1-1-1985

# Categorization of fraction word problems.

Pamela T. Hardiman
*University of Massachusetts Amherst*

# CATEGORIZATION OF FRACTION WORD PROBLEMS

A Dissertation Presented

By

PAMELA THIBODEAU HARDIMAN

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

DOCTOR OF PHILOSOPHY

September 1985

Psychology

CATEGORIZATION OF FRACTION WORD PROBLEMS

A Dissertation Presented

By

PAMELA THIBODEAU HARDIMAN

Approved as to style and content by:

Arnold Well

Arnold Well, Chairperson of Committee

George W. Cobb

George Cobb, Member

Carolyn B Mervis

Carolyn Mervis, Member

Alexander Pollatsek

Alexander Pollatsek, Member

Melinda A. Novak
Acting Chair

Seymour Berger, Department Head

Psychology

Dedicated to all students of mathematics

# ACKNOWLEDGMENTS

The research was conducted with the most appreciated assistance of Arnie Well for Experiment I and Leslie Arriola for Experiment II. Leslie Arriola also helped to enter the data for Experiment II.

Many thanks to the members of my committee who aided in the development of the ideas presented in this dissertation. Carolyn Mervis has been supportive of this research from the earliest stages; her enthusiasm gave me confidence in the logic of my ideas. George Cobb provided the perfect model of the reader; just distant enough to the research to remind me that I'd better write things so they will be clear to George. Alexander Pollatsek has been a second advisor to me: the ideas for Experiment I were aided by discussions with Sandy during the fall of 1983. The thought and writing of this dissertation were greatly aided by the privacy of Sandy's office, whose use I was entrusted with during his sabbatical leave.

I would like to express my great appreciation to my advisor, Arnold Well. I have often appreciated the freedom I was given to develop my own ideas, which often may have seemed unusual. Arnie has always been supportive and willing to discuss new articles and research ideas. He has also been an ever patient reader of my writing.

No graduate student ever survives without a full

support staff, including spouse, relatives, friends, and dog. I have been lucky to have had the companionship of Cindy Connine, Susan Lima, and Albrecht Inhoff as officemates and friends through the years. My family has provided me with the models and inspiration to be a well-rounded person, and have waited patiently for many years as graduate school precluded for a time their opportunity to become grandparents. My husband, Chris, has been a constant source of encouragement through my higher education. Having survived six years of marriage and graduate school together, we should be ready for anything. Last but not least, I want to give a rub on the belly to my forever patient, kind and faithful dog, Dirt.

ABSTRACT

CATEGORIZATION OF FRACTION WORD PROBLEMS

September 1985

Pamela Thibodeau Hardiman, B.A., Hampshire College
M.S., University of Massachusetts, Ph.D., University of Massachusetts

Directed by : Professor Arnold Well

The present research was designed to: 1) determine whether categorization of fraction word problems can be explained by a theory of categorization developed to predict natural object categorization, the Best Examples theory (Mervis and Rosch, 1980), and 2) to determine whether this approach provides any new advantages.

In Experiment I, the effects of similarity of story-line and operation on categorization by nonexperts and experts were investigated. Subjects saw items with four alternatives that matched a standard in story-line and operation, operation alone, story-line alone, and neither dimension, and were to choose the two that required the same operation as the standard. On eight additional items, they indicated the operation needed to solve a word problem. The results supported a Best Examples interpretation: 1) most experts and nonexperts chose the story-line and operation match, which had the highest cue overlap

with the standard, however 2) for their other choice, nonexperts often chose the alternative matching in story-line alone. Experiment I also indicated fraction multiplication word problems are hard to identify, and that there are systematic differences in the difficulty of identifying problems that require the same operation.

Experiment II was designed explore these problem differences and determine whether the results were replicable with a younger population, eighth graders. A third level of problem structure was proposed to account for problem differences, termed middle level structure. Problems with the same middle level structure have the same algebraic open sentence description, match in presence or absence of action cues and are solved in similar ways.

Subjects were given three tasks: 1) middle strcture match task, requiring the choice of one of four alternatives that used the same operation as a standard, 2) identification of operations needed to solve whole number and fraction problems, and 3) a modification of the task in Experiment I. The results suggested that: 1) matching middle level structure aids catgeorization, and may be a more basic level of categorization than operation,,and 2) the difficulty of categorizing fraction problems is not related to difficulty categorizing corresponding whole number problems. Instructional materials should make finer distinctions among problems.

TABLE OF CONTENTS

LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

# OVERVIEW

To many people, mathematics is a topic to be avoided when possible, particularly if fractions are involved. Yet, proficiency in mathematics is becoming increasingly necessary in today's technology-oriented world. Thus, investigation of the acquisition of problem solving skills critically important.

The present research was designed with two intentions: 1) to determine whether categorization of fraction word problems can be explained by a theory of categorization originally developed to predict natural object categorization, the Best Examples theory (Mervis and Rosch, 1980), thus extending the scope of that theory, and 2) to determine whether there is any advantage to using a Best Examples approach, i.e., does it allow one to gain any new information about how students solve fraction word problems.

## What is an expert?

How does one define expertise within a domain of knowledge? One characteristic of expert problem solvers is their ability to "attend to the important features of a problem and then select the action which will lead to the solution of that problem" (Lewis and Anderson, 1985, pg. 26). In particular, organization of knowledge and the ability to categorize problems and integrate new information have been identified as distinguishing characteristics of experts in complex domains, such as physics and mathematics. However, if one considers that people can be experts in areas other than those which require what is generally thought of as problem solving, then this characterization seems overly limited in its generality. For example, most adults are reasonably expert speakers, readers, and negotiators of the physical world, but these domains are rarely construed as problem solving domains. Adults rarely trip when negotiating the physical world, whereas toddlers who are less expert negotiators of the physical world frequently fall.

The domains in which people may be said to be expert form a continuum, from those fields in which relatively few people are expert, such as physics and computer science, to fields such as arithmetic in which many people are expert, to the categorization and naming of everyday objects, in which most people are reasonably expert. Recently, fairly detailed characterizations have been made of experts in technical fields such as chess, physics, and computer science (Adelson, 1981; Chi, Feltovich, and Glaser, 1981; Larkin, McDermott, Simon, and Simon, 1980; McKeithen, Reitman, Reuter, and Hirtle, 1981). The general findings of these studies are: 1) novices and experts organize concepts differently; novices rely more heavily on surface features, while experts organize around deep structural cues, such as general physics principles, and 2) novices and experts use different strategies to solve problems; when attempting to solve physics problems, novices primarily use means-ends analysis, while experts try to proceed forward and develop the information given.

At the opposite end of the spectrum of domains, interest in differences in child versus adult labeling of objects has generated a considerable body of theory and data (Bowerman, 1980; Clark, 1973; Mervis, 1980). A major indicator of how quickly and accurately an object can be identified as a member of a particular category is the typicality of the object, i.e. how representative it is of the category. Typical members of a category share many features with other members of the category and few features with non-category members. Children who are able to sort objects taxonomically categorize the typical members of a category in the same way an adult would. Discrepancies between child and adult labeling tend to occur with the poorer examples of a category, which may share many features with neighboring categories.

These characterizations of expert and novice problem solvers and adult and child object categorizers are somewhat different in nature. There are several reasons why they may be different, including: 1) the difference in ages of the subjects, 2) the likelihood of becoming an expert in the domain, and 3) the fact that researchers have tended to view these areas as distinctly different areas of research. However, one might argue that these differences are related mainly to the subject matter, making it possible to develop a definition of expertise that encompasses a broader range of domains. This would facilitate discussion of expertise in the middle range of domains, such as the domain of fraction word problems, the topic of concern for this paper.

One step toward defining expertise more broadly might be to expand upon the finding that categorization is an important component of problem solving. First, as has been noted, a substantial literature on categorization of objects already exists. This literature has emphasized the ways in which experts and novices are similar and how differences between experts and novices are resolved. Therefore, this orientation seems amenable to speculation about how a novice problem solver might become an expert (for example, see Mervis and Crisafi, 1982 or Mervis and Mervis, 1982). A definition of expert object labeling may be considered as follows: A category is said to have been formed when "two or more distinguishable objects or events are treated equivalently" (Mervis and Rosch, 1980). A child may be considered to have developed an adult understanding of a category when the child applies the category label to all objects an adult would apply the label to and no others, i.e. neither under- nor over-extending the label. Using this definiton, if "objects and events" were supplemented with the more abstract notion of "problems," and the words "child" and "adult" were replaced by 'novice' and 'expert', the result would be consistent with Lewis and Anderson's (1985) characterization of expertise, but would be somewhat more comprehensive: an expert problem solver is one who is able to correctly categorize problems according to the action which should be taken to solve the problem. Although there are certainly other skills associated with expertise in various domains, the ability to categorize information appropriately will be the main characteristic of concern.

## The Present Research

The hypothesis guiding the present research is that learning about abstract categories, such as the conditions where arithmetic operators may be successfully applied, involves essentially the same processes as learning when object category names apply. In both cases, people must learn what conditions indicate membership in a certain category. In general, the major tasks of the novice are to learn which cues are correlated with membership in a category, and to learn to distinguish arbitrary and non-arbitrary cues.

