn your attention to the definitions of autonomy, automaticity, and modularity.

Autonomy, Automaticity, and Modularity

The topic of automaticity as it relates to retrieval of basic arithmetic facts is a more difficult concept to define than it might seem at first. Use of the term automatic, or automaticity, is very loose in much of the mathematics literature as well as in the literature in other domains. Zbrodoff and Logan (1986) set out to test whether the processes that are used by adults to solve simple addition and multiplication problems were autonomous. In the introduction to their study they do an excellent job of defining the distinctions among automaticity, autonomy, and modularity. They define a process as being autonomous if it can be activated without intention and once it has been activated it runs to completion, unable to be arrested (Zbrodoff & Logan used the term "ballistic"). The concept of automaticity is closely linked to autonomy, it actually has the properties of autonomy within its definition. Recent theorist have defined processes that are automatized as being fast, effortless, unconscious, and autonomous (e.g., Laberge & Samuels, 1974; Logan, 1980; Posner & Snyder, 1975) and seem to believe that automaticity is a unitary phenomenon. According to Zbrodoff and Logan, this is where theorists start to run into trouble. Zbrodoff and Logan believe that there was no reason that all of the above mentioned properties had to co-occur and other researchers (Regan, 1981; Paap & Ogden, 1981) have backed up this claim. Work such as this has led to belief in the dissociation of

automaticity components; even though the components **may** be highly intercorrelated that cannot be assumed and different methods of measuring speed of processing, obligatory activation, capacity usage, and intentionality must be used (Stanovich, 1990) for each of these components.

Modularity of processing and autonomy have a relationship that is analogous to the relation between automaticity and autonomy. Modular processes are autonomous by definition, they do, however, have many properties in addition to autonomy (functional autonomy, encapsulation, or cognitive impenetrability are all terms used to describe autonomy as it relates to Modularity Theory [Stanovich, 1990]). According to Fodor, (1983) there are nine properties other than autonomy that compose modular processes. Zbrodoff and Logan point out that determining whether a process is autonomous or not can (and should) be pursued separately from both determining whether processes are automatic and/or modular.

Arithmetic, Autonomy, and Automaticity

What **did** Zbrodoff and Logan (1986) initially conclude regarding the autonomy of the processes of simple addition and multiplication? From the outset, based on previous research on cross operation confusion effects (e.g., Miller et al., 1984), they determined that addition and multiplication processes could be started unintentionally and that these two arithmetic operations must be either partly or fully autonomous. In order to assess whether they were fully

autonomous, they conducted a number of experiments, the first using a verification task. Subjects in one condition were given a pure block of addition problems with some associative lures in them such as 3 + 4 = 12 (if you replace the plus sign with a multiplication sign the statement is true). Subjects in a second condition were given a block of mixed problems that from trial to trial could be either multiplication problems or addition problems, that had associative lures mixed in as well. At the outset, both groups of subjects were told what types of problems they were about to receive; the pure block group was told they were only going to receive addition problems and the mixed group was told they would receive both types of problems. The addition only group therefore should never have intended to do multiplication. The researchers reasoned that if the same amount of interference was evidenced in both conditions, then the process of addition was not affected by intention. If the interference (RT's) was greater in the second condition, it would provide evidence that in the second condition the subjects were intending to perform one operation as often as the other and associative lures would have more effect than when subjects were only concerned with one operation. The experiment was also repeated using multiplication blocks as the pure condition.

The results showed that interference was not equal in both the mixed and pure blocks, although the difference was smaller in the experiment that used pure blocks of

multiplication. Because of these results, Zbrodoff and Logan concluded that answering the question of whether arithmetic processes are autonomous or not is not best served by a dichotomous response; perhaps a better answer would be to place the process on an autonomy continuum. They concluded that arithmetic was at least partially autonomous.

Zbrodoff and Logan then carried out an experiment to test the second part of the autonomous process definition, whether the process of mental arithmetic would continue in a ballistic fashion once it was started. The paradigm used was a stop sign condition where subjects were shown arithmetic problems on a computer screen and were interrupted during some trials when a tone was sounded 100, 300, 500, or 700 ms after an arithmetic problem disappeared. The stop sign signaled that the subject should try to inhibit their processing of that problem. Subjects later completed a memory test to determine which problems subjects recognized. Two experiments were conducted, one in which subjects had to verify answers to arithmetic problems and one in which subjects had to produce answers to problems. The hypothesis in each experiment was that subjects would remember significantly fewer problems when they tried to stop their processing. The results of the memory test supported the hypothesis that subjects could stop their processing and that these simple arithmetic procedures were not ballistic. At least one question lingers, however; is it possible that the

response to the stop signal disrupted memory whether the arithmetic process was inhibited or not?

More recently, Klapp, Boches, Trabert, and Logan (1991) have redefined automaticity. These researchers criticize "property-list" approaches to automaticity such as Laberge and Samuels (1974) who define automaticity using the terms fast, effortless, and autonomous, because these approaches are merely descriptive and do not supply a mechanism that would allow for predictions about what needs to be done to attain automaticity. Klapp et al. also criticize resource theory definitions of automaticity because even though they allow the properties of automaticity to be deduced they, like the property-list approaches do not specify a learning mechanism.

The definition of automaticity that Klapp et al. embrace is a memory retrieval definition. In this type of theory (e.g., Logan, 1988) automaticity is achieved when performance is a result of single-step retrieval of solutions from memory. According to this theory, memory retrieval is always in competition with algorithms. Algorithmic procedures win the race when retrieval of answers is slow and inefficient, as for example, when a child is in the beginning stages of learning multiplication; memory retrieval will be attempted but will usually be beaten out by algorithmic procedures. Automatic retrieval is therefore fast because memory retrieval is fast, is obligatory because memory retrieval is obligatory, and is effortless compared to algorithms because

memory retrieval only wins the race when associative strength is high and conditions ripe for retrieval are present (Klapp, et al., 1991). Let us now move to a discussion of obligatory activation of processes.

Arithmetic and Obligatory Activation

Originally tasks called Stroop (1935) tasks were used in domains other than mathematics, such as word identification, as evidence for the existence of automaticity. As mentioned above, the idea that obligatory activation of processes always co-occurs with resource free processing is not necessarily true. In the case of word recognition, some children as early as first grade were reaching asymptote on Stroop measures of automaticity which was in direct conflict with the general belief that the development of prelexical automaticity was responsible for comprehension development over a long period of time (Stanovich, 1990). We now know that the Stroop task is a good indicator of intentionless activation but that it does not necessarily directly indicate capacity use.

Some studies using Stroop and Stroop-like tasks have been conducted recently in the domain of arithmetic to assess obligatory activation. One such study was conducted by Lefevre et al. (1988). The Stroop tasks for their subjects involved a verification task. Subjects were shown an addition problem and immediately following the addition problem a probe digit was presented. It was the subject's responsibility to press a yes button if the probe matched one

of the addends in the addition problem and to press a no button if the probe matched neither of the addends in the addition problem. Time between the presentation of addition problem and probe (termed stimulus onset asynchrony, or SOA) was varied. The crucial trials in this experiment were the "no" trials that involved a probe number that was actually the correct sum to the presented addition problem (e.g. 5 + 2; probe is 7). If addition is an obligatory process, one would expect reaction times for correct sum probes to be greater than non correct sum probes for the no trials. That is exactly what the researchers found for both the 60 and 120 millisecond SOA conditions. They also found the effect for word presentations as opposed to digit presentations (five + two; probe is seven) and for trials where the plus sign was removed (5 2; probe is 7). Later studies revealed that (college) subjects that were skilled at multi-digit arithmetic showed more obligatory activation for the probe type problems mentioned above than subjects who were less skilled at multi-digit arithmetic (LeFevre & Kulak, 1994), and similar effects have been demonstrated for elementary school children (Lemaire, et al. 1994), as well as elderly subjects (Rogers & Fisk, 1991).

Koshmider and Ashcraft (1991) wanted to document the change in this obligatory process over development and used a priming paradigm and students at four different grade levels to investigate this. Subjects were given true-false multiplication verification problems that were each preceded

by a number prime. The number prime was either a correct answer to the problem (relevant prime), an incorrect answer to the problem but a correct answer to another multiplication problem (irrelevant prime), or a line of two dashes (neutral prime). One would expect that relevant primes would reduce response times, irrelevant primes would interfere with verification judgments and raise RT's, and the neutral primes would have no effect. Another prediction would be that the effects would become greater over time as one becomes more proficient at arithmetic. The researchers conducted statistical analyses on the benefits (relevant prime condition RT's compared to neutral prime condition RT's) and costs (irrelevant prime condition RT's compared to neutral condition RT's) of priming. The results were somewhat mixed. For easy multiplication problems the benefits across the 225, 450, and 1400 millisecond SOA's were significant for (all) grades 3, 5, 7, and college students. Hard problems showed significant benefits for the three older age groups at the two longer SOA's. Koshmider and Ashcraft noted that these findings contradicted the findings of Campbell (1987b) using an SOA of 300 ms, but that this contradiction was probably due to processing differences between the verification task used in this experiment and the production task used in Campbell's study.

Analyzing costs, Koshmider and Ashcraft found that easy problems at all SOA's were significant across all grade levels. The medium and difficult problems showed a trend

toward non-significance as SOA and grade level increased, a finding that is in agreement with earlier presented findings of LeFevre et al. (1988) and Lemaire et al. (1994). Koshmider and Ashcraft drew two main conclusions: 1.) the experiment showed that interference is not the only priming effect, facilitation also occurs, a finding consistent with Campbell (1987b), and 2.) that facilitation is larger with increased SOA indicating that conscious processes might be aiding in the activation in the network. In terms of development, benefit increases for hard problems at longer SOA's and older ages while costs are present at all ages for easy problems, but diminish as age and difficulty of problem increases.

In summary, there is much evidence to show that arithmetic procedures are at least partially autonomous. As the work of Koshmider and Ashcraft (1991) and Zbrodoff and Logan (1986) point out, however, in certain contexts these autonomous effects can be altered.

Working Memory and Arithmetic

Working memory is a construct that is related to automaticity; it is the amount of resources used during processing. Working memory in the context of this paper can basically be described as the ability to keep information in memory while mentally acting on that information or other information related to it. The trading or carrying of numbers in complex mental addition reflects such a process; in mentally computing 24 + 28 one might decompose the problem

into two simpler problems such as (20 + 20) + (8 + 4). In order to complete this problem one must compute 8 + 4 = 12, hold that answer in working memory, use working memory to solve 20 + 20 = 40, and then combine the two answers. Available working memory resources are an important factor in children being able to successfully solve more complex types of mathematical problems. I would like to discuss several studies that try to assess the working memory resources that are needed to perform mental arithmetic. Ashcraft, Donley, Halas, and Vakali (1992), Kaye et al. (1989), Logie et al. (1994), and Klapp, et al. (1991) have all studied this, although in slightly different ways.

Ashcraft et al. (1992) examined how much working memory resources were used during single and double digit addition. It has already been established that double digit addition taxes working memory resources (Hitch, 1978) so that aspect of the study about to be described is really a replication of previous work. The arithmetic task used in this study, as in most of Ashcraft's studies, was a verification task. Addition problems were presented with either incorrect or correct answers and subjects were instructed to press one button if the sum was correct and another if the answer was incorrect. In order to assess how much working memory was required to complete this task, Ashcraft et al. had subjects engage in concurrent tasks (a dual task paradigm), specifically letter and word tasks that consumed working memory resources but that did not interfere with the

arithmetic task. One of the letter tasks, the control task, was simple letter repetition where subjects would see four of the same letter appear on a computer screen and simply had to repeat the letter until the arithmetic problem came on the screen and the person solved it. The tasks that were used to manipulate working memory load were a word generation task and an alphabetization task. In the word generation task subjects would see four letters appear on a computer screen and were required to rapidly generate words that began with those letters until the arithmetic problem appeared and was solved. In the letter alphabetization condition four letters again appeared on the screen and subjects were required to recite them as they appeared on the screen and then, as rapidly as possible, recite them in their alphabetic order.

The conclusions that Ashcraft et al. (1992) were able to draw from the statistical analysis was mixed (as an aside, only correctly answered arithmetic problems were used in the analyses). There was no task by problem difficulty interaction found. One interpretation of this was that having to perform a task concurrently with a mental addition verification task did not affect arithmetic performance (speed). Problems did arise in this interpretation when the details of the reaction times for each concurrent task were analyzed and when subject behavior was taken into account. The word generation and alphabetization task increased RT about 400 and 600 ms respectively while the letter repetition control task only increased RT about 250 ms. It was also

noted during the experiment that subjects often slowed their verbal response rate in the word generation and alphabetization conditions (even after reminders that they should do this task as rapidly as possible). Hence the 400 ms and 600 ms increases may have been larger in the two conditions if the subjects had kept their vocalization speeds to a maximum. Ashcraft et al., because the dual task load may actually have been greater in the above experiment if subjects had not slowed their vocalizations during the alphabet and word task, left open the possibility that some working memory resources are needed for people to solve simple addition problems.

Kaye et al. (1989) carried out research on working memory and arithmetic as well, although it differed in many respects from the design used in Ashcraft et al. (1992) above. In order to assess the amount of working memory resources that were required to solve simple addition problems, Kaye et al. used an auditory probe detection task that included a unique procedure for assessing processing loads at different stages of the solution process. The addends of an addition problem were presented on a computer screen 750 ms after the beginning of a trial, and the answer to be verified was presented 1750 ms after the beginning of a trial. The primary task was for subjects to verify whether the addition statements were true or false. On half of the trials, an auditory probe was presented. The subjects' secondary task was to respond to this auditory probe by

pressing a button. The onset of the auditory probe was varied (it occurred 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, or 2500 ms after the addends were presented) allowing the researchers to assess the demands on working memory at different times during the solution process. For example, if the auditory probe was presented when the addends were presented, the elevation in response time to the probe as compared to probe response time with no processing demands was a measure of addend encoding demands. Another main objective of the experiments was to see how working memory demands changed over development. In order to evaluate the developmental change, subjects from grades 2, 4, and 6, and college students were included in the study.

In their analyses, Kaye et al. used combined RT's for the arithmetic verification task and the probe detection task because the probe task affected performance on the primary task, and therefore dual task reaction time was a more accurate measure of working memory load. A very definite developmental pattern was found. Second graders showed the most inefficient processing as evidenced by large increases in RT as a function of onset of the probe which would be expected, as numerous studies have shown that at this grade few students have begun to retrieve number facts directly, and counting strategies are very prevalent (e.g. Kaye, Post, Hall, & Dineen, 1986). Fourth graders showed that early processing demands of the task were greater than later demands. Sixth graders had less of a processing demand than

the fourth graders, and demands were equal during both the encoding and computational stages of processing. College subjects showed very efficient performance (mean RT = 649 ms), but there were still effects of probe interval on combined response time. Overall the results seem to indicate that even for college students that presumably are most skilled at arithmetic fact retrieval, some working memory resources are needed to verify the truth of addition facts.

A third study conducted by Logie, Gilhooly, and Wynn (1994) using adults and a double digit mental addition task yielded similar results. The researchers found that certain secondary tasks lead to disruption of double digit mental addition and that this disruption occurred whether subjects were administered mental addition problems aurally or visually.

The last dual task study I would like to mention is Klapp et al. (1991). The arithmetic task used in their study was quite unique. The general hypothesis they wished to test was whether overtraining subjects on an addition task would lead to reduced interference when overtrained subjects were then asked to verify addition problems while performing a concurrent task; in other words does overtraining lead to the saving of working memory resources when trying to solve addition problems. Klapp and his colleagues wanted to have three groups of subjects to compare: novices, subjects that had automatized addition problems, and subjects that had practice beyond automatization. One way to examine the

acquisition of automaticity in addition is to select and compare children of different ages. There is a problem, however, in that general processing speed varies greatly from grade to grade making it difficult to determine when the addition process becomes automatic. Another problem is that the amount of practice may vary widely from child to child within grades with no way to control it.

To sidestep these problems a task called alphabet arithmetic was designed. In alphabet arithmetic (AA) one addend of a problem is a number and the other is a letter. The answer to a problem is found by starting at the letter addend and moving forward through the alphabet the number of letters specified by the number addend in the problem. So if the problem is D + 4 = ?, one must count from D four letters forward: E is one, F is two, G is three, H is four, H is the answer. The beauty of this task is that the researchers can be assured that the amount of practice a subject receives is tightly controlled and it bears a close resemblance to normal addition; people have to use counting algorithms to initially solve problems but eventually progress to where answers can be retrieved. It also is beneficial to use this task because adults who have processing speeds that are generally quick and relatively stable can be used as subjects.

The basic nature of the experiments was as follows. Klapp et al. trained groups of subjects to three different levels of skill. One group was a novice at AA, one group was trained to automaticity, and one group was trained beyond

automaticity. Automaticity was considered to be present when the memorability of a problem no longer depends on the size of the addends in the problem. In other words, if response time and digit magnitude are plotted linearly the slope of that line should approach zero; if someone were using a counting strategy to solve AA problems the slope of this line would be approximately 400-500 ms (Klapp, et al., 1991). Each group was required to perform two different types of secondary tasks while performing the primary task of verifying the correctness of AA problems. One concurrent task was repetitive speech (repeating January over and over) and in another experiment nonrepetitive speech (reciting the months in order) was used. The results of the experiments were in line with predictions. Novice subjects experienced interference when doing either concurrent task (increased RT and errors compared to a control condition), the automatized group experienced interference with the AA task only during the sequential month saying task, and the overlearned group did not experience interference while performing either concurrent task.

In addition to the working memory studies outlined above, other related research has been conducted that shows there is a relationship between another working memory measure, digit span, and mental arithmetic strategy choice and abilities (Geary, et al., 1993; Ellis and Hennelly, 1980; Ellis, 1992). These studies involved children of different nations and showed that digit span length is related to the

pronunciation rate of number words in different languages. Children that have longer digit memory spans because of shorter number words may progress to using more advanced counting strategies (including retrieval) sooner than other children.

What is the significance of these studies? They have provided us with evidence that while some working memory resources <u>are</u> used during basic computation, mental arithmetic processes are at least partially automatic and working memory demands can be reduced by overlearning basic arithmetic facts.

In summary I would like to echo one of the concluding statements in Kaye et al. (1989) that "The next steps to be taken in this type of chronometric research on the development of mathematical ability would involve direct measurement of the processing savings accrued while subjects are performing more complex mathematical tasks that require efficient arithmetic computations as part of their solution" (p. 479). A more general statement would be that we need to determine the link between degree of automatization and ability to solve more complex mathematical problems.

<u>Basic Number Facts and Mathematics Achievement</u>

In a number of articles, researchers have correlated measures of basic number fact mastery with scores on different achievement tests. Does automaticity of basic math facts correlate with performance on achievement test sections that deal with more complex problems? A number of studies

suggest that knowledge about basic arithmetic facts is related to computational scores and mathematics concept scores on achievement tests. In Siegler's (1988a) work, with good students, not-so-good students, and perfectionists mentioned earlier there were significant differences between the good and perfectionist groups compared to the not-so-good students on the math computation and math problem solving subtests of the Metropolitan Achievement Test. Students that were either quickly able to retrieve arithmetic problems from memory or that were able to quickly retrieve arithmetic problems on some problems and quickly count using a backup strategy for others, performed better on arithmetic computational and problem solving achievement measures.

Another study, reported by Resnick and Ford (1981) gave an example of how computer aided instruction in math computation for students in grades 1-6 helped to significantly improve Stanford Achievement Test scores from one year to the next. Compared to a control group, students improved not only on sections testing their computational ability, but also on sections that tested their knowledge of mathematics concepts and how to apply them.

Still other studies have examined the relationship of both basic math fact abilities and reading abilities with word problem solving abilities. For instance, Muth (1984) studied the relationship of reading and computational skills to the ability to solve arithmetic word problems. She conducted regression analyses using 6th graders' scores on a

15 item word problem test as the criterion variable and using scores on the reading comprehension and arithmetic computation subtests of the Comprehensive Test of Basic Skills as predictor variables. A significant amount of variance (54%) was accounted for by the reading ability and computational skills together with 8% of that being unique to computation. Balow (1964) conducted a similar analysis using subtests of the Stanford Achievement Test as measures of reading, computational, and mathematical reasoning abilities. After controlling for IQ, Balow found, like Muth, that reading ability and computational abilities were significant predictors of mathematical reasoning abilities.

Research on the relationship between mental arithmetic abilities and word problem solving abilities extends into research with subpopulations as well. Zentall (1990) conducted a study with normal, learning disabled, and attention deficit disorder students in the seventh and eighth grades. Zentall was attempting to determine the relationship between reading comprehension, cognitive skills, behavioral scores (measures such as bottom/torso movements that were operational definitions of attention), and math-fact retrieval time with number of problems correctly solved on a math word problem test. Only behavioral scores and math-fact retrieval time correlated with percentage of word problems correctly solve and only math-fact retrieval time correlated significantly with absolute number of word problems correctly solved.

How might arithmetic skill and its development in turn affect the development of problem solving abilities? At a somewhat obvious and mundane level, accuracy and speed of arithmetic problem solving aids in accuracy and speed of more complex problem solving--the more accurately and more quickly a person can carry the solution process of a word problem the more problems that person will get correct on a test of mathematic abilities--especially if it is timed. There is evidence, however, that accuracy probably does not account for much of the variability in word problem solving past grade 5. Previous research has shown that by grade 6 students have achieved about 95% of adult accuracy on simple arithmetic problems but only about 65% of adult speed (Mercer, 1979). It is safe to assume that incorrect problem solution will infrequently occur due to failure to do computations correctly. Another piece of research by Morales, Shute, and Pellegrino (1985) supports this claim as well. These researchers found that for their fifth/sixth grade sample of students, somewhere between 80-90% of incorrect solutions of word problems were due to conceptual not computational errors.

Another more theoretically interesting hypothesis does exist, however. Hiebert (1990) offered the suggestion that making the basic math facts automatic frees more space in working memory to think about how to **apply** facts in a problem. Several other researchers have echoed this thought (e.g. Geary, 1994; Geary & Widaman, 1992; Silver, 1987; for

an opposing viewpoint see Rabinowitz & Woolley, 1995). Resnick and Ford (1981) have also offered a similar answer, stating that "number facts...need to be developed to the point of automaticity so they can avoid competing with higher-level problem-solving processes for limited space in working memory" (pp. 32-33) and drew an analogy to research in reading where automaticity of word recognition is associated with higher levels of reading comprehension (LaBerge & Samuels, 1974; Perfetti & Hogaboam, 1975).

Geary (1994) in his review of the cognitive literature involving basic mathematical abilities and mathematical reasoning abilities came to a similar conclusion, that people that have well developed mathematical reasoning abilities are able to do three things: 1.) they are able to quickly and automatically solve basic arithmetic problems, 2.) they have developed schemas to help their problem solving, and 3.) they are able to hold things in working memory while carrying out other procedures.

Where might these working memory savings help in solving a problem? What I propose is that the working memory load reduction allows a child to keep representational aspects of a problem in mind longer so that those representational aspects may more likely be linked to the specific strategy or procedure that is used to eventually solve the problem. It also should help, especially in multi-step problems, when children have to keep the representational aspects of a problem in mind, translate that information into a solution

procedure, and finally carry out that procedure. If a person is slow in retrieving arithmetic facts to carry out the solution procedure, he or she will have a hard time keeping the steps of the procedure in mind and will be more likely to make an error. So I propose that the working memory savings helps within a problem by allowing a person to keep a solution procedure in mind long enough to solve a problem and helps between problems in that a person is able to concentrate on linking specific types of problems to specific types of solution procedures and will more readily solve analogous problems in the future.

CHAPTER 3

PURPOSE AND METHODOLOGY OF THE CURRENT STUDY

Purpose

The research conducted had two main purposes: to study the relationship of basic addition, subtraction, and multiplication abilities (speed and accuracy) with ability to solve more complex computational problems and, more importantly, to solve single and multi-step word problems. In addition, it was important to examine the relationship these simple arithmetic abilities had to standardized mathematics computation and problem solving/concept tests. In order to accomplish these objectives arithmetic accuracy and response time measures, word problem and complex computational measures, standardized math test scores, and reading ability measures were collected.

Three main aspects of the current study distinguish it from other studies in this domain. The first novel aspect of this study was that simple mental arithmetic measures were recorded using a computer. Other studies have looked at the relationship of more complex computational abilities (e.g., the ability to solve problems such as 212 x 37 = ?, or 3/7 + 5/8 = ?) with complex mathematical problem solving abilities (e.g., Balow, 1964; Muth, 1984) and/or have used imprecise paper and pencil measures to assess computational ability (e.g., Balow, 1964; Muth, 1984; Zentall, 1990). By recording simple mental arithmetic response times via a computer it was possible to examine if a more basic level relationship

exists, the simple mental arithmetic/problem solving relationship, and to do so with a very precise measurement tool, the computer.

Secondly, in this type of research the dilemma about what to do with subjects' error trials persists. In the past, researchers have either deleted children's error trials from analyses (e.g., Geary, Brown, & Samaranayake, 1991; Koshmider & Ashcraft, 1991; Geary & Brown, 1991), assumed the speed accuracy trade-off is insignificant and have used all trials in analyses (e.g. Ashcraft & Fierman, 1982), or deleted error prone subjects and assumed the speed accuracy trade-off is insignificant (e.g., Ashcraft, et al., 1992; Graham & Campbell, 1992). A method of combining accuracy and response time data has been developed recently in the LATAS lab at the University of Massachusetts. If this combined accuracy/response time measure proves to be more predictive of complex mathematical problem solving abilities in this study, this new variable may be used in subsequent related studies and could help to solve the error trial dilemma.

The third and final distinguishing feature of this study was the addition of a developmental component. In previous work it has often been the case that only one grade is sampled for study (e.g., Muth 1984) and no conclusions can be drawn about how the arithmetic/complex mathematical problem solving relationship may change from grade to grade. In this study four grade levels were sampled and it will be possible to look at how that relationship changes over development.

Based on the previous research conducted on mental arithmetic, math problem solving, and working memory reviewed in the introduction, several interesting predictions can be made about the results of this study.

1.) Arithmetic response times should change over grades, starting out fairly high at the 5th grade and gradually approaching a low asymptote with development (grade).

2.) Arithmetic accuracy should change somewhat, starting out lowest at the 5th grade and gradually increasing with each grade.

3.) Arithmetic measures (collected using the Computer-based Academic Assessment System, or CAAS) and grade will be significant predictors of complex computational ability and the CAAS arithmetic measures will be better predictors of the criterion at the early and/or middle grades and poorer predictors at the later grades.

4.) CAAS arithmetic measures, reading comprehension, and grade will be predictive of complex math problem solving abilities and the CAAS arithmetic measures again will be better predictors of the criterion at the early grades and/or middle grades and poorer predictors at the later grades.