The domain chosen for this study was one-step arithmetic problems involving proper fractions, i.e. word problems which present two quantities and information about the operation which is required, and ask for a third, unknown quantity. This domain is relatively constrained, in that there are nominally only four possible categories: the problem may require the operation of either addition, subtraction, multiplication, or division. These categories are considerably more abstract than those which have been studied previously, presenting an opportunity to determine whether the usefulness of the Best Examples theory may be expanded to a broader range of categories.

The study of the domain of fraction word problems has considerable educational relevance as well; National Assessment of Educational Progress (Carpenter, Corbitt, Kepner, Lindquist, and Reys, 1980) results indicate that fractions and operations with fractions are not well understood, particularly in the context of word problems. Yet, an understanding of fractions is critical to higher levels of mathematics, such as algebra and calculus. Obviously, a desirable educational goal is to make the acquisition of expertise in the domain of fractions both more probable and more efficient. Thus, the domain of fraction word problems is important because it invovles a small and constrained set of abstract concepts, enabling a test of the relevance of theories based on object category formation as a general definition of expertise, and the topic is of educational concern.

A central goal of both studies was to learn what cues subjects use when they must decide that two problems are similar, and whether experts and novices rely on the same kinds of cues in making this decision. For this purpose, a variation on the oddity task was used. The oddity task has been used by Rosch and her associates in several studies on object categorization (see Rosch and Mervis, 1977; Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976; Mervis and Crisafi, 1982). In this research, the subject is presented with three pictures of objects and is asked which picture does not belong with the other pictures. Two of the pictures belong to the same category, while the third picture is related in some way to one or both of the other two pictures, but is not a member of the category. For example, in order to determine whether a child was able to categorize objects as animals, one might present a dog, a horse, and a house. A young child might incorrectly group the dog and the house because dogs are typically seen in houses, while horses are not.

In the present studies, subjects saw a fraction word problem and had to indicate which two of four alternative problems required the same operation. The alternatives were structured so that one matched in both story context and operation, a second matched in operation alone, a third matched in story context alone, and the fourth matched in neither dimension. According to the Best Examples theory, both nonexperts and experts should choose the alternative that matches in both story context and operation, since it has a high cue overlap with the standard. For their second choice, experts should choose the alternative that matches the standard in operation, whereas novices may choose the alternative that matches only in surface structure because of the arbitrary surface similarity. This task will be referred to as the similarity judgment task.

Experiment I

Experiment I was intended to address three major issues, using a college age population: 1) Do experts and nonexperts use the same cues to decide whether two problems require the same operation for solution? 2) Is cue use related to the operation required for solution? and 3) Does the form of the presentation of a problem influence the ease of categorizing problems according to operation for solution, i.e. word problems versus numerical expressions? On the basis of NAEP (Carpenter et.al., 1980) results, a difference due to operation was expected, with addition problems being easier to identify as similar than multiplication problems. No predictions could be made about division problems, since none were given on the NAEP. Research by Lesh, Landau, and Hamilton (1983) suggested word problems paired with numerical alternatives should be easier to match than other presentation forms.

There were also two additional issues: 1) Does the complexity of the fractions involved in a problem, and hence the difficulty of the computation, influence ability to categorize problems? and 2) Does the general form of the problem statement influence ability to categorize problems? Intuitively, it seems that word problems with less complex numbers should be easier to recognize as similar. Concerning the second issue, research with young children solving addition and subtraction problems indicates there are differences in the ease of solution of a problem that are related to the type of problem statement (Briars and Larkin, 1984; Carpenter and Moser, 1983; Riley, Greeno, and Heller, 1983). As it was not feasible to investigate all possible types of fraction problems, two basic forms of problems were chosen to be investigated in this study: 1) dynamic problems which portray a transfer of an item, making the end state different from the start state, and 2) static problems which concern the relationship between two quantities, with no change in the start state.

In addition to the similarity judgment task, the subjects were given two tasks to assess their level of mastery of fractions problems and verbal skills, the operations assessment task and the Peabody Picture Vocabulary Test. A set of fraction word problems were given on the operations assessment task, and the subject was to state what operation should be used to solve each problem. The PPVT was given to ensure that any differences between relatively expert and nonexpert subjects were not simply due to different levels of verbal skills.

The results of Experiment I support a Best Examples interpretation: 1) both experts and nonexperts frequently chose the alternative which matched the standard in both story line and operation, i.e. had a high degree of cue overlap, and 2) for their other choice, most experts correctly chose the alternative matching in operation alone, whereas nonexperts often chose the story line only match. Thus, nonexperts were often mislead by cue overlap.

In addition, the results indicated there were differences among problems requiring different operations; addition problems yielded the highest rate of correct responses, while multiplication yielded the lowest rate. There were also differences within operations between dynamic and static problems. Performance in classifying dynamic problems was generally somewhat better, except in the case of multiplication. There were striking differences in number of correct responses for mutiplication versus all other operations, and in the types of responses to dynamic and static problems. Over sixty percent of the subjects misclassified the dynamic multiplication problem as a subtraction problem, whereas approximately half the subjects misclassified the static multiplication problem as a division problem. This result suggests that subjects may not view these two types of problems as members of the same category.

## Experiment II

Experiment I suggested that there may be systematic differences in approaches to word problems that require the same operation, particularly in the case of multiplication. Therefore, Experiment II was designed to investigate within- and between-operation differences more systematically. A second aim was to try to extend the results to a younger population to ensure that the results were not peculiar to adult nonexperts.

More specifically, there were three issues concerning within- and between-operations differences. The first issue concerned when subjects would be likely to choose the alternative that matched in story line rather than operation. Experiment I indicated that nonexperts often chose incorrect alternatives for problems structured so that the story line match appeared highly confusable with the standard. Would subjects also err if the story line match required an operation which subjects did not tend to confuse with the operation required by the standard? To address these questions, the similarity judgment task was again used, with the change that the distractor items were paired so that every operation appeared with every other operation. The operations assessment task was used to predict when subjects would err.

Second, Experiment I provided evidence that the presence of an action cue affects problem categorization: dynamic and static multiplication problems were not perceived as the same type of problem by most subjects. In order to determine how important the dynamic-static dimension is, word problems were paired with sets of four alternatives that shared cover stories and were all either dynamic or static word problems. The sets of alternatives were paired with given problems such that the dynamic-static dimension either matched or did not match the given problem. The subject was instructed to indicate the alternative that required the same operation as the given problem.

Third, there was a concern that the difficulties subjects had in determining correct operations for fraction problems may have merely reflected difficulty in dealing with such problems in general. Therefore, a set of whole number problems was added to the operations assessment task to determine whether subjects also had difficulty categorizing whole number problems.

The tasks were presented in the following order: 1) the task assessing the importance of the action cue, 2) the operations assessment task, 3) the Peabody Picture Vocabulary Task, and 4) the similarity judgment task.

The results of Experiment II suggest action cue is an important dimension of problems: both age groups were more accurate in choosing the alternative that matched in operation when the problems matched in presence of an action cue. This result, together with a logical analysis of problem types, suggest that all word problems which require the same operation may not perceived as members of the same category. There appear to be critical differences among subtypes of fraction word problems. The advantage gained through matching middle level structure suggests the middle level may be a more basic level of categorization than the level of operation. Hence, teaching might be more efficient by exploiting this middle level, rather than treating operation as basic.

The results of the operations assessment task rule out the possibility that subjects fail to solve fraction word problems because of difficulties understanding similar problems with whole numbers. Thus, there are difficulties uniquely related to fraction word problems, suggesting the basic level also includes distinctions based on the type of numbers involved in the problem.

Finally, the results of the similarity judment task suggest that for at least younger subjects, there is a relationship between choice of an incorrect alternative and confusability of operations. Subjects tended to choose the surface structure distractor when the operation of that alternative matched their incorrect response made in the fractions assessment task. In Section II, both age groups were more accurate in choosing the alternative that matched in operation when the dynamic-static dimension matched, suggesting that this dimension is an important feature of fraction problems. On the fractions assessment task, nearly all subjects responded with the correct operations for whole number problems, eliminating the argument that the operation was not generally well understood.

In conclusion, the results of the present experiments support a Best Examples interpretation, thus extending the range of categories to which this theory may be applied. In addition, the idea of levels of categorization, and of a basic level, has proven to be a useful concept in research and may change the way children are taught to solve problems as well.

# CHAPTER II

# DEFINING EXPERTISE

## Requirements for a Definition

The goal of the present research is to develop a framework for investigating skilled behavior in a range of natural domains, including fraction word problems. Ideally, a framework for investigating expertise should: 1) be applicable to a broad range of ages and domains, 2) allow levels of expertise to be distinguished, and 3) result in a theory of how a novice might become an expert. These properties will be examined in more detail in this section, followed by a discussion of the literatures on experts and novices and on object categorization, and a proposal for a research framework.

According to Webster's Third New International Dictionary (1976), an expert is one who has "special skill or knowledge derived from training or experience." Two assumptions seem to be implied in this definition: 1) the average person does not have this special skill or knowledge, and 2) expertise is acquired over a relatively long period of time. In contrast, a novice is "one who has no previous training or experience in a specific field or activity."

The literature concerning experts and novices has generally accepted these definitions, with the minor qualification that novices have had some short period of training or experience in the domain such that they can perform the operations, but not skillfully. Complex domains that require a considerable body of specialized knowledge, such as chess (Chase and Simon, 1973), Go (Reitman, 1976), physics (Chi, Feltovich, and Glaser, 1981; Larkin, McDermott, Simon, and Simon, 1980), and computer programming (Adelson, 1981; McKeithen, Reitman, Rueter, and Hirtle, 1981), have commonly been chosen to investigate differences between novices and experts. However, most human beings have acquired skill in a considerable range of activities involved in their daily lives. Although they may not be specialized in the sense that only a few people are capable of such behavior, such activities as verbal communication, walking, and driving are skills, since they require a long period of training, as can be attested to by any parent. If expertise were defined in a way that encompassed a broader range of human activity, then not only would the results of any investigation potentially be more widely applicable, but a fruitful interaction might result from considering literatures of both specialized and less specialized abilities.

The definition of expertise also should be capable of generating a theory of how a novice might, or might not, become an expert. The expert-novice orientation has focused primarily on identifying patterns of behavior associated with the ends of a spectrum of abilities, with little concern as to how progression along the dimension might occur. However, the role of learning in cognition recently has begun to generate interest among cognitive psychologists interested in problem solving (see Anderson, 1982; Langley and Simon, 1981; Lewis and Anderson, 1985).

There appear to be two reasons why learning seems to have assumed a somewhat more central position in psychological research. The first is relevant to psychological theory: the human mind is an inhererently adaptive system which is constantly changing, and therefore knowledge and strategies can never truly be fixed. Simon (Anzai and Simon, 1979; Langley and Simon, 1981) has proposed that if an invariant of human behavior existed, then a theory of learning might supply the necessary invariants. Since expertise can only result from a continuing process of learning, any general framework for investigating expertise must incorporate the notion of dynamic change in a system. The second is a more practical concern: if we know more about the process of learning, then it should be possible to design educational programs that lead to more efficient acquisition of skills taught by explicit instruction.

Once a framework has been developed that incorporates the notion of dynamic change, the capacity to distinguish levels of expertise in a theory-relevant manner should be inherent in the framework.

## Experts and Novices

Learning has only recently begun to be a topic of major significance for cognitive psychologists interested in problem solving. Following general trends in cognitive research, the focus of the first phase of problem solving research was concerned with characterizing the general problem solving capabilities of the adult and employed abstract puzzle problems. The second phase of problem solving research focused on the differences between experts and novices solving problems in complex content domains. A third phase that could probably be characterized as the study of the adaptive learning system seems to be emerging. The present research is in the spirit of the third phase.

### Phase I - General Problem Solving

During the first phase of problem solving research, a major goal was to build similations of human problem solving behavior that were as powerful as possible without assuming any built-in domain-specific knowledge. The simulations of human problem solving of that era, the

General Problem Solver (Newell, Shaw, and Simon; 1960) and Human Problem Solver (Newell and Simon; 1972), solved problems using means-ends analysis. Means-ends analysis is a problem solving strategy which involves determining the current state and the end state, and finding an appropriate method to reduce the distance between them (Hays, 1981). Research using a variety of puzzle-type tasks, such as Tower of Hanoi (Simon and Hayes, 1976), Missionaries and Cannibals (Reed, Ernst, and Banerji, 1974), and cryptarithmetic (Newell and Simon, 1972), indicated means-ends analysis is a primary strategy for solving problems.

Phase II - Experts and Novices

The orientation of problem solving research shifted when it became desirable to model problem solving behavior in complex real world domains. To solve a problem in a complex domain, large amounts of domain-specific knowledge are needed to generate expert-like behavior. Yet, as Bhaskar and Simon showed (1977), a general problem solving system with an encyclopedia of domain-relevant knowledge merely appended is not sufficient to generate a good approximation to expert behavior; the organization of this knowledge also is important. Therefore, different kinds of questions about problem solving became important: How is information encoded by experts and novices? How is an expert's knowledge of a domain structured? How do experts access their knowledge?

Three types of tasks have been employed in attempts to access underlying cognitive structures: perceptual tasks, memory tasks, and categorization tasks. In the perception task used by Chase and Simon (1973) to study differences between expert and novice chess players, subjects were shown a board of chess pieces and were asked to replicate the board while it was in full view. There was no difference in the number of pieces placed after one glance, but the expert was considerably faster than the novices at reconstructing boards from actual games. There was no difference in performance on the random boards. Presumably, the expert's knowledge of typical chess positions enabled faster performance.

Memory tasks have been used to examine behavior in a wider range of domains, including chess (Chase and Simon, 1973), Go (Reitman, 1976), and computer programming (Adelson, 1981; McKeithen, Reitman, Rueter, and Hirtle, 1981). The general procedure is to provide examples of both normal and randomly arranged stimuli for some short period of time, and then ask subjects to recall the items. Subjects in chess and go experiments see actual game board configurations, while computer programmers see programs in either normal or scrambled order. As with the perception task, no differences between experts and novices are reported when the stimuli are randomly arranged. However, with ordered materials, experts recall more items overall. Taking number of items recalled before a significantly long pause as a measure of chunk size, there is no difference between experts and novices in number of chunks recalled, but experts recall larger chunks. The chunks for experts and novices differ in nature as well: games experts recall pieces together which have some functional relationship such as an attack configuration, while computer programmer experts group items that share semantic information. Novices recall items that are spatially close to each other or that share syntax.

The memory experiments suggest that much of the power of the expert lies in the ability to quickly recognize meaningful stimuli as members of various functional categories. Hinsley, Hayes and Simon (1978) showed that college students can categorize algebra word problems into types very quickly, and that these types provide information that is helpful in solving the problem. Thus, skilled problem solvers may have developed problem "schemata" or categories that indicate possible solution strategies for problems conforming to the solution type. Based on this and similar results, Chi, Feltovich and Glaser (1981) hypothesized that expert-novice differences may be related to the "poorly formed, qualitatively different, or nonexistent categories in the novice representation."

To investigate this hypothesis, Chi et. al. (1981) asked novice and expert physicists to perform a variety of tasks, including categorization tasks, to learn how their knowledge of the domain was structured. Their results revealed that experts tend to sort physics problems on the basis of major physics principles, while novices sort on the basis of the entities involved in the problem statement or the surface structure. They suggest that both groups consider the surface cues present in the statement of the problem, but that experts then engage in a higher level of analysis that yields information about the physics principles involved and that may be contrary to the expert's first impression based on the surface features, whereas novices use the cues as they exist.

Physics experts also engage in a type of analysis that is different from means-ends analysis and is not used by novice physicists, i.e. forward knowledge development (Larkin et.al. 1980). A novice using a means-ends analysis strategy would search for an equation which would yield as a result the desired quantity and then attempt to determine the information needed to fill in the equation, whereas an expert engaging in knowledge development would invoke physics principles only when some new piece of information could be generated from the immediately available information.

In summary, the literature on experts and novices indicates that the advantage of the expert lies not merely in the possession of more domain-related knowledge, but in an organization of that knowledge that is based on deep, functional relationships between concepts. Experts engage in an initial period of qualitative analysis that yields structural information useful in guiding problem solving.

Phase III Adaptive Learning Systems

Several researchers in problem solving have recently come to believe that learning must be incorporated as a central feature in models of problem solving (see Anzai and Simon, 1979; Anderson, Greeno, Kline, and Neves, 1982; Lewis and Anderson, 1985). Anzai and Simon (1979) simulated the learning of increasingly sophisticated methods of solving the Tower of Hanoi problem through the process of solving the problem several times.

More relevant to the present research, Anderson et.al. (1982) attempted to model the acquisition of a complex cognitive skill in a manner that is compatible with the expert-novice framework. Anderson et. al. suggested that students learning to solve geometry proof problems obtain information from two sources: declarative rules and worked out examples from the text. Similarity of new problems to worked out examples is important to novices in many domains. However, Anderson et.al. (1982) note that the analogies novices make are commonly based on superficial relationships. Thus, the novices can easily be led down the wrong solution path. Once a novice has successfully solved several problems using analogy relations, a problem schema or structure for understanding the problem in more general terms is developed. In this system, this structure undergoes a process of transition from declarative to procedural knowledge. Finally, the novice learns to distinguish problems for which the schema is applicable from problems for which it is not.

Although Anderson et.al.'s (1982) simulation was not formally completed or extended to other domains, the adaptive systems orientation offers an insight not made during either Phase I or Phase II: analogies may be an important general mode of problem solving for novices and perhaps experts on some occasions (see Clement, 1981 ). Novices may fail to solve many problems because they do not have sound bases for making analogies and rely on features that are not predictive of the problem category. The object categorization literature will now be examined as an example of another set of domains where people must learn to discriminate predictive and non-predictive features of a stimulus.

## Categorization of Natural Objects

One of the major conclusions resulting from the expert-novice research is that both experts and novices categorize problems according to solution type, and such categorization may yield information which is useful in solving the problem. Differences in performance between novices and experts are thought to result mainly from differences in the organization of the categories.