Methodology

Subjects

Twenty seven grade 5, 28 grade 6, 23 grade 7, and 22 grade 8 students were selected from classrooms from a local middle school. The classrooms in the school were not grouped

by ability level (tracked) so it was assumed that students from different classrooms in the same grade did not differ, on average, in mathematic ability or intelligence.

Materials

Pencil and Paper Test

A mathematical skills test was administered to all of the students. This test was constructed in conjunction with the guidelines set forth by the principal and math teachers in the participating school (see Appendix, pages 152-155). Fifty questions were constructed, 22 of which were complex computational problems and 28 of which were word problems. The word problems were patterned after problems from previous word problem research (Compare, Change, and Equalize problems from Greeno, Riley, & Heller, 1983; relational problems from Lewis & Mayer, 1987, and Lewis, 1989), problems that appear in the 6th grade Iowa Test of Basic Skills (ITBS) booklet, Form K Level 12, as well as problems that were suggested by the Belchertown math teachers. The computational problems involved the addition, subtraction, multiplication, and division of single and multi-digit whole numbers, decimals, and fractions and again were patterned after Iowa Test problems and problems suggested by the math teachers. An instruction sheet and answer sheets with ample space to work out problems were also provided to students.

Reading Measure

Because previous research (e.g., Balow, 1964; Geary, 1994; Muth, 1984; Zentall, 1990) has noted the importance of

reading ability in problem solving, scores from a reading test that were administered earlier in the same school year were collected as well. The reading tests were 96 item, grade appropriate Sentence Verification Technique (SVT) reading tests developed by James M. Royer at the University of Massachusetts. Several articles exist that detail the construction and psychometric properties of the test and give examples of passages and questions included in the test (e.g., Royer, 1990; Royer, Carlo, & Cisero, 1992; Royer & Sinatra, 1994)

In general, SVT tests are usually composed of 6 story passages that range from a half to a full page in length. Each passage has 16 test sentences associated with it and students are instructed to mark "yes" on their answer sheet if the test sentence they read means the same thing as a sentence read in the story and to mark "no" if it does not. Composing the 16 sentences are 4 sentences selected from 4 different categories of items. Two categories of sentences preserve meaning while the other two involve meaning changes. Sentences that preserve meaning are either exact duplicates (original) of a sentence from the story or are sentences from the story that have had some or all of their words changed without affecting the original meaning of the sentences (paraphrase). The meaning change sentences are either sentences from the story that have had a word or words altered to change their meaning (but still would fit the overall topic of the passage) or are sentences that have no

relation to the passage or any of the sentences in it (distracter).

Computer Testing

Each student was also administered a mental arithmetic test using the basic skills component of the Computer-based Academic Assessment System (CAAS). A Gateway 2000 Colorbook laptop computer with a 5.5" x 7.5" monitor was used to administer the CAAS mental arithmetic battery. A total of 4 different computational content areas were sampled: addition, subtraction, and multiplication (two tests, "easy" and "hard multiplication") of whole numbers.

Arithmetic test stimuli appeared in the middle of the computer screen in black against a white background. The appearance of each stimulus immediately triggered a timing mechanism in the computer and each stimulus remained on the screen until the student voiced an answer to the problem into a microphone interfaced with the computer. The voicing of an answer stopped the timing mechanism thereby recording the response time for each trial, accurate to +/- 2 milliseconds.

A scoring box was also interfaced with the computer and was used by the researcher to record whether the student's response for a trial was correct or incorrect. The researcher pressed the left button on the box to indicate a correct response and the student subsequently heard a bell and saw the word "correct" appear in the upper right hand corner of the computer screen. The researcher pressed the right hand button to indicate an incorrect response and the

student heard a buzz and saw the word "wrong" appear in the upper right hand corner of the screen. If the microphone picked up a noise that was not a student's intended answer (e.g., a cough, background noise, the child counting out the solution to a problem) the researcher pressed both buttons simultaneously and the response time for that problem was not recorded. In this instance students heard a double buzz and saw the word "error" appear in the upper right hand corner of the screen. Once an answer had been scored, the screen went blank for three seconds until the next stimulus appeared. The response time and accuracy information for each problem for each student was written to a score file for each task that was readily accessible for conducting subsequent analyses.

Before each test, directions appeared on the screen that explained the nature of the task, informed the student about what to expect after a correct, incorrect, or spoiled trial, encouraged the student to respond as quickly as possible while still getting the correct answer, and asked if the student had any questions. Five practice questions were also constructed for each test and were administered after the directions to ensure each student was clear on what he/she would see and was expected to do. Upon completion of the practice problems each student was again asked if he/she had any questions and then proceeded to the actual test stimuli for which response time and accuracy were recorded.

Computer Stimuli

Addition Subtest. Each problem consisted of two whole numbers separated by an addition sign. There were two categories of problems. The first category included problems where both the addend and augend were single digit whole numbers greater than 0. The second category included problems where either the addend or augend was a single digit whole number greater than 0, and the other number was a two digit whole number less than or equal to 20. In the test bank of problems for category 1 there were 45 single digit plus single digit problems plus each problem's commutative complement (e.g., 2 + 5 and 5 + 2) for a total of 90 problems. In the test bank of problems for category 2 there were 99 single digit plus double digit problems (from 1 + 10 up to 9 + 20, inclusive) plus each problem's commutative complement for a total of 198 problems.

During the running of the task, a total of 20 problems was selected randomly by the computer, 10 randomly selected without replacement from the bank of problems in category 1 and 10 randomly selected without replacement from the bank of problems in category 2.

Subtraction Subtest. Each problem consisted of two whole numbers separated by a subtraction sign. There were two categories of problems. The first category included problems where both the minuend and subtrahend were single digit whole numbers greater than 0. The second category included problems where the minuend was a two digit whole

number less than or equal to 20 and the subtrahend was a single digit whole number greater than 0 but less than the minuend. In the test bank of problems for category 1 there were a total of 45 single digit minus single digit problems (from 1 - 1 up to 9 - 9, inclusive). In the test bank of problems for category 2 there were a total of 165 double digit minus single digit problems (from 10 - 1 up to 20 - 20, inclusive).

During the running of the task, a total of 20 problems was selected randomly by the computer, 10 randomly selected without replacement from the bank of problems in category 1 and 10 randomly selected without replacement from the bank of problems in category 2.

Multiplication Subtest. Each problem consisted of a single digit whole number greater than 0 separated by a multiplication sign. In the test bank of problems there were 45 single digit times single digit problems plus each problem's commutative complement (e.g., 2 x 5 and 5 x 2) for a total of 90 problems. During the running of the task, a total of 20 problems was randomly selected without replacement by the computer from the test bank of problems.

Hard Multiplication Subtest. Each problem consisted of one single digit whole number and one double digit whole number less than 20 separated by a multiplication sign (e.g., 17 x 7). For half of the problems the double digit appeared first and for the other half the single digit appeared first. The test bank contained a total of 20 problems. During the

running of the task problems were randomly selected without replacement from the test bank until the problem set in the test bank was exhausted.

Procedure

The CAAS basic arithmetic assessment was the first test administered and was given between November and February of the school year (it was conducted on a grade by grade basis starting with grade 5 and ending with grade 8). An area of a hallway near the principal's office was partitioned off and students were called out of their classes to be tested during a time when the hallways were quiet and relatively empty. The researcher began the computer portion of the experiment by introducing himself to the student and by briefly explaining what the student was about to do and how long it would take to complete. Each student completed the addition task first, followed by the subtraction, and multiplication tasks in that order. The researcher read the directions out loud as the student read them silently. After the instructions any questions the student had were answered before moving on to the practice problems. Once the practice problems were completed the student again had the opportunity to ask questions before moving on to the actual testing for which data were recorded. A short break between each of the subtests was provided if needed. The same procedure of reading directions, solving practice problems, and solving test problems was followed for each subtest until all 4 of them were completed. It took approximately 30 minutes to

complete the whole battery. Once a student finished testing he or she was thanked for participating. The written mathematics test was administered in March of the school year. All of the grades were administered the written test on the same day during class time, which was approximately 45 minutes.

SVT reading comprehension tests were administered earlier, towards the beginning of the school year one grade at a time in conjunction with a separate project that was being conducted at the middle school.

SVT tests, as they were answered as yes/no on an answer sheet that could be scanned, were computer scored and total scores for each student were subsequently computed by adding up the number of correct answers each student had on the test. The written math test was hand scored by the author and subtest scores were totaled for each student as the total number of items correctly solved. Item by item scoring for each student for the two tests was also recorded so that reliability estimates could be computed. CAAS arithmetic accuracy and response time data were recorded in score files by the CAAS program. After the written tests were scored and totaled, the three sets of measures were then matched by student and merged into one data file in preparation for forthcoming analyses.

In order to test the aforementioned predictions it was first necessary to gather reliability data on the instruments used to measure complex mathematical ability and reading

ability to determine if the tests had sound psychometric properties. Analysis of the data then proceeded in the following manner:

1.) Response time and accuracy data were analyzed by grade to note any developmental trends in this area.

2.) Response time and accuracy measures were combined into a single measure (to be described) to determine if any benefit might be gained from using such a measure in place of response time alone as a predictor of complex math problem solving ability.

3.) Multiple regression analyses were conducted to determine the significant predictors of written computational problem solving and word problem solving abilities as well as the other measures of complex mathematical abilities obtained from the school.

4.) Regression analyses were performed to determine whether there was a developmental trend in the basic mental arithmetic/complex mathematical problem solving relationship.

CHAPTER 4

RESULTS

Student Attrition

Before going into a detailed report of the data and the analyses conducted, first it is necessary to report some of the problems that were encountered while collecting both the CAAS and written test data. It will then be clear when the different pieces of data are reported and examined why the number of students fluctuated (at times dramatically) or differed from the 100 students that were originally administered the CAAS arithmetic tests.

The first problem encountered in data collection was during the administration of the CAAS arithmetic tasks. Several students did not complete all of the tasks due to a variety of interruptions. For example, one fifth grade student was interrupted by a fire drill after only having finished the addition task.

Another problematic event involved the administration of the written math test. Because the written test was given to all students in all grades on the same day as decided by the principal, some students never took the test due to absence, dismissal, etc. In total, only 83 of the students that completed the CAAS arithmetic battery also completed the written test.

The third and most debilitating problem in data collection involved the group of sixth grade teachers and the written math test. Apparently there was some

"miscommunication" between the principal and the sixth grade teachers at the middle school because, before giving the tests back to be scored the teachers intentionally removed all student identification information from the written tests. Without any way to match up written test scores with CAAS data, the 6th grade could not be included in the regression analyses involving CAAS arithmetic measures and the 50 item complex computation/word problem test that was developed for the study. A request was made to the principal to obtain standardized math test scores so that analyses could be done using all of the data from the CAAS arithmetic tasks. Unfortunately, scores from the ITBS were provided for the 6th grade only. Therefore, two sets of analyses will be reported, one set involving the 5th, 7th, and 8th grades and a separate set of analyses involving data from the 6th grade. Such are the perils of data collection.

Descriptives for Arithmetic and Written Test Data

CAAS Data

A summary of the accuracy and response time means and standard deviations for the CAAS arithmetic tasks is provided in Table 6 on page 111. The data reported there are for **all** students that took the CAAS arithmetic test. It was discovered during the testing that many of the students found the hard multiplication task too demanding and therefore many students did not attempt or complete that task. In total, barely half of the students were able to complete this

Table 6. CAAS arithmetic task means and standard deviations by grade.

Operation Grade	N	Mean (%) Accuracy	Accuracy Std. Dev.	Mean RT (sec.)	RT Std. Dev.
Addition					
Grade 5	27	92.5	8.94	2.50	1.06
Grade 6	28	96.6	5.62	1.82	.54
Grade 7	23	94.2	8.04	1.71	.95
Grade 8	21	94.6	6.68	1.84	.87
Subtraction					
Grade 5	26	89.9	7.23	2.48	.83
Grade 6	28	96.8	4.79	1.64	.42
Grade 7	23	90.9	9.81	1.67	1.10
Grade 8	21	94.7	6.02	1.84	.84
Multiplication					
Grade 5	24	82.3	13.89	3.75	2.20
Grade 6	27	96.1	5.63	1.97	.78
Grade 7	22	92.3	10.14	1.82	.59
Grade 8	21	86.7	11.71	1.97	.72

arithmetic task, and therefore only data from the other three arithmetic measures will be reported here.

There are several trends in the arithmetic data worth noting. As mentioned previously it has been noted (Mercer, 1979) that by 6th grade students on average have attained 95%

of adult accuracy (and only 65% of adult speed) on simple arithmetic tasks. In this sample, with the possible exception of multiplication, students from each grade were very accurate on average, scoring around and sometimes well above 90% on each of the arithmetic tasks. Accuracies, however, did not increase monotonically with grade level as was expected.

The response time data, however, was quite different. A developmental pattern was evident in this data. Arithmetic response times, as was predicted, decreased as a function of increase in grade. More specifically, Bonferonni t-tests indicated that the 5th grade response times for all arithmetic measures were significantly higher when compared individually to each of the other three grades (all t's > 2.90). None of the pairwise differences for the 6th through 8th grades were significant.

There are at least two, and possibly three explanations for the observed difference in arithmetic response time means. The first explanation is that there is a practice effect. Older students have had more time and opportunity to practice solving basic arithmetic problems both in and out of school. A second possible explanation involves the development of children's general cognitive processing speed. As Kail (1991) has demonstrated, overall cognitive processing speed increases over development according to an exponential function.