The literature on problem solving has tended to consider categorization of problems within complex domains, such as physics and computer science, in isolation. There have as yet been no attempts to link these results with more general theories of categorization. However, given that categorization may be considered to be one of the most basic of human cognitive traits, it may be possible to enhance our understanding of the organization of expert and novice knowledge by considering behaviors related to problem solving in the context of a model of categorization.

Recent models of categorization behavior have been based on research with natural objects. A compelling advantage of considering categorization of problems as possibly analogous to categorization of natural objects is that acquisition of object categories has been studied extensively, and the findings may be applicable to the acquisition of categories in problem solving. A claim will be made here that is similar to an argument made by Lewis and Anderson (1985): the acquisition of problem solving categories, i.e. classes of problems that require the same operator, involves basically the same processes as the acquisition of natural object categories.

## The Nature of Catgeories

A category is said to exist whenever two or more distinguishable objects, events, or problems are treated in the same way. These treatments may include labeling two entities with the same name, performing the same physical action with two objects, and using the same operator to solve two problems (Mervis, 1980; Mervis and Rosch, 1981).

While there has been little argument concerning the conditions under which a category may be said to exist, there has been less unanimity in determining a general definition of "category." There have been two major thrusts to define category: the first will be referred to as the Traditional theory, and the second as the Best Examples theory (the formulation of these two theories follows Mervis, 1980; and Mervis and Rosch, 1981).

Traditional Theory. The Traditional theory (as conceived of by Bourne, 1968 and presented by Mervis, 1980, and Mervis and Rosch, 1981) defines a category by a set of criterial attributes which all members of a category must possess and no nonmembers of a category may possess. The ideal definition is minimal, consisting of only those attributes that are necessary to distinguish exemplars of one category from exemplars of another category. Hence, if the world consisted of only red and green circles, the definitions of red circles and of green circles would only need to include the color attribute, since shape is not a distinguishing characteristic. Given that membership in a category is specified by possession of the set of criterial attributes, the boundaries of the category are well defined, i.e. a stimulus can always be unambiguously classified as a member of a certain category. The criterial attributes definition also implies that all stimuli with these properties are equally good members of a category, so a large green circle and a tiny green circle would be considered equally good examples of green circles. The final tenet of the traditional theory is that the world is a total set, which means that values of any attribute occur may occur in combination with every other value of every attribute present. So in a world with two colors and two sizes, large green circles, tiny green circles, large red circles, and tiny red circles would occur equally often. Thus, the way in which one chooses to divide the world into categories is completely arbitrary.

In the world of real objects, as opposed to worlds with arbitrarily defined characteristics, it is much more difficult to state the defining features of a category. Although it cannot be proven impossible, no one has yet determined a single set of features that are criterial for the category "bird". Yet, there are features which are commonly associated with birds, such as wings, feathers, beaks or bills, and the ability to fly. As these features are not strongly associated with any other category, one might say these features are correlated with the category "bird". To address these concerns about real world categories, the Best Examples theory was developed.

Best Examples Theory. The Best Examples theory (as presented by Mervis, 1980, and Mervis and Rosch, 1981) proposes that membership in real world categories is defined by possession of a subset of a family of features, rather than a full set of criterial attributes. This family of features may consist of a set of overlapping attributes which are common to many members of the category, but each individual feature need not be possessed by all members of the category. For example, ducks have bills, wings, feathers, and can usually fly, while chickens have beaks, wings, feathers, and usually can't fly very far. These members of the family of features are not necessarily limited to members to a single category, so that nonmembers of the category may share some of the features as well. A bat can soar and a platypus has a bili, but neither are members of the category bird.

The fact that members of a category need not share a single obvious set of criterial features suggests that category members may vary in how typical they are of the category. In fact, numerous studies have indicated that certain members of a category can be classified much more quickly than others. In a seminal study, Smith, Shoben, and Rips (1974) showed that people can respond "true" faster to the statement "A robin is a bird." than to "A chicken is a bird."

There is generally a high degree of agreement about which members of the category are the most typical members. These highly typical members, or good examplars, generally share many features with other members of the category, particularly general shape characteristics, and few features with noncategory members. Less typical members, or poor exemplars, share fewer features with members of the category and may share more features with nonmenbers of the category. Therefore, poor examples are harder to categorize than good examples, and easier to misclassify.

This featural overlap among categories yields "fuzzy" category boundaries, rather than well-defined boundaries. Bowerman (1980) argues that the boundaries of a category are sometimes culture-specific: different cultures place boundaries of word meanings in different places. For example, one can "open" and "close " a water faucet in Spanish, but this use of open" and "close" is not allowed in English.

Disagreements that people have concerning membership of items in a category will usually concern the boundaries of the category, and not the highly typical members of the category. Heider (now Rosch, 1973) found high agreement on the most typical value of red, even amongst cultures which had no color word for red, but much less agreement on the boundary for red. The lack of well-defined boundaries does not imply that the boundaries of a category are arbitrary. Rather, the Best Examples view is that categories are defined by natural, although less than perfect, breaks in correlated clusters.

The Traditional theory and the Best Examples theory each address different domains of concepts, well-defined artificial categories versus natural object categories that cannot be defined by a single rule. One might argue that abstract concepts, such as arithmetic operations, have many of the properties associated with artificial categories, and hence any theory concerning operation categorization should be guided by the Traditional theory. However, there is evidence that subjects see even artificial categories structured in the manner that Best Examples theory suggests: i.e., there is substantial agreement on which examplar in artificial categories is the best examplar

(Rosch, Simpson, and Miller, 1976).

Given the difficulty that many students experience in learning to apply arithmetic operators correctly to word problems, there seems little doubt that the relationship between the wording of an arithmetic problem and the operation which should be applied is probabilistic from the student's point of view. Thus, although either the Traditional theory or the Best Examples theory might conceivably provide an adequate framework for analyzing categorization of arithmetic problems, the Best Examples theory may be more appropriate.

A reasonable extension of the Best Examples theory to more abstract categories would say that students treat mathematical categories in much the same way they treat object categories, i.e., they look for specific words or words with similar meanings to those which have appeared before in mathematical problems known to be solvable using a certain operator.

There is anecdotal as well as experimental evidence that suggests people do rely on correlated features when attempting to solve word problems. The word "of" often appears in conjunction with fraction multiplication problems. This correlation seems so compelling that students actually are taught that the presence of the word "of" indicates one should multiply. However, note that one can easily write a problem using the word "of" that requires addition, subtraction, or division. Several studies of algebra and physics word problem solving indicate that decisions about solution type are made on the basis of surface structure cues, like the word "of", particularly by poorer problem solvers (Chi, Feltovich, and Glaser, 1981; Hinsley, Hayes, and Simon, 1977; Larkin, McDermott, Simon, and Simon, 1980; Silver, 1979, 1981). Experts presumably also attend to surface structure cues, in reading the problem. However, they are probably also able to gain abstract structural information from the problem that can be used to confirm any initial hypothesis concerning which operator should be used.

## A New Description of Expertise

If one were to try to state what is common to experts in varied domains, perhaps the most salient attribute would be that experts are able to categorize new information appropriately. For a speaker of English, this would mean labeling an object correctly that one had never seen before: an expert problem solver would be able to categorize new problems appropriately. Although there may be other characteristics that distinguish novices and experts, the abilit to categorize information appropriately seems of fundamental importance.

In general, one might say a novice has developed an expert understanding of a category when the category label is applied to all and only those stimuli to which a community of experts would apply the label. Therefore, an expert word problem solver is one who is able to categorize word problems correctly according to the operator or operators which are appropriate for solving the problem. This description of an expert satisfies the first criterion a framework for investigating expertise should have: it can be applied across a wide range of ages and domains.

The second issue that must be addressed is how a novice might reasonably become an expert. This issue might be approached by considering the development of expertise within a certain domain, labeling natural objects. The Best Examples theory has been applied to predict the kinds of errors children make before they become experts at labeling common objects. Such information may be informative in addressing the more general issue of how a novice becomes an expert.

Bowerman (1980) has suggested that parents generally select highly typical objects as the first referents for a word. Thus, the core of the child's category would be the same as the adults from the beginning. In fact, Mervis and Pani (1980) showed that category learning is much easier when the initial exemplars are good rather than poor examples. Adults may be sensitive to this, and choose their referents accordingly.

According to Bowerman (1980), children then try to apply the category label to other objects that share features with the first referent. Bowerman's (1980) hypothesis is quite similar in nature to a model proposed by Lewis and Anderson (1985). They propose that after novices are shown the first example of a word problem in which a certain operator is applied, they will try to make analogies to the first referent when encountering new problems. An argument has been made by Michener (now Rissland, 1978) that mathematics training should basically exploit the tendency of novices to solve problems through the use of analogy by providing a good stock of typical examples .

According to the Best Examples theory, once novices have identified the central tendency of a category, they generally will fail to apply category labels consistently correctly because they have not established the boundaries of the category. Thus, novice problem solvers become more expert as they learn to distinguish problems which require different operations for solution, but have similar sounding surface structures. Conversely, novices may fail to become experts because they are unable to distinguish arbitrary aspects of the problem statement, such as the problem setting, from indicators of the operation required for solution.