The third possible explanation involves the CAAS arithmetic results of the 6th graders sampled in this study. As an inspection of Table 6 indicates, 6th graders had the highest average accuracies and the lowest standard deviations for both accuracy and response time measures (with the exception of multiplication). As mentioned before, their response times on each of the arithmetic tasks were significantly faster than the 5th graders and were on par with the 7th and 8th graders. It is possible that the 6th graders sampled here may have been somewhat more adept at basic arithmetic than would be predicted. Perhaps during their schooling more emphasis was placed on automatizing basic arithmetic facts, and if a different group of 6th grade students were sampled the response times and their standard deviations may have been higher. In summary, it is probably likely that the arithmetic response time patterns obtained in this study resulted from a combination of differential practice, general cognitive speed differences, and the precocity of the 6th grade sample.

Written Math and SVT Test Scores

A summary of the word problem and complex computational subtests scores from the written math test as well as grade appropriate SVT tests scores are provided in Table 7 on page 114. Sixth grade written math test scores were not reported for reasons mentioned in the introduction to this chapter. Both sets of test items were scored dichotomously as correct or incorrect and the total number of correct items on each

Table 7. Mathematics test and SVT test means and standard deviations by grade.

Test Grade	N	# Correct	Standard Deviation
Word Problem Subtest			
Grade 5	23	12.3	4.72
Grade 7	18	19.1	4.87
Grade 8	13	16.1	5.22
Computation Subtest			
Grade 5	23	6.4	3.55
Grade 7	18	11.6	4.00
Grade 8	13	8.8	3.76
SVT			
Grade 5	22	75.9	12.57
Grade 6	25	77.8	9.99
Grade 7	17	73.5	9.89
Grade 8	13	78.2	7.89

test represented a student's score on that test. Cronbach's alpha and split half reliability measurements were estimated for both the math and SVT tests. For the written math test the alpha was .92 and the split half estimate of reliability was .82 and .91 when corrected for length of the test using

the Spearman-Brown prophecy formula. Reliability measures were calculated for each grade level SVT separately. Cronbach's alpha for the four SVT tests ranged from .64 to .77 with an average of .69. The split-half estimates ranged from .46 to .74 with an average of .59 (average Spearman-Brown correction was .73).

The above reliability estimates suggest that the internal consistency of the items on the math test were high, that is, a large proportion of the variance in students' test scores was a result of true score variance on this test. SVT reliabilities used in this study were considerably lower than expected, especially given the reliability data already collected on the SVT that yielded higher reliability estimates (e.g., Royer, 1990). Further examination of the test items is warranted if these particular versions of the SVT are to be used in measuring reading comprehension in the future.

In addition to conducting analyses on total scores, an analysis of errors was also planned. Students were given ample space on their answer sheets and were instructed to work out their answers to the math problems in that space. Unfortunately, many students on the majority of problems did not (systematically at least) write down their work and therefore a reliable categorization of errors was not possible and therefore was not pursued.

Turning to the data, again several points are worth noting. SVT tests, as they were grade appropriate in

difficulty, showed similar characteristics from grade to grade; overall means changed little by grade. The patterns of math subtest means at first, however, appear to be odd. As the two math subtests were largely based on material similar to what would appear on a 6th grade ITBS, one would expect that students in the upper grades would be able to solve more problems than students in the younger grades. The data show that means did increase on both math subtests from 5th to 7th grade but then dropped slightly for the 8th grade. This drop, however, can be explained at least partially by the shortened testing time the 8th graders were given. The eighth grade teachers explained after administration of the test that their testing time was unexpectedly cut short by approximately 10 minutes due to a school function. In light of this information the slight drop in means from 7th to 8th grade on the math subtests is understandable.

Due to the absence of sixth grade test data, scores from the ITBS math and reading sections were provided by the school reported in the form of national percentiles (eventually converted to standard scores for upcoming analyses). Summary descriptives statistics found that on average students from grade 6 in this study scored in the 65th percentile on the Math Computation subtest (Math Comp.), the 69th percentile on the Math Concepts and Estimation subtest (Math Conc.), the 64th percentile on the Math Problem Solving and Data Interpretation subtest (Math PSDI), and the 61st percentile on the Total Reading battery (Read. Tot.),

which was a combination of the Reading Comprehension and Vocabulary subtests of the ITBS.

Arithmetic and Written Test Score Raw Correlations

All of the CAAS arithmetic accuracy and response time data as well as written math and SVT data were correlated before regression analyses were conducted. One important piece of information that was needed from inspection of the correlation matrix was the strength of relationship the combined accuracy and response time measure had with complex mathematical measures in this study. The formula developed in the LATAS lab for combining accuracy and response time measures into a single index is as follows:

$$\text{Combined} = \sqrt{\{[(100 - \text{Accuracy})/SD_A]^2 + [\text{Response Time}/SD_{RT}]^2\}}$$

In effect, the index first changes the accuracy measure to an inaccuracy measure thereby making changes in accuracy mean the same thing as changes in response time. That is, greater inaccuracies mean poorer performance, and the lower inaccuracies mean better performance; the same can also be said for response time. The formula then divides both measures by their standard deviations which equates their scales and allows them to be squared, combined, and then rooted.

Correlation matrices using the arithmetic accuracy, response time, and combined measures, and written math and reading test measures were computed for the 5th, 7th, and 8th

grade sample (hereafter to be called the 578 sample) and the 6th grade sample separately. These two correlation matrices appear in Table 8 and Table 9 on pages 119 and 120, respectively. For ease of inspection, the correlations of the arithmetic response time measures and the combined arithmetic measures with the written math test variables and standardized math measures are presented in bold.

Examination of the arithmetic accuracy/written math test relationships revealed correlations that were not consistent with what would be expected if students were trading accuracy for speed on the CAAS tasks. For the 578 sample, correlations among arithmetic task accuracies and their corresponding response times ranged from -.17 to -.37, and for the 6th grade sample the correlations ranged from -.19 to -.69. These correlations show that the students that were faster at solving the basic arithmetic problems on average (decrease in RT) also tended to be the students that were more accurate at solving those problems (increase in accuracy). If many students were trading accuracy for speed the correlations would reflect that faster students also tended to be less accurate. In other words, one would expect the arithmetic accuracy/RT relationships to be close to zero or possibly weakly positive. Such was not the case.

Turning to a discussion of the correlations of arithmetic variables with the written math tests, I would first like to note arithmetic accuracy/written test relationships. As can be seen from Tables 8 and 9, the

Table 8. Correlations among arithmetic, math test, and reading variables for grades 5, 7, and 8.

	WP Score	Comp. Score	Add. RT	Add. Acc.	Add. Comb.	Sub. RT	Sub. Acc.	Sub. Comb.	Mult. RT	Mult. Acc.	Mult. Comb.
Comp. Score	.76										
Add. RT	**-.56**	**-.52**									
Add. Acc.	.14	.06	-.17								
Add. Comb.	**-.43**	**-.32**	.67	-.82							
Sub. RT	**-.67**	**-.55**	.77	-.13	.53						
Sub. Acc.	.20	.04	-.03	.23	-.21	-.20					
Sub. Comb.	**-.45**	**-.29**	.39	-.25	.42	.65	-.86				
Mult. RT	**-.58**	**-.51**	.77	-.04	.47	.85	-.08	.48			
Mult. Acc.	.35	.26	-.26	.32	-.36	-.37	.51	-.57	-.37		
Mult. Comb.	**-.54**	**-.46**	.63	-.21	.50	.73	-.36	.64	.82	-.82	
SVT	.14	.04	-.19	-.06	-.04	-.25	.09	-.21	-.12	0	0

Table 9. Correlations among arithmetic and ITBS math and reading variables for grade 6.

	Math Comp.	Math Conc.	Math PSDI	Read. Tot.	Add. RT	Add. Acc.	Add. Comb.	Sub. RT	Sub. Acc.	Sub. Comb.	Mult. RT	Mult. Acc.	Mult. Comb.
Math Conc.	.79												
Math PSDI	.72	.76											
Read. Tot.	.54	.52	.61										
Add. RT	**-.65**	**-.79**	**-.70**	-.26									
Add. Acc.	.13	.26	.37	.03	-.38								
Add. Comb.	**-.41**	**-.60**	**-.61**	-.21	.83	-.81							
Sub. RT	**-.67**	**-.57**	**-.62**	-.39	.75	-.25	.60						
Sub. Acc.	.23	.33	.55	.19	-.37	.56	-.58	-.19					
Sub. Comb.	**-.43**	**-.43**	**-.70**	-.35	.67	-.54	.72	.70	-.80				
Mult. RT	**-.51**	**-.56**	**-.50**	-.09	.68	-.14	.51	.46	-.33	.44			
Mult. Acc.	.25	.39	.41	.18	-.45	.41	-.40	-.31	.43	-.48	-.69		
Mult. Comb.	**-.40**	**-.51**	**-.45**	-.11	.67	-.19	.50	.43	-.39	.62	.98	-.83	
SVT	.51	.43	.65	.62	-.43	.06	-.26	-.31	.20	-.32	-.28	.23	-.26

correlations between arithmetic accuracies and both math tests were positive and tended to be low ranging from .06 to .35 for the 5th ,7th, and 8th sample, and ranging from .13 to .55 for the 6th grade sample.

Arithmetic response time measures on the other hand had a much stronger relationship with the written math tests in both samples. For the 578 sample the correlations with the written math test measures ranged from -.51 for multiplication RT to -.67 for subtraction RT, and for the 6th grade sample from -.50 for multiplication RT to -.79 for addition RT; both were higher than what was recorded for the accuracy measures. Looking at the relationships between the combined arithmetic measures and written test performance in bold in the tables it is worth noting that these values tend to be (nonsignificantly) smaller than or equal to arithmetic RT/written test correlations. Taking into account what we already know about the accuracy data, namely that arithmetic accuracies in general were high, there appears to be little or no accuracy for speed trade-off, and that accuracies tended to have low positive correlations with the written math tests, it makes sense that the combined arithmetic measure in these samples is not a better predictor of written math test performance than arithmetic response time alone.

This is not to say that the combined measure is not useful, however, in other contexts or with other samples of students. For instance, in learning disabled populations the combined measure has been very useful because student

accuracies tend to be much lower and variable than in non-disabled populations, and response time alone is not the best indicator of arithmetic performance and possibly more complex problem solving performance (e.g., Royer & Tronsky, 1997). The combined measures also may be informative in students from elementary grades where strategy use and accuracies may be better indicators of arithmetic/mathematics performance (e.g., Geary & Burlingham-Dubree, 1989). Nevertheless, for the purpose of this study it is safe to conclude that response time alone is at least as good a predictor of complex computational and word problem solving ability as the combined measure and therefore will be used in the regression analyses that follow.

Statistical Analyses

Due to the many problems that were encountered during the administration and scoring of the CAAS arithmetic and written math tests it was necessary to alter some of the planned analyses. As Meyer and Well (1995) have noted, if we are to take seriously the multiple correlations that result from regression analyses it is important that the ratio of number of cases compared to number of predictor variables is large (in some cases 30 or more according to Meyer and Well). When the aforementioned ratio is small a researcher is capitalizing on chance when reporting R_{sample} as it grossly overestimates $R_{population}$ (other regression issues will be taken up later in the concluding chapter).

Because the expected N/p ratio was greatly reduced due to subject/task attrition and the inability to match the word problem/complex computation written test results to CAAS arithmetic data for the 6th grade, the three arithmetic RT variables were combined. This was done by first standardizing each RT variable (converting RT's to z-scores) using the mean and SD of each arithmetic variable across the whole sample. Then the z-scores were added together to form one arithmetic RT variable to be used in all analyses which from now on will be referred to as the "arithmetic aggregate variable (ArithAgg)." In addition to addressing the N/p issue the combining of arithmetic response time variables also addresses another issue in regression, the issue of multicollinearity. If predictor variables are highly intercorrelated it can lead to inflated standard errors of regression coefficients. Returning to the correlation matrices in Tables 8 and 9, it is evident that the three arithmetic response time variables are highly intercorrelated--most of the intercorrelations are .68 or higher. Combining the three arithmetic RT's into one variable eliminates the potential problems of using highly intercorrelated predictor variables in regression analyses.

Analysis of 5th, 7th, and 8th Grade Data

Regression Analyses

For a quick summary of the regression analyses for the 578 data refer to Table 10 (computation problem subscore as the criterion variable) and Table 11 (word problem subscore

as the criterion variable) on pages 125 and 127. For each of the analyses information is given about criterion variables, predictor variables, sums of squares, mean squares, F ratios and their associated probabilities, and R^2 and adjusted R^2 values.

<u>Computation Subscore</u>. Number of computation problems solved correctly by each student was used as the criterion variable in the first set of analyses. When entered into a regression analysis separately both ArithAgg [$F(1, 49) = 21.31$, $p < .0001$] and grade [$F\ (2, 51) = 9.39$, $p < .001$] were significant predictors of computation problem solving (grade was coded as a dummy variable in this and subsequent regression analyses). Only SVT when entered alone did not capture a significant proportion of variance [$F\ (1, 49) = .10$, ns].

When SVT was entered in a regression analysis along with ArithAgg, the partial F revealed that SVT did not account for any variance in the criterion over and above ArithAgg and in total the two variables accounted for approximately 26% of the variance in computation subscores. When ArithAgg, SVT, and grade were then entered into the same regression equation together both ArithAgg [$F\ (1, 43) = 5.44$, $p < .05$] and grade [$F\ (2, 43) = 8.89$, $p < .01$] accounted for a significant proportion of variance in the criterion, while SVT failed to account for a significant amount of variance in the criterion [$F\ (1, 43) = 2.02$, ns]. In total, the three predictor

variables accounted for 46% of the variance in computation subscore.