Finally, the third requirement, that one be able to meaningfully distinguish levels of expertise, may be easily addressed within this framework. Novices who are relatively more expert are able to demarcate more accurately boundaries between related concepts.

# CHAPTER III

## THE DOMAIN OF FRACTION WORD PROBLEMS

### The Educational Problem

The concept of a "fractional part" is ubiquitous, both in common speech and in mathematics. In fact, it has been described as one of the more important and complex ideas that children encounter during their elementary school years (Behr, Lesh, Post, and Silver, 1983). Competence in applying fraction concepts is essential for a mature understanding of rational number, which in turn is critical for understanding of algebraic operations. Yet, despite the centrality of rational number concepts, growing evidence indicates large numbers of high school students (see Carpenter, Corbitt, Kepner, Lindquist, and Reyes, 1980) and adults (see Watson, 1980) are not able to use rational number concepts in a fluid manner.

In addition to the need for educationally relevant research on fraction word problems, the domain of fraction word problems has several other qualities that make it an appropriate candidate for studying categorization behavior within the Best Examples framework. The domain is fairly circumscribed and abstract, and the fact that fractions are commonly taught in elementary school should ensure the existence of large numbers of competent and less competent subjects.

Three Aspects of Competence

Competence in reasoning with numerical quantities has been viewed by educational studies (such as the National Assessment of Educational Progress, 1980, and the Rational Number Project, 1983) as being composed of three interrelated aspects: 1) knowledge of relative sizes of numbers, including equivalencies, 2) procedures for computing the results of arithmetic operations, and 3) the ability to set up the computation necessary to solve a word problem. The present research focused mainly on word problems, with the intent of determining which characteristics of problems lead people to classify two or more problems as similar. However, because these three aspects of knowledge are interrelated, to gain an adequate picture of the abilities of the nonexpert it is necessary to consider competence in each of these areas.

The results of two major educational studies, the 1978 National Assessment of Educational Progress (NAEP, reported in Carpenter et.al., 1980) with 9, 13, and 17 year-olds, and the Rational Number Project (RNP, Lesh, Landau, and Hamilton, 1983) with fourth through eighth graders, indicate many students have difficulty with all three aspects of fraction knowledge, but have more difficulty solving fraction word problems than fraction computation problems. These results corroborate the results of the previous NAEP examination, and have been interpreted as indicating that although students may have a rote understanding of computational routines, they have little insight into problem structure. The results of interview studies support this interpretation (see Behr et.al., 1983 and Hunting, 1984).

Knowledge of Relative Sizes. Many students appear to have poor intuitions concerning the relative sizes of fractional numbers. Only 58% of a large sample of Australian 14 year-olds successfully answered the following question: Which of the following fractions is closest to 3/16? a) 1/16 b) 1/4 c) 3/8 d) 7/8 e) 1/2 (Bourke, Mills, Stanyon, and Holzer, 1981, as reported in Hunting, 1984). Considering that performance is generally better when the numbers are halves, fourths , and eighths than other sets of fractions, this result is not encouraging. On a more difficult task given in the NAEP, only 12% of a sample of 17 year-olds were able to order 5/8, 3/10, 3/5, 1/4, 2/3, and 1/2 correctly from smallest to largest.

Hart (1981) has suggested that junior high school students do not see fractions as extensions of whole numbers, and therefore have difficultly locating fractions on a number line and judging their relative sizes. Exercises on both the NAEP and RNP asking students to locate numbers on a number line support this observation. Some of the difficulty of indicating placement on a number line may be a result of a confusion between the interpretation of a fraction as a quantity and the fraction as an operator (see Behr et.al., 1983). Thus, some children might place a mark meant to indicate 1/2 in the middle of the number line they were given, rather than at 1/2, because they saw 1/2 as an operator. However, this category of errors does not seem to account for all types of errors.

Performance in identifying pictorial representations of fractions is somewhat better when the arrangement is not a number line. The NAEP results indicate most 13 year-olds can identify fractions represented by simple pictorial representations, although Peck and Jenks (1981) suggest students are less successful when they try to draw their own pictures. The results on the RNP were similar, but in addition indicate that more complex pictures, whose units are odd shapes or are compositions of more than one separate shape, make the identification of the fraction much more difficult. Thus, many students display evidence of a limited ability to comprehend fraction knowledge through pictures.

Computational Procedures. Given that many students have only moderately adequate intuitions about the relative sizes of fractions, it is not surprizing that computational procedures should appear to be applied in a rote fashion (Carpenter et.al., 1980). Although computational skills probably receive a disproportionate amount of instructional time, as compared to relational sizes and word problems, the time spent is not as well reflected in test scores as such emphasis warrants.

The NAEP test results indicate over 80% of 17 year-olds can add two fractions with like denominators successfully; however, performance drops to 65% or less when the denominators are different. Somewhat counter to intuition, if the denominators are not the same, the complexity of the fractions had no effect on level of performance. Students seem to have a procedure to be applied when unlike denominators arise and they can apply it with similar results to any set of fractions.

The most common error made in addition on both the NAEP and the RNP was adding both the numerators and denominators of the fractions involved. The pervasiveness of this error was reflected in pictorial representations as well: 38% of the students taking the RNP indicated a picture of 2/9 when asked to choose a picture representing 1/6 + 1/3. Only 19% chose the correct picture of 1/2.

Those students answering 2/9 seem to lack well developed models of relative size, which could rule out 2/9 on the basis that it is not even as large as 1/3. The lack of well developed models of size is also reflected by performance on another NAEP item: only 37% of 17 year-olds were able to estimate an answer to 12/13 + 7/8.

The NAEP and RNP reported fewer results for subtraction problems, but in general, patterns of solution for subtraction follow the same trends as performance on addition problems, with slightly lower overall performance. Performance on multiplication problems was just slightly lower than performance on addition problems when proper fractions were involved: approximately 2/3 of the 17 year olds taking the NAEP could multiply two proper fractions or a fraction by a whole number. Performance dropped to below 50% when one or both of the numbers was a mixed number (Note - division problems were not included in these studies). Thus, although students may have an algorithm for performing a particular computation, these results seem to indicate they may not have a good understanding of the underlying concepts and processes.

Word Problems. More convincing evidence that students do not understand underlying concepts may be found in an examination of performance on word problems. Word problems tend to be more difficult for students than computational problems. However, performance on addition and subtraction word problems was slightly better than performance on computational problems that required the same operation for solution and had similar fractions. Carpenter et.al. (1980) suggest that the word problem actually assists in the computation for addition and subtraction problems: the context may rule out certain possibilities.

The relationship between performance on word problems and computational problems is quite different for multiplication problems. Only 30% of the 17 year-olds who took the NAEP were able to solve a single step fraction multiplication word problem, while 70% were able to solve a similar computational problem. Performance on the RNP was similarly poor. Carpenter et.al. (1980) interpret this performance as evidence that computational algorithms, at least in the case of multiplication are applied mainly in a rote manner. Because the difference in performance between word problem solving and computational problem solving is so much greater for multiplication than addition and subtraction problems, it suggests that children do not have well developed conceptions of the kinds of situations which are associated with multiplication.

<u>Sources of Difficulties</u>

The results of test questions in all three aspects of fraction knowledge serve to confirm the common observation that knowledge of fractions is difficult to acquire and use effectively. At least part of the difficulty may be attributed to "buggy" algorithms (see Brown and Burton, 1978; Brown and VanLehn, 1980) for a discussion of bugs in subtraction of whole numbers). A number of students responded to questions on the NAEP and the RNP in a way that seemed to reflect an algorithm with a missing or imperfect step. For instance, when multiplying two fractions, some students multiplied the denominators and added the numerators. Adding the numerators is just a slight deviation from the actual algorithm which requires one to multiply the numerators. In this case, the student is obviously unable to rule out the result on a logical basis.

However, difficulties with fractions cannot be attributed to algorithmic confusions alone. In solving a word problem, the problem solver must first decide what the appropriate action should be, or alternatively, be able to choose the correct action once several actions have been considered. The ability to choose a correct operator requires access to problem schemata that indicate the range of situations to which the operator is applicable. The schema for multiplication problems, in particular, appears to be not well formed for many students, given the difficulty encountered with multiplication word problems.

One of the limitations of the NAEP and the RNP data sets is that although they provide an indication of problem areas for many students, they are not comprehensive enough to to provide much information about the characteristics of students' problem schemata. A goal of the present research is to characterize students' arithmetic word problem schemata and to determine what information from a problem is used to categorize it. For fraction word problems, possible cues for problem classification might include the characters and setting depicted in the problem, key words commonly associated with an operation, such as "of" or "all together", and the general pattern of actions in the problem statement. The NAEP and RNP results on word problems suggest students may rely on different kinds of cues when they attempt to solve multiplication problems than with addition or subtraction problems. Thus, one might expect a study of problem categorization to find differences among operations.