Table 10. Summary of regression analyses for 5th, 7th and 8th grade data with computation subscore as the criterion.

Source	df	Sum of Squares	Mean Square	F	p	R^2	Adj. R^2
ArithAgg	1	255.32	255.32	21.31	< .0001	.30	.29
Residual	49	587.03	11.98				
SVT	1	1.93	1.93	.10	ns	0	0
Residual	49	958.11	19.55				
Grade	2	265.04	132.52	9.39	< .001	.27	.24
Residual	51	719.77	14.11				
ArithAgg	1	238.47	238.47	18.60	< .001		
SVT	1	.06	.06	0	ns		
Regression	2	241.67	125.69	9.42	< .001	.30	.26
Residual	45	577.00	12.94				
ArithAgg	1	51.64	51.64	5.44	< .05		
SVT	1	19.22	19.22	2.02	ns		
Grade	2	168.74	84.37	8.89	< .01		
Regression	4	408.52	102.13	10.26	< .0001	.50	.45
Residual	43	438.01	9.95				
ArithAgg	1	141.41	141.41	22.67	< .001		
SVT	1	27.77	27.77	4.45	< .05		
Grade	2	25.17	12.59	2.02	ns		
Arith x Grade	2	152.54	76.27	12.23	< .001		
Regression	6	512.60	102.52	13.99	< .0001	.69	.64
Residual	41	255.72	6.24				

Finally, when the interaction of ArithAgg with grade was added to the other three variables in the regression equation, the partial F-test revealed that the interaction accounted for significant amount of variance in the criterion (approximately 19%) over and above the other three predictor

variables [$F(1, 41) = 12.23$, $p < .001$]. A total of 64% of the variance in the criterion was accounted for using the four variables as predictors.

This final analysis indicates that the relationship between arithmetic response time and computation subscore changed from grade to grade. Looking at correlations grade by grade revealed that in the 5th grade the correlation between the two variables was moderate (.39), sharply increased for 7th graders (.74), and once again dropped off in the 8th grade (.44). Due to the small samples at each grade (n's were 20, 18, and 13 for the 5th, 7th, and 8th grade, respectively) it is dangerous to draw any hard and fast conclusions from the changing correlations. A possible explanation will be detailed, however, in the conclusion to this analysis section.

<u>Word Problem Subscore</u>. Number of word problems solved correctly by each student was used as the criterion variable in the second set of regression analyses. Again, when entered into a regression analysis separately both ArithAgg [$F(1, 49) = 32.93$, $p < .0001$] and grade [$F(2, 51) = 9.88$, $p < .001$] were significant predictors of computation problem solving. Only SVT when entered alone did not account for a significant proportion of variance [$F(1, 49) = .95$, ns].

When SVT was entered in a regression analysis along with ArithAgg, the partial F-test revealed that SVT did not account for any variance in word problem subscore over and

Table 11. Summary of regression analyses for 5th, 7th and 8th grade data with word problem subscore as the criterion.

Source	df	Sum of Squares	Mean Square	F	p	R^2	Adj. R^2
ArithAgg	1	644.68	644.68	32.93	< .0001	.40	.39
Residual	49	959.24	19.58				
SVT	1	230.50	30.50	.95	ns	.02	0
Residual	49	1566.25	31.96				
Grade	2	473.03	236.52	9.88	< .001	.28	.25
Residual	51	1220.30	23.93				
ArithAgg	1	538.64	296.10	14.84	< .0001		
SVT	1	16.64	16.64	.82	ns		
Regression	2	589.88	294.94	14.46	< .0001	.40	.36
Residual	46	918.00	19.96				
ArithAgg	1	157.41	157.41	10.03	< .01		
SVT	1	86.71	86.71	5.53	< .05		
Grade	2	243.37	121.69	7.76	< .01		
Regression	4	833.25	208.31	13.28	< .0001	.55	.51
Residual	43	678.17	15.41				
ArithAgg	1	175.56	175.56	12.23	< .01		
SVT	1	100.39	100.39	6.99	< .05		
Grade	2	11.13	5.56	.39	ns		
Arith x Grade	2	86.00	43.00	3.00	ns		
Regression	6	919.25	153.21	10.67	< .0001	.61	.55
Residual	41	588.56	14.36				

above ArithAgg, and in total the two variables accounted for approximately 36% of the variance in computation subscores. When ArithAgg, SVT, and grade were then entered into the same regression equation together ArithAgg [F (1, 43) = 10.03, p < .01], grade [F (2, 43) = 7.76, p < .01], and SVT [F (1, 43 = 5.53, p < .05] accounted for a significant proportion of unique variance in the criterion. In total, the three

predictor variables accounted for 51% of the variance in word problem subscore.

Finally, when the interaction of ArithAgg with grade was added to the other three variables in the regression equation, the partial F-test revealed that the interaction did not account for significant proportion of variance in the criterion (approximately 4%) over and above the other three predictor variables [$F(1, 41) = 3.00$, $p = .06$]. A total of 55% of the variance in the criterion was accounted for using the four variables as predictors.

While the final analysis with the interaction term only approached significance, it did indicate that the relationship between mental arithmetic response time and word problem solving ability might be changing with development. An examination of the correlations between the two variables by grade yielded r's of .60 at the 5th grade, .71 at the 7th grade, and .34 at the eighth grade. Once again it must be kept in mind that it is dangerous to draw any hard and fast conclusions from the changing correlations due to each grade's small sample size.

Analysis of 6th Grade Data

Data Preparation

Before analyses could be conducted, the ITBS scores for each child on each of the four standardized tests needed to be converted from national percentiles into z-scores. This was done by finding the z value from a z distribution that

corresponded to the national percentile that a student received on a particular test.

Descriptions of Standardized Tests

The three math subtest scores from the ITBS that were used as criterion variables in regression analyses in this study need to be explained in more detail. The Math Computation subtest was very similar to the researcher generated written computation test in that it contained problems involving the addition, subtraction, multiplication, and division of (multidigit) whole numbers, fractions, and decimals. The Math Concepts and Estimation subtest involved, but was not limited to, problems that tested students' number sense, ability to estimate (sums, products, and quotients), knowledge of geometry (polygons, similar figures, angles, area, etc.), knowledge and conversion of decimals and fractions, ability to solve open problem sentences, and number pattern recognition. The Math Problem Solving and Data Interpretation subtest required students to solve word problems similar to those in the researcher produced written word problem test and also required students to solve problems by reading and interpreting graphical information. The Reading Total test was composed of a Vocabulary and Reading Comprehension subtest. For the Vocabulary subtest, students were given a phrase with a bold-faced word in it for which they were to select from a set of 4 choices an appropriate synonym. For the Reading Comprehension subtest, students were required to answer multiple choice questions

after reading either several paragraphs of text or several lines of poetry.

Some of the problems in both the Concepts and Estimation and Problem Solving and Data Interpretation sections were either computationally simple or asked questions that were void of computations altogether. Several questions in the Problem Solving section asked the student to determine what information other than what had already been given in the problem was needed to solve it. The researcher generated written test differed in this regard as it contained problems that all needed to be solved using computation at some point. The results of these analyses will be of note in that a strong relationship between basic arithmetic response time and performance on the conceptual sections of the ITBS Math test will offer further evidence that the relationship is not merely due to the fact that the conceptual problems required computations in their solutions.

<u>Regression Analyses</u>

For a quick summary of the regression analyses for the 6th grade data refer to Table 12, Table 13, and Table 14 on pages 131, 132, and 133, respectively. For each of the analyses information is given about criterion variables, predictor variables, sums of squares, mean squares, F ratios and their associated probabilities, and R^2 and adjusted R^2 values.

<u>Math Computation</u>. When the arithmetic aggregate (ArithAgg) variable was used as a predictor alone it

accounted for 45% of the variance in the criterion ITBS Mathematics Computation subscore [F (1, 24) = 21.36, p < .001].

Table 12. Summary of regression analyses for 6th grade data with ITBS Math Computation score as the criterion variable.

Source	df	Sum of Squares	Mean Square	F	p	R^2	Adj. R^2
ArithAgg	1	8.11	8.11	21.36	< .001	.47	.45
Residual	24	9.11	.38				
Read. Total	1	5.25	5.25	10.41	< .01	.29	.27
Residual	25	12.61	.50				
ArithAgg	1	5.67	5.67	18.79	< .001		
Read. Total	1	2.18	2.18	7.23	< .05		
Regression	2	10.29	5.15	17.07	< .0001	.60	.56
Residual	23	6.93	.30				

Reading Total as a lone predictor variable accounted for 27% of the variance in Math Computation subscore [F (1, 25) = 10.41, p < .01]. When the two predictors were entered into a regression equation together they accounted for 56% of the variance in the criterion [F (2, 23) = 17.07, p < .0001]. ArithAgg also accounted for a significant portion of unique variance in the criterion when Reading Total was entered in the regression equation [F (1, 23) = 18.79 p < .001]. Similarly, Reading Total also accounted for a significant proportion of variance in the criterion when ArithAgg was held constant [F (1, 23) = 7.23 p < .05].

Math Concepts and Estimation. The ArithAgg variable when entered alone as a predictor variable accounted for 50% of the variance in the criterion ITBS Math Concepts and Estimation subscore [F (1, 24) = 26.42, p < .0001].

Table 13. Summary of regression analyses for 6th grade data with ITBS Math Concepts and Data Estimation score as the criterion variable.

Source	df	Sum of Squares	Mean Square	F	p	R^2	Adj. R^2
ArithAgg	1	8.43	8.43	26.42	< .0001	.52	.50
Residual	24	7.66	.32				
Read. Total	1	4.47	4.47	9.26	< .01	.27	.24
Residual	25	12.06	.48				
ArithAgg	1	6.13	6.13	23.75	< .001		
Read. Total	1	1.73	1.73	6.69	< .05		
Regression	2	10.16	5.08	19.69	< .0001	.63	.60
Residual	23	5.93	.26				

Reading Total entered alone as a predictor accounted for 24% of the variance in the criterion [F (1, 25) = 9.26, p < .01], and when both predictors were entered together in a regression equation they accounted for 60% of the variance [F (2, 23) = 19.69, p < .0001]. ArithAgg accounted for a significant amount of variance in the criterion with Reading Total held constant [F (1, 23) = 23.75, p < .001], and Reading Total accounted for a significant amount of variance with ArithAgg held constant [F (1, 23) = 6.69, p < .05].

Math Problem Solving and Data Interpretation. The same three regression analyses run with the previous two criterion variables were also run with Math PSDI subscore as the criterion with similar results. ArithAgg alone accounted for 44% of the variance in the criterion [F (1, 24) = 20.98, p < .001], Reading Total alone accounted for 35% of the variance in the criterion [F (1, 25) = 14.92, p < .001], and the two

Table 14. Summary of regression analyses for 6th grade data with ITBS Math Problem Solving and Data Interpretation score as the criterion variable.

Source	df	Sum of Squares	Mean Square	F	p	R^2	Adj. R^2
ArithAgg	1	8.02	8.02	20.98	< .001	.47	.44
Residual	24	9.17	.38				
Read. Total	1	6.79	6.79	14.92	< .001	.37	.35
Residual	25	11.38	.46				
ArithAgg	1	5.25	5.25	19.88	< .001		
Read. Total	1	3.09	3.09	11.70	< .01		
Regression	2	11.11	5.55	21.01	< .0001	.65	.62
Residual	23	6.08	.26				

predictor variables entered together accounted for 62% of the variance in the criterion [F (2, 23) = 21.01, p < .0001]. Once again ArithAgg accounted for a significant proportion of variance in the criterion with Reading Total held constant [F (1, 23) = 19.88, p < .001], and Reading Total accounted for a significant proportion of variance with ArithAgg held constant [F (1, 23) = 11.70, p < .01].

Summary of Data Analysis

In analyses for both samples of data with all of the criterion measures one result stands out above all of the others: the arithmetic response time aggregate variable accounted for a significant proportion of variance in the criterion measures. It accounted for a significant proportion of variance whether entered alone or in combination with other predictor variables. The response time aggregate was a significant predictor regardless of whether the criterion was a computational or problem solving measure and whether the criterion measure was performance on a well known standardized test or a researcher/teacher constructed word problem test.

Other results were not so consistent. It was expected that the reading measures used in the two samples should function differently depending on the criterion measure. It was expected that the reading measures would have a strong relationship with any mathematics problem solving ability measure that involved processing any amount of text--a finding that has already been well established (e.g., Balow, 1964; Kintsch & Greeno, 1985; Muth, 1984). Conversely, it was predicted that any measure such as the written complex computation measures that were void of text would not have a significant relationship with the reading measures. Results of the data sets for the two samples varied within and between samples. Within the 578 sample, SVT did or did not account for a significant amount of unique variance in the

criterion depending on the variables that were entered in any given regression equation. This held true whether the criterion was computation or word problem subscore.

The data analyses for the 6th grade, in turn, differed from the 5th, 7th, and 8th grade analyses mentioned above. In these analyses it made no difference which of the three criterion measures was being used and whether or not the arithmetic response time aggregate variable was entered as a predictor along with the reading measure. In each case, both the arithmetic response time and reading measures were significantly related to the computational and conceptual/problem solving dependent variables.

It is not difficult to explain why the reading measure might be a significant predictor of computation subscore on the ITBS. The reading total measure may have served as a proxy for a number of variables, most likely either IQ or, more interesting in the context of the present study, working memory resources. What is more difficult to explain is why the reading measures in both samples did not function similarly. The low correlations of SVT and the researcher generated written math test (.04 and .14) is in stark contrast to the very high correlations of reading measure and mathematics measures from the standardized tests (.52, .54, and .61) even taking into account the different sample sizes. It would be tempting to claim that the two reading measures were functioning differently in the two samples, however, the two reading test were also fairly highly correlated (.62).

Another possibility is that the reading difficulty of the researcher constructed written math test was lower for the 578 sample than the ITBS math tests were for the 6th graders. Because the researcher constructed word problem test was largely based on a 6th grade ITBS problem solving test section, it could be that the reading difficulty of the test was not much of a factor for the older (7th and 8th) grades that took the test, thus lowering the correlation of SVT with the math subscores. One additional explanation for the discrepant reading correlations is that the reliability estimates of the SVT tests were somewhat low overall which leads to an underestimation of the relationship between reading comprehension and the two dependent variables.

The final result that deserves mention is the pattern of arithmetic response time/written test subscore relationships across grades that was found. The interaction of grade and ArithAgg accounted for a significant proportion of variance above other predictors when computation subscore was the criterion and just missed conventional significance when word problem subscore was the criterion. When these relationships were examined at different grade levels, the patterns of correlations across grades was similar, starting out moderate at grade 5, peaking at grade 7, and dropping down again at grade 8. Small sample sizes notwithstanding, this pattern is close to what might be expected.

As was mentioned in the introduction to this study, the argument is that automatizing basic math facts frees up

working memory to concentrate on more complex aspects of problems whether they be word problem or complex computational problems. At younger grade levels where students are still using solution strategies other than retrieval for arithmetic problems that are slow and produce highly variable response times, RT measures might yield low or moderate correlations with more complex mathematical problem solving abilities. Once a larger proportion of children at later grades automatize the arithmetic facts, response time may become a more powerful predictor of complex math problem solving and yield much higher correlations. Once all or almost all children have automatized arithmetic facts at later grades (and general processing speeds have become faster and less variable with development) and the range of response times is restricted, it leads to a decrease in the arithmetic RT/complex math problem solving correlation. This might be what has been demonstrated by the pattern of correlations in this study across grades although with such a small sample such a generalization is not warranted without further investigation.

CHAPTER 5

DISCUSSION

The discussion section will be laid out in three main sections. First it is necessary to review some of the principles of regression to critically examine the significance of the results that have been reported. The second section of the discussion will then focus on what can be said with confidence about the results of the study and will place the findings of this study in their appropriate context within research already conducted in this domain. Third, and maybe most importantly, several recommendations about future research involving basic mental arithmetic, working memory, and higher order mathematical problem solving separately, and as a unit, will be examined.

The Limitations of Correlational Research

Regression analyses must always be interpreted and generalized from very carefully due to their nature. The most fundamental caution about correlational research (that is even today often overlooked) is that significant relationships that are found do not say anything about causation. I have been very careful in my claims not to state or imply that the research conducted here shows that students that are able to mentally solve arithmetic problems quickly **causes** them to be better complex mathematical problem solvers by freeing up working memory resources. If I were to conclude that I would be violating the most basic assumption about correlational research.

Another important issue in correlation has to do with sample size. Regression analyses are very susceptible to being influenced by a small number of data points when the number of subjects is small from the outset. Reference was made earlier to this issue of stability in regression and it was noted that if one really wants to take regression coefficients seriously, it may be necessary to have as many as 30 subjects for each predictor value entered in an analysis (Meyer & Well, 1995).

It has been indicated that student attrition was very high in the study. The first problem concerned an inability in some cases to run students through all the CAAS arithmetic tasks and absenteeism was a small problem in administering the written math test. Sample size issues were then made even more problematic when teachers made some of the 6th grade data unanalyzable. Steps were taken to try to reduce the instability created by the above data collection problems (and address the multicollinearity of predictor variables issue) by combining variables that were highly related. While the resulting N/p ratios were improved, they still were not as high as desired and therefore caution needs to be exercised when interpreting the results.

Sampling issues do not end with the N/p issue, though. Correlation coefficients are sample specific. That is, the estimation of the correlation of variables in a population are heavily influenced by the variability of the variables in a sample. If the current study was undertaken at a different

school, using different grades in the sample, and larger samples at each grade the results and conclusions could be slightly or very different.

A final issue I would like to mention is another sampling issue, the sample of variables that are included in analyses. The relative importance of a predictor variable's relationship to the criterion is heavily influenced by the other variables that are included in a regression equation. Usually variables are at least somewhat correlated with other variables being examined in a particular study, or for that matter other variables that were left out of the study. The predictive power of variables in a regression analysis will fluctuate, often wildly, depending on their relationship to the other predictor variables also entered in the analysis. It would be very difficult to conclude by this study whether or not reading is an important predictor of computation and/or word problem solving abilities, especially looking at the results of the 5th, 7th, and 8th grade analyses. Depending on which variables were entered into those regression analyses, the SVT variable was at times a significant predictor of the criterion and at other times was not. Thankfully there is existing evidence that has been accumulated that may help to explain these fluctuations.

Conclusions and Implications of the Current Study

The most important result that was obtained in this study was that the arithmetic response time variable that was created was a significant predictor of complex mathematical

problem solving ability as measured by two separate tests. Which variables were included in the several regression analyses did not affect the significance of the relationship. In light of what was mentioned above about correlation of predictor variables, an argument could be made that if other variables were included in the analyses the relationship of arithmetic RT and math problem solving would be rendered negligible or, at the very least, would be greatly diminished. Certainly a claim could be made that IQ as a predictor would greatly reduce the variance accounted for by arithmetic RT.

I do not deny that IQ would also be highly predictive of complex problem solving abilities in math. One of the subtest of the WISC-R is an arithmetic subtest so the two variables are inextricably linked. I am going to argue, however, that several pieces of evidence currently exist indicating that basic numerical abilities are related to complex mathematical skills after the effects of IQ are partialed out. In the aforementioned research conducted by both Balow (1964) and Zentall (1990) IQ effects were partialed out of subsequent analyses. In Balow's study complex computational measures along with reading ability were still strongly related to a standardized measure of math conceptual/problem solving ability. In Zentall's study the number of simple arithmetic problems solved on a paper and pencil test was (the only) significant predictor of number of word problems correctly solve when IQ was partialed out.

Other evidence also exists in previous studies of individual differences. As Geary (1994) notes, there have been numerous factor analytical studies conducted that identify a number or number facility factor that includes basic arithmetic abilities. Several researchers (e.g., French, 1951; Pawlik, 1966) have gone so far as to claim that this factor is the best confirmed and clearest aptitude factor of all. Other researchers (Jensen, 1994; Spearman, 1927) have noted that arithmetic skills have much that is shared over and above general intelligence. For Spearman, who was a strong supporter of the explanation of individual difference in mental abilities by *g* or general intelligence, this is a pretty strong statement.

Other findings from the present study that have already been mentioned are somewhat more muddled. The fact that the nature of the relationship between reading ability and word problem solving has already been firmly established in several other studies (e.g., Jerman & Mirman, 1974; Kintsch & Greeno, 1985; Muth, 1984) and in this study in the 6th grade, tends to overshadow the findings that SVT was not highly correlated with word problem subscore. The low correlation is further deemed suspect as SVT is strongly correlated with the standardized measures of problem solving (most highly with the word problem and data interpretation subtest, r = .65) as shown by the 6th grade data in Table 9. Reading ability's relationship to computational ability is still unknown as previous research on that relationship has not

been conducted. The conflicting results in this study do not help much although the possibility that reading measures are measuring working memory resources and therefore should be related to computational solving abilities would also be an interesting hypothesis to test.

A brief comment on the significant interaction effect is appropriate here as well. Again let me underscore that the changing arithmetic/complex math ability relationships over grade are highly suspect due to the very small samples available for analysis at each grade. It does raise some interesting possibilities for future research, though. Would this same effect be found in a more stable analysis with larger samples and if so when (and why) does this relationship change over development? Possibly the correlation change is showing when most children have switched from solving many basic arithmetic problems by slow counting methods to only or mostly using a retrieval strategy which is faster.

In summary, at the very least it has been established that there is a strong relationship between speed of mental arithmetic problem solving and more complex mathematical problem solving. Previously, the idea that computation abilities are predictive of more complex problem solving abilities has either been approached using complex computation abilities (Balow, 1964; Muth, 1984) as predictors or more inaccurate paper and pencil measures of basic arithmetic (Zentall, 1990). This study has established that

an even more basic arithmetic/complex problem solving relationship exists and warrants further investigation in a controlled experimental setting. Now let me again turn to an exposition of reasons why the mental arithmetic/math problem solving connection may exist.

Working Memory, the Development of Skilled Reading and the Development of Skills in Mathematics

There is really nothing to add to the working memory account of why the arithmetic/math problem solving relationship exists as it was extensively outlined in the introduction to the study and nothing in this study can directly add to that account. I would, however like to borrow an example from the domain of reading to illustrate how skilled reading develops, how an analogous progression may occur in mathematics, and how working memory is an important measure in the development of skill in both domains.

According to Royer and Sinatra (1994) there are several component skills a person must possess if he or she is to be a good reader. Those component skills are:

1.) Enabling skills--phonological awareness and the ability to identify letters.

2.) Word identification--the ability to map phonemes on to letters and take those letter-sound combinations and identify them as words. This is highly dependent on working memory and must eventually become automatic for someone to become a fluent reader.

3.) Activation of meaning--development of automatic activation of word meanings

4.) Syntactic and semantic processing--syntactic processing is usually fully developed through speech experience

free from language impairment before the development of reading skills, and semantic processing or text modeling requires several higher level skills such as making unconscious and conscious inferences to pull meaning of texts together and accessing related knowledge from memory.

5.) Prior knowledge and metacognitive processing--meaning of text is constructed through an interaction of the message within a text, the prior knowledge of the reader, and the particular context in which reading occurs. Metacognitive processing is the monitoring of one's comprehension to determine if/where it is failing.

These five component processes are arranged in a hierarchy and at least up through the level of syntactic processing are seen as encapsulated skills that cannot be affected by conscious strategic activities (in a skilled reader). Failure at one of the lower component skills leads to a mushrooming impairment in reading termed the "Matthew effect" (Stanovich, 1986). People that do not develop the necessary automatic skills at the word recognition level, for instance, develop overall reading skills and educational competencies that significantly lag behind those people that develop automatic word identification skills, and without remediation this disparity rapidly widens. The delay in development or lack of automatic word identification skills has an enormous impact on other components of the reading process, and as each successive component skill depends on the previous skill(s), it is not difficult to see why the "Matthew effect" occurs. It is usually the case as well that even if repair at a lower level of skill occurs it does not necessarily transfer to the repair of higher level skills. Direct repair of the higher level skill may also be needed in

addition to repair of the lower level skill for such a person to fully become a competent reader.

It is a very distinct possibility that something similar happens in the domain of mathematics--that there is a hierarchy of more and more complex skills, each one building on previously learned skill(s) that leads to the building of a competent mathematician. The simplified path of skills might go something like this:

number representation---> counting knowledge---> arithmetic ability---> more complex procedural and conceptual knowledge---> higher mathematical reasoning (problem solving)

It should be noted that this is a simplified version of what happens. Many other factors feed into this path of skills. For instance, (less complex) conceptual and procedural knowledge influences counting knowledge and arithmetic ability and working memory probably influences several skills such as arithmetic skills and even higher mathematical reasoning (see Geary, 1994, for a detailed diagram and Kaye, 1986, for a similar account of the above path of mathematics skills).

Although simplified, the above path should function similarly to the previously mentioned hierarchy of the development of reading skills. It has already been shown that arithmetic problems have many of the automatic skill properties that have been attributed to word recognition (e.g., Lefevre et al., 1988; Lefevre et al., 1994; Rogers & Fisk, 1991) and the idea of information encapsulation

therefore may be true in mathematics as well. The development of skill in math in general should be similar to reading in that if acquisition of any of the component skills shown above is impaired or delayed, there may be significant implications for the learning of higher order skills in the domain.

An interesting addendum to what has been mentioned above is that both the arithmetic and mathematical reasoning abilities of people in the United States has in recent years been shown to lag behind other nations (e.g., Stevenson, Chen, & Lee, 1993). Geary (1994) has also noted that arithmetic abilities are poorer than what they were 40 or 50 years ago when drill and practice of basic arithmetic facts was much more prevalent in the classroom. Another interesting finding from Geary, Salthouse, Chen, and Fan (1996) was that a comparison of samples of older American and Chinese adults (57 to 85 years of age) showed no differences in arithmetic abilities, perceptual speed, and spatial ability while younger American adults (college age) performed significantly poorer on arithmetic task than Chinese younger adults but performed equally on the other two ability measures. These findings seem to indicate that Americans' poorer arithmetic knowledge is a recent phenomenon.

More evidence is coming to light about the inadequacy of adults' basic arithmetic abilities. Until recently, many researchers have assumed that by adulthood all of the basic arithmetic facts have been automatized (e.g., Ashcraft, 1992;

Widaman & Little, 1992). Recent research suggests that this is not the case; it appears that adults still use counting strategies rather than retrieval to solve many simple arithmetic problems (Lefevre, Sadesky, & Bisanz, 1996; Lefevre, Bisanz, Daley, Buffone, Greenham, & Sadesky, 1996; Geary, 1996). It could be coincidental that both arithmetic and math reasoning abilities have worsened over the years in America reflecting overall educational neglect. It seems very likely, however, that the two go hand in hand. At least in part, declines in complex math problem solving ability are most likely attributable to similar declines in arithmetic ability.