The results of these assessments raise some important issues for the present study of problem categorization. First, since students perform poorly when asked to compute answers to multiplication word problems, one might expect students to also have difficulty recognizing a multiplication problem, since the NAEP results indicate this poor problem solving performance is not due to poor computational skills alone. Therefore, the first question is: Are differential rates of categorizing problems related to skill in recognizing problems of particular types?

Second, there is a rather surprizing result that if the denominators are not the same, computational problems with more complex fractions are not significantly harder to solve than problems with simpler fractions. Therefore, does the complexity of the fractions make a difference when one is attempting to categorize a word problem? (i.e., do less complex numbers facilitate understanding the action pattern in a problem?). These domain relevant questions will be incorporated in the present study on problem categorization.

# CHAPTER IV

## DISTINCTIONS AMONG WORD PROBLEMS

The limited conclusions which can be drawn from the NAEP and RNP examinations indicate that a thorough investigation of all three aspects of fraction knowledge may not be a practical possibility. In fact, more useful knowledge might be gained by a more thorough investigation of a more limited area, such as fraction word problems. Given that mathematics problems do not generally present themselves as computations, but must be formulated as word problems in the real world, and given that students are presumably taught computational techniques in order to be able to solve word problems, the present studies were designed to focus mainly on word problems.

One fundamental question is whether there are any systematic differences in ease of solution among word problems that require the same operation for solution. Research on addition and subtraction word problems with young children indicates there may be considerable variablility in the ease of solution of different types of word problems. Given that "A word problem identifies some quantities and describes a relationship among them" (Riley, Greeno, and Heller, 1983), the relationship may vary in the degree to which it is easily modeled and hence, solved. In other words, if as Briars and Larkin (1984) suggest, children solve word problems by acting out the problem with counters in their heads, it may be more difficult to act out certain problems correctly.

For example, the following two problems (from Riley et.al., 1983) could both be solved by adding the two numbers given in the problem. However, the rates of success by first graders in solving these problems are remarkably disparate: Problem A was solved by 100% of the subjects, while Problem B was solved by only 6% of the subjects.

| | |
|---|---|
| | Joe had 3 marbles. |
| Problem A | Then Tom gave him 5 more marbles. |
| | How many marbles does Joe have now? |
| | |
| | Joe has 3 marbles. |
| Problem B | He has 5 marbles less than Tom. |
| | How many marbles does Tom have? |

There are several differences between these two problems, including the presence of an exchange, the person mentioned in the question statement, and whether or not there is a comparison. Although less obvious, these problems also differ in the "algebraic open sentence" (or abstract algebraic equation) which most closely represents the problem. If X is the first known quantity mentioned, Y the second known quantity mentioned, and ? the unknown quantity, then Problem A seems to be most aptly described by 'X + Y = ?' and Problem B by 'X = ? - Y'. Although the final step in finding the solution to these two open sentences is the same, Problem B obviously requires the extra step of reformulating the problem by moving Y to the other side of the equation if it is to be solved by addition.

There are many ways that one might try to analyze word problems to predict ease of solution. However, recent research (i.e. Carpenter and Moser, 1982; Riley et.al., 1983; Briars and Larkin, 1984) is in agreement on one major point: children's improved ability to solve word problems as they become older is related to an improved ability to understand the more complex situations depicted in problems such as Problem B. Let us examine the ways in which children's attempts to solve addition and subtraction word problems have been studied.

## Review of Research on Addition and Subtraction

Early attempts to study the causes of differential rates of solution for addition and subtraction word problems focused on surface features of the presented text, such as number of words in the problem, presence of key words, and the size of the numbers involved (see Briars and

Larkin, 1984 for a review). Although these factors obviously have some influence on problem solving, they do not account for all of the observed differences.

Kintsch and Greeno (1985) have recently argued that understanding a problem text requires both specific information derived from the text, such as the numbers of objects involved, and a situational model that is developed based upon the reader's understanding of conceptual relations among quantities. They claim that the problem solver has schemata for various types of relations among quantities, which may be cued, perhaps by key words, and substantiated during the course of problem understanding. Although there has not been complete agreement concerning how specific the schemata must be (e.g. Briars and Larkin's modeling approach proposes much more general schemata based on verb understanding), recent research has supported the claim that the reader's understanding of the relations is important, and has attempted to account for differences in problem difficulty by distinguishing different types of problems that may be solved by roughly equivalent processes (see Carpenter and Moser, 1982; Nesher, 1982; Riley et.al., 1983; Briars and Larkin, 1984).

One way in which problems have been distinguished according to solution type is by the open sentence that most closely matches the problem statement, i.e. Problem A is most closely represented by X + Y = ? (see Briars and

Larkin, 1984 or Riley et.al., 1983 for a review). Although this distinction seems to capture some of the critical distinctions between problems such as A and B, it does not completely account for differences in rates of solution. The following two problems can both be represented by the same open sentence, A - B = ?, but Problem C is solved by 100% of a sample of kindergarteners, while Problem D is solved by 4% of the same subjects (from Riley et.al., 1983):

| | |
|---|---|
| | Joe had 8 marbles. |
| Problem C | Then he gave five marbles to Tom. |
| | How many marbles does Joe have now? |

| | |
|---|---|
| | Joe has 8 marbles. |
| Problem D | Tom has 5 marbles. |
| | How many more marbles does Joe have than Tom? |

These two problems differ on the dimension of action: Problem C involves an active transfer of a set of objects from Person X to Person Y, while in Problem D the number of objects each person has remains the same and the sizes of the sets are compared.

Thus, another way of distinguishing problems is by the presence of an action. However, it is also possible to show through such examples that two problems that both have an action cue but have different open sentence structures would have different rates of solution. Thus, the presence of an action cue alone is not sufficient to predict problem solving success: both the form of the open sentence and the presence of action cues are necessary to predict problem solving success.

The most recent research on arithmetic word problem solving has involved both the form of the open sentence and the presence of action cues in problem organization structures, although the details of these models are slightly different (i.e. Riley et.al., 1983 versus Briars and Larkin, 1984). Both organizations differentiate betweenproblems with action cues, problems in which two sets are combined, and problems in which two sets are compared. Riley et.al. (1983) further separate problems involving action into those in which the intent of the action is to make two sets equal and those in which it is not. In addition, both organizations consider which component is unknown: the initial quantity, the change to that quantity or comparison quantity, or the resultant quantity.

Ideally, it would be advantageous to include all types of the problems which have been previously defined. However, given that these studies have distiguished a dozen or more types of addition and subtraction problems, it was not feasible to conduct a thorough investigation of all types of problems in a first set of studies. Hence, the problems used for the present study were limited to those in which: 1) the result was the unknown quantity and 2) the fractions were proper fractions, i.e. had values less than one. This produced four types of problems, two problems in which one would add and two problems in which one would subtract. Furthermore, one of the two problems for both cases had an action cue, termed a <u>dynamic</u> problem, and the other did not, a <u>static</u> problem.

These four types of problems were: 1) the result of an incrementing change in ownership, 2) the result of two sets combined, 3) the result of a decrementing change in ownership, and 4) a comparison of the size of two sets.

The range of operations was expanded to include multiplication and division problems in which the result was unknown. Again, one problem had an action cue and one did not. The four multiplication and division problems were: 1) the result of a fractional transformation of a fractional part, 2) the relationship of a fractional part of a fraction to a whole, 3) the result of partitioning a fractional quantity, and 4) the size of a quantity given a fractional part. Examples of all eight types of problems are given in Table 4.1.

TABLE 4.1 - Eight Types of Fraction Word Problems

Addition

| | |
|---|---|
| Dynamic | Charlie had 3/6 of a can of cake frosting. His neighbor gave him another 1/6 of a can of cake frosting. How much frosting did Charlie have then? |
| Static | Rachel tested 7/16 of the lab animals, while Harry tested 3/8 of the lab animals. What fraction of the lab animals have they tested together? |

Subtraction

| | |
|---|---|
| Dynamic | Hansel began the trip with 3/4 of a pound of bread. He used 1/4 of a pound of bread to mark the trail. How much bread did Hansel have then? |
| Static | Hemingway used 47/80 of a box of typing paper last week, while Orwell used 34/40 of a box of typing paper. How much more paper did Orwell use than Hemingway? |

Multiplication

| | |
|---|---|
| Dynamic | Margret has 2/5 of a gallon of ice cream. She gave 1/5 of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive? |
| Static | 7/10 of the beds in the garden were planted with flowers. 2/7 of the flowers were tulips. What fraction of the garden was planted with tulips? |

Division

| | |
|---|---|
| Dynamic | Grace had 3/4 of a pound of chocolate bits. She needed 1/4 of a pound of chocolate bits to make a batch of cookies. How many batches of cookies could Grace make? |
| Static | Arlen mixed up 5/12 of a bucket of birdseed. He found he had enough birdseed to fill 7/12 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders? |

# CHAPTER V

## EXPERIMENT 1

### Introduction

The primary purpose of Experiment I was to try to determine what types of cues are used by nonexpert problem solvers in deciding what operations are necessary to solve fraction word problems and whether the cues that are used vary from problem to problem. An argument was developed earlier in this dissertation that one of the distinguishing characteristics of expertise is the ability to classify new problems and information appropriately according to the deep structure of the problem. In the present experiment, the presence of two types of cues was systematically manipulated in the context of fraction word problems: 1) surface structure cues, a similar story line for word problems and the same numbers for numerical equations and 2) deep structure cues, which indicate what operation is required to solve the problem.