The Focus of Future Research in Arithmetic and Mathematics

Working Memory Issues

In light of the research evidence presented above, it is apparent that additional research is needed involving working memory and its role in simple mental arithmetic. Adult participants have been used in experiments that have attempted to measure the amount (or lack of) working memory resources that are used during simple mental arithmetic. It has been assumed that the adults sampled in the aforementioned studies had already committed simple arithmetic problems to memory (e.g., Ashcraft, Donley, Halas, Vakali, 1992; Lemaire, Abdi, Fayol, 1996). Several pieces of evidence indicate that many adults still use counting strategies to solve some simple arithmetic problems (Lefevre et al., 1996; Lefevre, Sadesky, & Bisanz, 1996; Geary, 1996).

Arithmetic working memory resource studies may therefore be estimating working memory loads that are imposed when adults are still using counting strategies, not when adults are using retrieval. Recent work conducted by Klapp et al. (1991) using alphaplication has shown that when adults were trained to the point of automaticity and "beyond" on their arithmetic-like task, the amount of working memory resources used when a dual task load was imposed was reduced to an insignificant amount. What needs to occur in the future are studies that involve training participants in mental arithmetic to the point of automaticity and beyond, measuring of participants' strategy use/degree of automaticity, and then assessing working memory load in mental arithmetic via various dual task methodologies.

Arithmetic and Complex Mathematics: Experimental Evidence

Obviously, to better study the arithmetic/complex mathematics ability relationship it is necessary to move the research into the experimental realm. To my knowledge only one experimental research project has been conducted that gave practice to children on computational problems via a computer and sought to test improvement in computation and conceptual knowledge in mathematics as indicated by a standardized test (Suppes, Jerman, & Brian, 1968; Suppes & Morningstar, 1972). This project, however, was conducted using practice on complex computational problems. Experiments need to be conducted where children are trained on the basic math facts and records are kept of their

progress via daily graphing and periodic assessment using a computer and the CAAS or some other similar system. After training, students' performance on tests of more complex computational and conceptual/word problems must then be compared to a control group that did not receive training to see if basic arithmetic practice does produce better complex mathematical problem solving.

In addition, the level of analysis needs to be more exact. Several researchers have given explanations of why improving arithmetic abilities improves more complex abilities. The explanations have all been along the lines of "it frees up working memory resources to be applied to more complex aspects of problem solving." That assessment is no longer adequate. We need to dig deeper in our analysis to determine exactly how it is that this resource savings might be working and how it is being applied. We also need to determine if improvement in arithmetic fact retrieval is more beneficial (or only beneficial) on certain types of problems such as those problems requiring calculations versus those problems where calculations are not required.

Working Memory and the Mathematically Disabled

According to many researchers (e.g., Geary, 1994) the systematic study of the mathematically disabled (MD) has lagged well behind the study of reading disabilities even though math disabilities may be as prevalent or more prevalent. An increasing amount of research has been conducted recently on the arithmetically disabled that has

revealed that working memory problems as well as procedural problems (such as counting procedures/strategies) account for much of MD students problems in math. Research has recently started to pick up in the area of problem solving and MD (e.g., Zentall, 1990; Zentall & Ferkis, 1993) as well. Advances have also been made in the study of working memory and MD and disabilities in general as researchers have been able to construct tasks that better capture the dynamic properties of working memory (e.g., Swanson, 1993). Attempts have also been made to categorize learning disabilities by way of working memory analysis (Swanson, 1990). This type of working memory research needs to continue and should at least in part be applied to the study of mathematic disabilities and arithmetic and mathematical problem solving. Some of the future research possibilities outlined in this final section are currently being pursued in our lab and these efforts will continue.

APPENDIX

WORD PROBLEM AND COMPLEX COMPUTATION TEST QUESTIONS

1.) Jill has 5 pieces of candy. Terry gives Jill 7 more pieces. How many does Jill have now?

2.) Paul has 12 comic books. Cindy has 5 comic books. How many more comic books does Paul have than Cindy?

3.) It is Jane's birthday and her brother wants to bake her a cake. He can't find the one cup measure but did find the 1/3 cup measure. How many times must he fill up the 1/3 cup if the cake recipe calls for 2 and 1/3 cups of flour?

4.) On a map of Springfield, 1 inch equals 3 miles. What is the distance between two parks that are 7 inches apart on the map?

5.) Frank has 4 different pairs of pants and 3 different types of shirts. How many different ways could Frank combine his pants and shirts?

6.) Monica has 16 pennies. She has 9 more than Laurie. How many pennies does Laurie have?

7.) Write the number that should replace the question mark in the following number sentence:

$$72 + (3 \times 6) - (12 + 9) = ?$$

8.) Maria rides her bicycle 9 miles every week. Sara rides 7 times as many miles as Maria. How many miles does Sara ride in 3 weeks?

9.) What number should replace the question marks to make the statement below true?

$$\frac{4}{7} = \frac{??}{35}$$

10.) Write a mixed number (a whole number and fraction) that has the same value as the improper fraction 17/6.

11.) 5 + 34 + 111 =

12.)
$$\begin{array}{r} 57 \\ \underline{\times\ \ 64} \end{array}$$

13.) 345 – 15 =

14.) $\frac{4}{9} \times 3 =$

15.) $84 - 10 =$

16.) $\frac{3}{4} \times \frac{8}{15} =$

17.) \$5 - 35¢ =

Tom and Karen work in a toy factory packaging toys. They can fit 40 toys in a large box, 30 toys in a medium box, and 20 toys in a small box. Use this information to answer the next three questions.

18.) A toy store in Amherst would like a small, medium, and large box of toys sent to them on Monday. How many total toys will they be getting?

19.) If Karen gets paid \$2.00 for every 100 toys she boxes and she packs 300 toys every hour, how much money does Karen make in an hour?

20.) Tom wants to finish packing 7 large boxes, 2 medium boxes, and 1 small box before lunch. How long will it take Tom if he can pack 120 toys every half hour?

21.) Kate has 13 pencils. Todd has 5 pencils. How many more pencils does Tom need to have as many as Kate?

22.) Apples are on sale at the market for \$1.06 per pound. A farmer is selling them for 10 cents (\$.10) more per pound. How much would it cost to buy 5 pounds of apples from the farmer?

23.) Ben is a mechanic and can fix 5 sets of brakes in a day. Susan is also a mechanic and can fix 6 sets of brakes in a day. How many sets can they fix together in 5 days?

24.) At Shell, oil costs \$1.25 per quart. At Exxon, oil costs 12 cents (\$.12) less per quart than it costs at Shell. How much would 4 quarts of oil cost at Exxon?

25.) Teddy has 90 raffle tickets to sell for a charity raffle. If he can sell exactly 12 tickets each day, how many days will it take him to sell all of his raffle tickets?

26.) Joseph delivers 25 newspapers every day. Tim delivers 1/5 as many papers as Joseph every day. How many papers does Tim deliver in <u>7 days</u>?

27.) $\begin{array}{r} 3451 \\ +\ \ 274 \\ \hline \end{array}$

28.) $\begin{array}{r} 743 \\ -\ \ 267 \\ \hline \end{array}$

29.) $\begin{array}{r} 714 \\ \times\ \ \ 23 \\ \hline \end{array}$

30.) $\frac{13}{41} + \frac{12}{41} + \frac{21}{41} =$

31.) 10.3 + 3.64 + 12.2 =

32.) 2.24 x 9 =

33.) $\begin{array}{r} .03 \\ \times\ \ .07 \\ \hline \end{array}$

34.) $\frac{3}{4} + \frac{6}{24} =$

35.) What number must replace the **?** to make the following number sentence correct?

(7 x ?) + 2 = 30

36.) What is the average (mean) of the following test scores: 85, 90, and 74?

37.) A local store in John's neighborhood was advertising football cards for sale. John could buy 3 packs for $4.00. Each pack of cards has 12 football cards. How many cards could John buy for $20.00?

38.) John wanted to sell some of his football cards so that he could buy tickets to a football game. John had 4 cards worth $5.00 each, 7 cards worth $2.00 each, and 6 cards worth $1.00 each. How many tickets could John buy if he sold all of these cards and if tickets cost $10.00 each?

39.) John went to a card show with $25.00 to spend. It cost $2.00 for admission and John bought an album to hold his cards in for $5.00. He then bought 2 old football cards with the money he had left over. If one of the football cards cost twice as much as the other, how much was the more expensive card?

40.) One quarter of the students in Mrs. Smith's math class are boys. If there are 15 girls in the class how many students are there total?

41.) What is the greatest common factor of (the largest number that divides into both) 36 and 72?

42.) If two boats 100 miles apart sail towards each other at the same time, how long will it take for them to meet if one boat travels 20 miles per hour and the other travels 30 miles per hour?

43.) Filene's is having a sale on women's clothing. Every item of women's clothing is 20% off the original price. Sarah buys a dress at Filene's on sale for $56.00. What was the original price of her dress before the sale?

44.) $$\begin{array}{r} 40000 \\ \underline{\times \quad 127} \end{array}$$

45.) 37% of 90 =

46.) $360 \div .8 =$

47.) $\frac{23}{27} - \frac{4}{9} =$

48.) $\frac{5}{3} \div \frac{25}{6} =$

49.) $7\frac{10}{24} - 3\frac{11}{24} =$

50.) $\frac{1}{3} + \frac{1}{6} + \frac{5}{8} =$

References

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Antell, S. E. & Keating, D. P. (1983). Perception of numerical invariance in neonates. Child Development, 54, 695-701.

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), Formal methods in developmental psychology: Progress in cognitive development research, (pp. 302-338). New York: Springer-Verlag.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. Cognition, 44, 75-106.

Ashcraft, M. H. & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. Journal of Experimental Psychology: Human Learning and Memory, 4, 527-538.

Ashcraft, M. H., Donley, R. D., Halas, M. A., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J. I. D. Campbell (Ed.) The nature and origins of mathematical skills, (pp. 301-329). Amsterdam: Elsevier.

Ashcraft, M. H. & Fierman, B. A. (1982). Mental addition in third, fourth, and sixth graders. Journal of Experimental Child Psychology, 33, 216-234.

Ashcraft, M. H. & Stazyk, E. H. (1981). Mental addition: A test of three verification models. Memory & Cognition, 9, 185-96.

Balow, I. H. (1964). Reading and computation ability as determinants of problem solving. The Arithmetic Teacher, 11, 18-22.

Baroody, A. J. (1984a). Children's difficulties in subtraction: Some causes and questions. Journal for Research in Mathematics Education, 15, 203-213.

Baroody, A. J. (1984b). The case of Felicia: A young child's strategies for reducing memory demands during mental addition. Cognition and Instruction, 1, 109-116.

Baroody, A. J. (1987a). Children's Mathematical Thinking. New York: Teacher's College Press.

Baroody, A. J. (1987b). The development of counting strategies for single-digit addition. Journal for Research in Mathematics Education, 20, 159-172.

Baroody, A. J. & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. Cognition and Instruction, 1, 321-339.

Baroody, A. J. & Ginsburg, H. P. (1983). The effects of instruction on children's concept of "equals": Elementary School Journal, 84, 199-212.

Baroody, A. J. & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics. Hillsdale, NJ: Erlbaum.

Becker, J. (1993). Young children's numerical use of number words: Counting in many-to-one situations. Developmental Psychology, 29, 458-465.

Bullock, M., & Gelman, R. (1977). Numerical reasoning in young children: The ordering principle. Child Development, 48, 427-434.

Campbell, J. I. D. (1985). Associative interference in mental computation. PhD. Thesis, University of Waterloo, Ontario, Canada.

Campbell, J. I. D. (1987a). The role of associative interference in learning and retrieving arithmetic facts. In J. A. Sloboda & D. Rogers (Eds.), Cognitive processes in mathematics, (pp. 107-122). Oxford: Oxford Press.

Campbell, J. I. D. (1987b). Network interference and mental multiplication. Journal of Experimental psychology: Learning, Memory, and Cognition, 13, 109-123.

Campbell, J. I. D. (1990). Retrieval inhibition and interference in cognitive arithmetic. Canadian Journal of Psychology, 44, 445-464.

Campbell, J. I. D. (1991). Conditions of error priming in number-fact retrieval. Memory & Cognition, 19, 197-20.

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process and acquisition. Canadian Journal of Psychology, 39, 338-366.

Carpenter, T. P., Hiebert, J. & Moser, J. M. (1981). Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems. Journal for Research in Mathematics Education, 12, 27-39.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective, (pp. 9-24). Hillsdale, NJ: Erlbaum.

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

Compton, B. J., & Logan, G. D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. Memory & Cognition, 19, 151-158.

De Corte, E., & Verschafell, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. Journal for Research in Mathematics Education, 18, 363-381.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. Cognition, 43, 1-29.

Ellis, N. C. (1992). Linguistic relativity revisited: The bilingual word-length effect in working memory during counting, remembering numbers, and mental calculation. In R. J. Harris (Ed.), Cognitive processing in bilinguals. Amsterdam: Elsevier.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. British Journal of Psychology, 71, 43-51.

Fodor, J. (1983). Modularity of mind. Cambridge: MIT Press.

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs (No. 5).

Fuson, K. C. (1982). An analysis of the counting-on solution procedure in addition. In T.P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective, (pp. 67-81). Hillsdale, NJ: Lawrence Erlbaum.

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

Fuson, K. C. (1992). Research on Whole Number Addition and Subtraction. In D. A. Grouws (Ed.), Handbook of Research on Mathematics Teaching and Learning, (pp. 243-275). New York: Macmillan Publishing Company.