The Best Examples theory would predict that when subjects are required to judge the similarity of several problems, both experts and nonexperts should correctly judge a word problem to be similar to a standard if both the surface structure and the deep structure are similar.

However, nonexperts should be more likely than experts to judge two problems to be similar if they share only surface structure similarity. In particular, nonexperts should be likely to base their judgments on surface structure similarity if the operation needed to solve one or both of the problems is not generally well understood. Thus, it would be more difficult to recognize a problem which shares only operation similarity with another problem. In contrast, experts should consistently state that two problems are similar only when they both require the same operation for solution.

## Possible Influences on Performance

Any two problems which require the same operation for solution must share some features of similarity, whether it is only on a deep, abstract level, or on a relatively surface level as well. Different types of similarity may influence how easily a problem can be correctly categorized, and hence solved. Four factors which may influence the ease of making a similarity judgment will be investigated in Experiment I. Differences in categorization performance may result from: 1) differences between arithmetic operations in ease of understanding the problems or modeling the actions, 2) differences among problems that require the same operation, but have different open sentence structures or differ in presence of an action cue, 3) difficulties related to the fact that a word problem must be compared with a word problem, rather than a numerical expression, or 4) difficulties related to manipulating the specific numbers involved in the problem.

Between Operations. There are two sources of information which suggest that fraction word problems requiring different operations are not equally difficult to solve, and therefore may not be equally difficult to recognize as requiring the same operation. First, for the 13-year-olds who participated in the NAEP, there was considerable variablity in performance in solving single step word problems with whole numbers in which the result was unknown. Ninety-six percent solved the addition problems correctly, 89% solved the subtraction problems, 77% solved the multiplication problems, and only 40% solved the division problems. Second, the results for one step fraction word problems in which the result was unknown revealed similar variability in difficulty among fractions: the rate of solution for addition and subtraction problems was close to 80%, with addition problems solved correctly slightly more often. The multiplication problems were solved by only 30% of the subjects. No division problems with fractions were given.

Based on these results, one would expect nonexpert subjects to have the least difficulty recognizing that two addition problems require the same operation, followed closely by subtraction problems. Multiplication should be considerably worse than either addition or subtraction, but data are not available to predict whether or not multiplication should be better than division.

Within Operations. There may be differences among problems that require the same operation, as well as differences between problems requiring different operations. As was discussed in Chapter IV, not all addition and subtraction word problems are equally difficult for young children. Even if the open sentences that best represent two problems are the same, it may not be equally as easy to say that these two problems require the same operation. The presence or absence of action cues influences ability to solve a problem, as well as the type of open sentence. Thus, another purpose of the first experiment was to investigate differences between fraction word problems that did and did not incorporate action cues.

In order to make the experimental design feasible, only problems for which the result was unknown were included in the study (i.e., the open sentences were of the form "A + B = ?,"A - B = ?," "A x B = ?," or "A / B = ?"). There were two types of problems within each operation: 1) dynamic problems that had action cues, and 2) static problems, which had no action cues.

Format of Problems. A third factor which might influence the ease of categorization is the format in which the problems are presented. The same computation could be presented either as a numerical sentence, such as 1/3 ÷ 1/4 = ?, or as a word problem.

If the subject first tries to determine what operation is required before finding another problem that matches in operation, it may be easier to recognize a word problem that requires the same operation as a numeric statement. The operation, in this case, need not be extracted from the second problem since it is already given.

According to this logic, subjects should perform better with a word problem standard when given numeric alternatives than when given alternatives that are word problems, since they must determine the operation required to solve only one problem. The next least complicated setting would be a numeric equation presented with word problem alternatives, followed by the comparison of word problems to word problems. However, subjects may not try to determine the operation required first and instead look for similarity in the wording, in which case the best performance should result with comparisons of word problems to word problems.

Computational Ease. A fourth factor which may influence the ability to categorize word problems is the ease of computing the actual answer. Although the NAEP study did not find any major differences in ability to solve problems that were related to the type of fractions involved, the result seems unintuitive. It seems that if the numbers involved in a problem are easy to think about or to visualize, the problem should be easier to solve. In fact, very young children can solve problems with very small numbers that they cannot solve with larger numbers (Gelman and Gallistel, 1978). To investigate whether the type of fraction involved did make a difference in ability to categorize problems, two types of fractions were used: 1) Easy fraction pairs, for which the common denominator was either 3, 4, 5, 6, 8, 9, 10, 12, or 16, and 2) Hard fraction pairs, for which it would have been difficult to find a common denominator, such as 12 and 13.

## Method

### Subjects

Sixty-three undergraduate students enrolled in psychology courses at the University of Massachusetts participated in Experiment I in exchange for course credit. Only those subjects who displayed above average verbal abilities, as indicated by a stanine score above 5 (of 9) on the Peabody Picture-Vocabulary Test (PPVT), were included in further analyses, to ensure that poor performances could not be attributed to verbal ability alone. Thus, analyses were conducted on the data of 47 subjects, 16 males and 31 females. Their average age was 20.8 years.

Five expert mathematicians also participated. They had either a degree in mathematics or considerable advanced mathematics training.

### Materials

The problems were presented in a booklet with wide margins for making notes or computations, if necessary (See Appendix I for test materials). The booklet was composed of two subsections: 1) Section I contained 56 randomly ordered items requiring judgments of similarity, and 2) Section II was composed of eight fraction word problems, to which subjects were to indicate the operation needed to solve the problem.

Section I contained two types of items, 32 single step arithmetic word problems (one operation was required to solve the problem) and 24 translations of single fractions that were included as filler items. All 56 items had a similar structure: a standard was presented with four alternatives. The subject was to choose which of the alternatives "went the best" and the "next best" with the standard. The instructions stated that the subjects should have labeled "best" that alternative which matched the standard in both the story line or numbers used and the operation required to solve the problem. They also stated that the second best match required the same operation or was the same fraction: it had no obvious similarity in the story line or numbers. A third alternative matched in the story line or numbers only, and the fourth alternative matched neither the story line or numbers nor the operation or fraction. For example:

Standard Ms. Gray reserved 4/8 of the seats in the theater. There were enough seats for 5/8 of the students in the school. What fraction of the theater was needed to fit all the students?

Best Match Ms. Gray reserved 5/10 of the seats in the theater. There were enough seats for 7/10 of

| | |
|---|---|
| | the students in the school. What fraction of the theater was needed to fit all the students? |
| Next Best | Nicole had 2/5 of a pound of popcorn. She used 1/5 of a pound of popcorn for each popcorn string she made. How many popcorn strings could Nicole make? |
| Story Line | Ms. Gray reserved 4/8 of the seats in the theater. She found enough students to fill 5/8 of the seats she had reserved. What fraction of the theater was filled by the students? |
| Neither | Matthew had 3/4 of a pound of radish seed. 1/4 of the seed did not sprout. How much seed did not sprout? |

Of the 32 arithmetic problems in Section I, there were 8 standards that required each of the 4 operations. Half of the problems had easy fractions and the other half had hard fractions. There were also four types of comparisons: 1) a dynamic word problem standard with word problems as alternatives, 2) a static word problem standard with word problems as alternatives, 3) a word problem standard with numeric problems as alternatives, and 4) a numeric problem with word problem alternatives.

The remaining 24 problems in Section I involved translations of fractions depicted in either a numerical, pictorial, or written problem format. These problems were used as filler material and were not analyzed separately. In Section II, subjects were given eight fraction word problems, and were asked to decide whether they should add, subtract, multiply, or divide the two numbers given in the problem to obtain the correct answer. There were two problems requiring each operation, one dynamic and one static.

Procedure

The subjects were run in small groups of no more than six people. The instructions were provided in the written booklet (see Appendix I for instructions and test questions). The subject was instructed to choose the alternatives that went "the best" and the "next best" with the given problem. Two sample items were provided to ensure that the subjects understood that their choice should be based upon the operation or the fraction. After they had read the instructions and the sample problems, subjects were told that it was not necessary to solve any of the problems, but it was permissable to write in the booklet. Any further questions were answered at this time. No time limit was given.

When the subject had completed half of Section I, there was a short break during which the PPVT was administered individually by a second experimenter in another room. The subjects then returned to the original room and completed the questionnaire.

Design

The data from the similarity judgment task involving the 32 word problems in Section I were analyzed according to a design with one between-subjects factor and four within-subjects factors. The between-subjects factor, Error Level, was a measure of word problem solving ability that was calculated on the basis of the subject's score on Section II. There were 10 subjects who misclassified the operation of only one or none of the eight problems in Section II. There were three groups of less expert subjects, 11 who made two errors, 14 who made three errors, and 12 who made four or five errors. The four within subjects factors were Operation, the Difficulty of the numbers involved (Easy or Hard), Type of comparison (Dynamic-Word, Static-Word, Word-Number, Number-Word), and Cue choice (correctness on best match, correctness on second best match).