Fuson, K. C., & Kwon, Y. (1991). Chinese-based regular and European irregular systems of number words: The disadvantages for English-speaking children. In K. Durkin & B. Shire (Eds.), Language in mathematical education: Research and practice, (pp. 211-226). Milton Keynes, PA: Open University Press.

Fuson, K. C., & Kwon, Y. (1992a). Korean children's single-digit addition and subtraction: Numbers structured by ten. Journal for Research in Mathematics Education, 23, 148-165.

Fuson, K. C., & Kwon, Y. (1992b). Korean children's understanding of multidigit addition and subtraction. Child Development, 63, 491-506.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. Cognition, 44, 43-74.

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. Journal of Experimental Child Psychology, 49, 363-383.

Geary, D. C. (1994). Children's mathematical development. Washington, D. C.: American Psychological Association.

Geary, D. C. (1996). The problem size effect in mental addition: Development of cross national trends. Mathematical Cognition, 2, 63-93.

Geary, D. C., Bow-Thomas, C. C., Fan, L., & Siegler, R. S. (1993). Even before formal instruction, Chinese children outperform American children in mental addition. Cognitive Development, 8, 517-529.

Geary, D. C. & Brown, S. C. (1991). Cognitive Addition: Strategy Choice and Speed-of-Processing Differences in Gifted, Normal, and Mathematically Disabled Children. Developmental Psychology, 27, 398-406.

Geary, D. C., Brown, S. C. & Samaranayake, V. A. (1991). Cognitive Addition: A Short Longitudinal Study of Strategy Choice and Speed-of-Processing Differences in Normal and Mathematically Disabled Children. Developmental Psychology, 27, 787-797.

Geary, D. C., & Burlingham-Dubree, M. (1989). External validation of the strategy choice model for addition. Journal of Experimental Child Psychology, 47, 175-192.

Geary, D. C., Fan, L., & Bow-Thomas, C. C. (1992). Numerical cognition: Loci of ability differences comparing children from China and the United States. Psychological Science, 3, 180-185.

Geary, D. C., Salthouse, T. A., Chen, G. P., & Fan, Liu (1996). Are east Asian versus American differences in arithmetical ability a recent phenomenon? Developmental Psychology, 32, 254-262.

Geary, D. C. & Widaman, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. Intelligence, 16, 47 -80.

Geary, D. C., Widaman, K. F., & Little, T. D. (1986). Cognitive addition and multiplication: Evidence for a single memory network. Memory & Cognition, 14, 478-87.

Geary, D. C., & Wiley, J. G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young and elderly adults. Psychology and Aging, 6, 474-483.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. Cognitive Science, 14, 79-106.

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

Gillund, G. & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-67

Goldman, S. R., Mertz, D. L., & Pellegrino, J. W. (1989). Individual difference in extended practice functions and solution strategies for basic addition facts. Journal of Educational Psychology, 81, 481-496.

Goldman, S. R., Pellegrino, J. W., & Mertz, D. L. (1988). Extended practice of basic addition facts: Strategy changes in learning disabled students. Cognition and Instruction, 5, 223-265.

Graham, D. J. (1987). An associative retrieval model of arithmetic memory: how children learn to multiply. In J. A. Sloboda & D. Rogers (Eds.), Cognitive processes in mathematics, (pp.123-141). Oxford: Oxford Press.

Graham, D. J. & Campbell, J. I. D. (1992). Network Interference and Number-Fact Retrieval: Evidence From Children's Alphaplication. Canadian Journal of Psychology, 46, 65-91.

Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.) The development of mathematical thinking, (pp. 153-196). Orlando, FL: Academic Press.

Groen, G. J. & Parkman, J. M. (1972). A chronometric analysis of simple addition. Psychological Review, 79, 329-343.

Hamann, M. S. & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. Cognition and Instruction, 3, 173-192.

Hatano, G., Miyake, Y., & Binks, M. G. (1977). Performance of expert abacus operators. Cognition, 5, 47-55.

Hiebert, J. (1990). The role of routine procedures in the development of mathematical competence. In T. J. Cooney & C. R. Hirsch (Eds.), Teaching and Learning Mathematics in the 90's,(pp. 31-39). Reston, VA: The National Council of Teachers of Mathematics, Inc.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. Cognitive Psychology, 10, 302-323.

Ilg, F., & Ames, L. B. (1951). Developmental trends in arithmetic. Journal of Genetic Psychology, 79, 3-28.

Jensen, A. R., & Whang, P. A. (1994). Speed of accessing arithmetic facts in long-term memory: A comparison of Chinese-American children. Contemporary Educational Psychology, 19, 1-12.

Jerman, M. E., & Mirman, S. (1974). Linguistic and computational variables in problem solving in elementary mathematics. Educational Studies in Mathematics, 5, 317-362.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. Psychological Bulletin, 109, 490-501.

Kaye, D. B. (1986). The development of mathematical cognition. Cognitive Development, 1, 157-170.

Kaye, D. B., Post, T. A., Hall, V. C., & Dineen, J. T. (1986). Emergence of information-retrieval strategies in numerical cognition: A developmental study. Cognition and Instruction, 3, 127-150.

Kaye, D. B., De Winstanley, P., Chen, Q., & Bonnefil, V. (1989). Development of efficient arithmetic computation. Journal of Educational Psychology, 81, 467-480.

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

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 194-209.

Koshmider, J. W. III & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. Journal of Experimental Child Psychology, 51, 53-89.

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lefevre, J. A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. Journal of Experimental Psychology: General, 125, 284-306.

LeFevre, J. A., Bisanz, J. & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. Memory & Cognition, 16, 45-53.

LeFevre, J. A., & Kulak, A. G. (1994). Individual differences in the obligatory activation of addition facts. Memory & Cognition, 22, 188-200.

Lefevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: reassessing problem size effects in adults. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 216-230.

Lemaire, P., Abdi, H., & Fayol, M. (1996). The role of working memory resources in simple cognitive arithmetic. European Journal of Cognitive Psychology, 8, 73-103.

Lemaire, P., Barrett, S. E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. Journal of Experimental Child Psychology, 57, 224-258.

Lewis, A. B. (1989). Training students to represent arithmetic word problems. Journal of Educational Psychology, 81, 521-531.

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

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. Cognitive Psychology, 12, 523-553.

Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95, 492-527.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. Memory and Cognition, 22, 395-410.

Mayer, R. E. (1985). Mathematical ability. In R. J. Sternberg (Ed.), Human abilities: An information processing approach (pp. 127-150). San Francisco: W. H. Freeman.

Mayer, R. E., Lewis, A. B., & Hegarty, M. (1992). Mathematical misunderstandings: Qualitative reasoning about quantitative problems. In J. I. D. Campbell (Ed.), The nature and origins of mathematical skills, (pp. 137-153). Amsterdam: Elsevier.

McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 377-397.

Mercer, C. D. (1979). Children and adolescents with learning disabilities. Columbus, OH: Charles E. Merrill.

Meyer, J. L., & Well, A. D. (1995). Research design and statistical analysis. Hillsdale, NJ: Erlbaum.

Miller, K., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 46-60.

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

Muth, K. D. (1984). Solving arithmetic word problems: Role of reading and computational skills. Journal of Educational Psychology, 76, 205-210.

Nesher, P., Greeno, J. G., & Riley, M. S. (1982). The development of semantic categories for addition and subtraction. Educational Studies in Mathematics, 13, 373-394.

Paap, K. R., & Ogden, W. C. (1981). Letter encoding is an obligatory but capacity-demanding operation. Journal of Experimental Psychology: Human Perception and Performance, 7, 518-527.

Pawlik, K. (1966). Concepts in human cognition and aptitudes. In R. B. Cattell (Ed.), Handbook of multivariate experimental psychology (pp. 535-562). Chicago: Rand McNally.

Perfetti, C. A., & Hogaboam, T. (1975). The relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 67, 461-469.

Piaget, J. (1965). The child's conception of number. New York: Norton.

Piaget, J. (1970). Psychology and epistemology. New York: W. W. Norton.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), Information processing and cognition: The Loyola symposium, (pp.55-85). Hillsdale, NJ: Erlbaum.

Rabinowitz, M. & Woolley, K. E. (1995). Much ado about nothing: The relation among computational skill, arithmetic word problem comprehension, and limited attentional resources. Cognition and Instruction, 13, 51-71.

Regan, J. E. (1981). Automaticity and learning: Effects of familiarity on naming letters. Journal of Experimental Psychology: Human Perception and Performance, 7, 180-195.

Resnick, L. B. & Ford, W. W. (1981). The psychology of mathematics for instruction. Hillsdale, NJ: Erlbaum.

Rogers, W. A., & Fisk, A. D. (1991). Age-related differences in the maintenance and modification of automatic processes: Arithmetic Stroop interference. Human Factors, 33, 45-56.

Royer, J. M. (1990). The Sentence Verification Technique: A new direction in the assessment of reading comprehension. In Legg, S. M., and Algina, J. (Eds.), Cognitive assessment of language and math outcomes. Norwood, NJ: Ablex.

Royer, J. M., Carlo, M. S., & Cisero, C. A. (1992). School-bases uses for the Sentence Verification Technique of measuring listening and reading comprehension. Psychological Test Bulletin, 5, 5-19.

Royer, J. M., & Sinatra, G. M. (1994). A cognitive theoretical approach to reading diagnostics. Educational Psychology Review, 6, 81-113.

Royer, J. M., & Tronsky, L. N. (in press). Addition practice with math disabled students improves subtraction and multiplication performance. In T. E. Scruggs and M. A. Mastropieri (Eds.), Advances in Learning and Behavioral Disabilities (Vol. 12). Greenwich, CT: JAI Press.

Siegler, R. S. (1986). Unities across domains in children's strategy choices. In M. Perlmutter (Ed.), Perspectives for intellectual development: The Minnesota Symposia on Child Psychology, (Vol. 19, pp. 1-48). Hillsdale, NJ: Erlbaum.

Siegler, R. S. (1987). Strategy choices in subtraction. In J. Sloboda & D. Rogers (eds.), Cognitive processes in mathematics, (81-106). Oxford: Oxford Press.

Siegler, R. S. (1988a). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. Child Development, 59, 833-851.

Siegler, R. S. (1988b). Strategy choice procedures and the development of multiplication skill. Journal of Experimental Psychology: General, 117, 258-275.

Siegler, R. S. (1989). Hazards of mental chronometry: An example from children's subtraction. Journal of Educational Psychology, 81, 497-506.

Siegler, R. S. & Jenkins, E. A. (1989). How children discover new strategies. Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds.), Developing cognitive competence: New approaches to process modeling. Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). Strategy choice in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), Origins of cognitive skills, (pp. 229-293). Hillsdale, NJ: Erlbaum.

Silver, E. A. (1987). Foundations of cognitive theory and research for mathematics problem-solving. In A. H. Schoenfeld (Ed.), Cognitive Science and Mathematics Education, (pp. 33-60). Hillsdale, NJ: Erlbaum.

Spearman, C. (1927). The abilities of man. London: Macmillan.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-407.

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, auotomaticity, and modularity. Developmental Review, 10, 72-100.

Starkey, P., & Cooper, R. G., Jr. (1980). Perception of numbers by human infants. Science, 210, 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. Science, 222, 179-181.

Stazyk, E. H., Ashcraft, M. H., & Hamann, M. S. (1982). A network approach to mental multiplication. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 320-335.

Steinberg, R. M. (1985). Instruction on derived facts strategies in addition and subtraction. Journal for Research in Mathematics Education, 16, 337-355.

Stevenson, H. W., Chen, C., & Lee, S. Y. (1993). Mathematics achievement of Chinese, Japanese, and American Children: Ten years later. Science, 259, 53-58.

Stigler, J. W., Fuson, K. C., Ham, M., & Kim, S. M. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. Cognition and Instruction, 3, 153-171.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. Journal of Experimental Psychology, 18, 643-662.

Suppes, P., Jerman, M., & Brian, D. (1968). Computer-assisted instruction: Stanford's 1965-66 arithmetic program. New York: Academic Press.

Suppes, P., & Morningstar, M. (1972). Computer-assisted instruction at Stanford, 1966-1968: Data, models, and evaluation of the arithmetic programs. New York: Academic Press.

Swanson, H. L. (1993). Working memory in learning disability subgroups. Journal of Experimental Child Psychology, 56, 87-114.

Swanson, H. L., Cochran, K. F., & Ewers, C. A. (1990). Can learning disabilities be determined from working memory performance? Journal of Learning Disabilities, 23, 59-67.

Widaman, K. F., & Little, T. D. (1992). The development of skill in mental arithmetic: An individual differences perspective. In J. I. D. Campbell (Ed.), The nature and origins of mathematical skills, (pp. 189-253). Amsterdam: North-Holland.

Woods, S. S., Resnick, L. B. & Groen, G. J. (1975). An experimental test of five process models for subtraction. Journal of Educational Psychology, 67, 17-21.

Wynn, K. (1990). Children's understanding of counting. Cognition, 36, 153-193.

Wynn, K. (1992a). Addition and Subtraction by human infants. Nature, 358, 749-750.

Wynn, K. (1992b). Children's acquisition of the number words and the counting system. Cognitive Psychology, 24, 220-251.

Zbrodoff, N. J. & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. Journal of Experimental Psychology: General, 115, 118-130.

Zbrodoff, N. J. & Logan, G. D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 83-97.

Zentall, S. S. (1990). Fact-retrieval automatization and math problem solving by learning disabled, attention disordered, and normal adolescents. Journal of Educational Psychology, 82, 856-865.

Zentall, S. S., & Ferkis, M. A. (1993). Mathematical problem solving for youth with ADHD, with and without learning disabilities. Learning Disability Quarterly, 16, 6-18.