Section II was also analyzed independently of Section I to determine what kind of errors subjects made. There were two within subjects factors in this design: arithmetic Operation and presence of Action Cue in a problem.

## Results

### Scoring

Data from Section I were coded in two ways: 1) according to the relationship between the standard and the subject's two choices and 2) for correctness, i.e. matching the standard in operation or not. Each of the two choices was transformed into a code that indicated what type of cues were present in the problem. The alternative that matched in both surface structure and deep structure will be referred to as B, the deep structure only match is D, the surface structure only match is S, and the remaining alternative that matched in neither dimension is N. These pairs of choices were tabulated over all subjects.

The second coding indicated whether these choices were correct. The subject received a score of 1 for the B Cue if either of the two choices was a B, and a 0 otherwise. A score of 1 was given for the D Cue if either of the two answers was a D. These scores were analyzed in a 4 (Error Level) x 2 (Cue types - B and D) x 4 (Operation) x 4 (Type - Dynamic-Word, Static-Word, Word-Number, Number-Word) x 2 (Difficulty - Easy, Hard) ANOVA.

For Section II, the correctness of the choice of operation to solve the problem was of interest. The type of incorrect operations chosen was also analyzed. Section II will be discussed first, since the results of this section were used to determine the subjects' levels of expertise.

Section II

Section II was developed to measure expertise. For this purpose, a subject received a score from 0 to 8 indicating number of errors in stating the operation to be used in solving each of the eight problems. The subjects were divided into four Error Level groups on the basis of these scores: 10 subjects made 0 or 1 error, 11 subjects made 2 errors, 14 subjects made 3 errors, and 12 subjects made 4 or 5 errors.

The responses to the items in Section II were also of interest independently of Section I (see Figure 5.1). A 2 (Action Cue - present or not present) x 4 (Operation) ANOVA revealed a main effect of Operation, $F(3,141) = 45.95$, $p < 0.0001$. The mean scores for problems requiring each operation were in the following order from highest to lowest: addition, subtraction, division, multiplication. The correct operation was stated for addition problems more often than for multiplication, $t(47) = 9.49$, $p < 0.0001$, or division, $t(47) = 6.37$, $p < 0.0001$. Responses to subtraction problems were also significantly better than responses to multiplication, $t(47) = 8.76$, $p < 0.0001$, or division, $t(47) = 4.7$, $p < 0.0001$, problems. The difference in performance on multiplication and division problems was

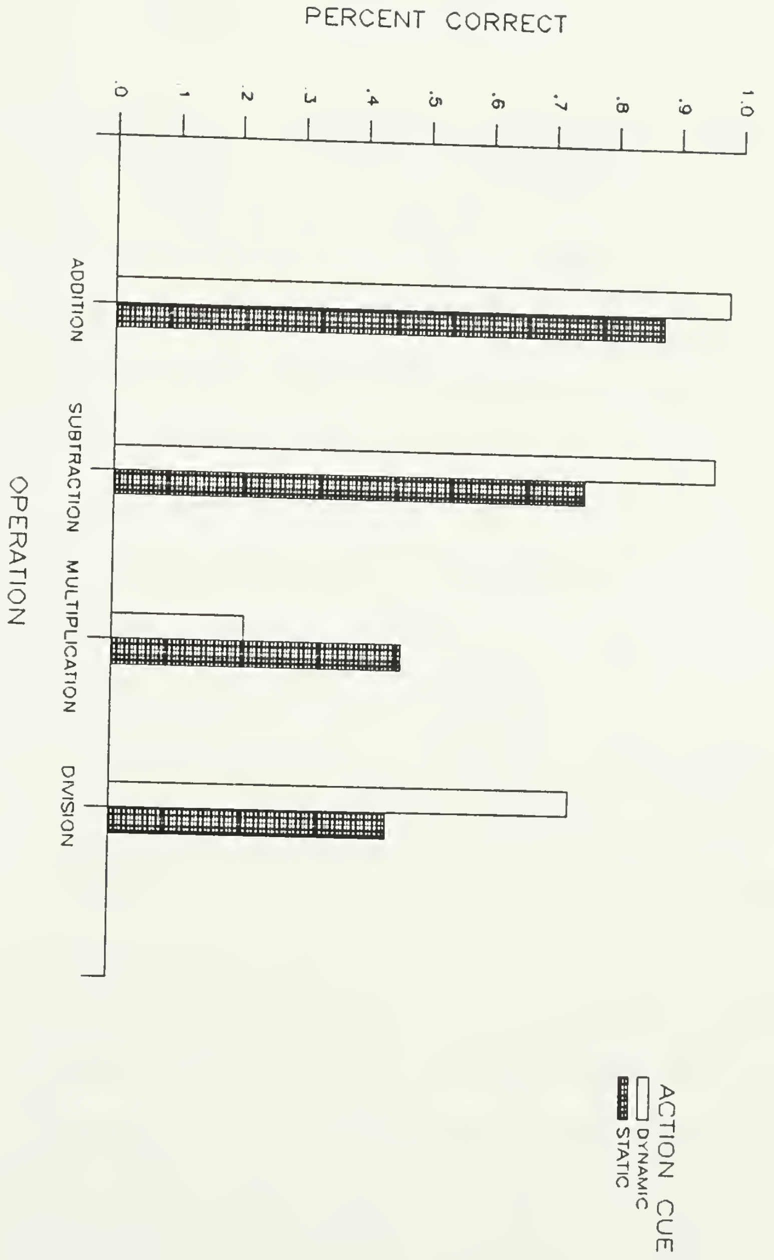



FIGURE 5.1 – SECTION II: ACTION CUE VS OPERATION

significant: more division problems were responded to correctly, $t(47) = 4.07$, $p = 0.0002$.

There was also a main effect of Action Cue, $F(1,47) = 6.02$, $p = 0.0179$, in which the correct operation was indicated more often for Dynamic problems than for Static problems.

Action Cue and Operation interacted, $F(3,141) = 10.81$, $p < 0.0001$), limiting the interpretability of the effect of Action Cue. Of the two problems given for each operation, the correct operation was chosen for the Dynamic problem significantly more often for every operation except multiplication: addition, $t(47) = 2.34$, $p = 0.0237$, subtraction, $t(47) = 3.14$, $p = 0.0029$, and division, $t(47) = 2.96$, $p = 0.0048$. For multiplication, performance was better for the static problem, $t(47) = -3.58$, $p = 0.0008$. The results of Section II are summarized in Figure 5.1.

The actual responses to each of the eight problems are given in Table 5.1. Three problems are of particular interest because of the low rates of correct responding and high concentration of answers on a single alternative: Dynamic multiplication (10 of 48 correct, 30 subtraction), Static multiplication (22 of 48 correct, 25 division), and Static division (21 of 48 correct, 22 multilplication). It is apparent that many subjects cannot reliably distinguish static multiplication and division problems: 16 subjects confused the operations of multiplication and division for

TABLE 5.1 - Responses on the Operations Assessment Task

| | Response | | | |
|---|---|---|---|---|
| Problem Type | Add | Subtract | Multiply | Divide |
| Addition | | | | |
| Dynamic | 47 | 0 | 0 | 1 |
| Static | 42 | 0 | 6 | 0 |
| Subtraction | | | | |
| Dynamic | 0 | 46 | 2 | 0 |
| Static | 1 | 36 | 1 | 10 |
| Multiplication | | | | |
| Dynamic | 1 | 30 | 10 | 7 |
| Static | 0 | 1 | 22 | 25 |
| Division | | | | |
| Dynamic | 1 | 1 | 11 | 35 |
| Static | 2 | 3 | 22 | 21 |

both of these problems. Of the 11 subjects who answered both problems correctly, 10 were in the 0-1 Error Level group. An $X^2$ test on the correctness of the responses for static multiplication and division problems indicates the choice of operation may be random, $X^2_1 = 2.106$, $p > 0.10$. This is not too surprizing, given the similarity of the computational algorithms.

There was a very different pattern of reponses to the Dynamic multiplication problem. Most (30 of 48) subjects interpreted this problem as a subtraction problem. However, they did not interpret either subtraction problem as a multiplication problem. Obviously, this confusion is asymmetric. More importantly, the two multiplication problems were not treated as members of the same category, since the rate and type of wrong answers were quite different.

## Section I

Sex Differences. Preliminary analyses revealed no main effect of sex, $F(1,35) = .43$, $p = 0.5165$. Of the 32 interactions that involved the sex variable, only two were significant: Difficulty x Operation x Sex, $F(3,105) = 3.16$, $p = 0.0278$, and Match x Operation x Sex x Error Level, $F(45,315) = 1.49$, $p = 0.0283$. One would expect 5 %, or 1.6 of the 32 tests to be significant by chance. Therefore, given the lack of a main effect and the possibility that two interactions could have occurred by chance, the sex grouping variable was eliminated to provide more power to test the remaining variables.

The Best Examples Theory. According to the Best Examples Theory, experts and nonexperts should perform similarly when there is a high degree of cue overlap between problems, i.e. they should both chose the B alternative. Differences in performance should result in choosing the alternative that only matches in operation, because experts are more able to distinguish the valid cues in the D alternative from the invalid cues in the S alternative.

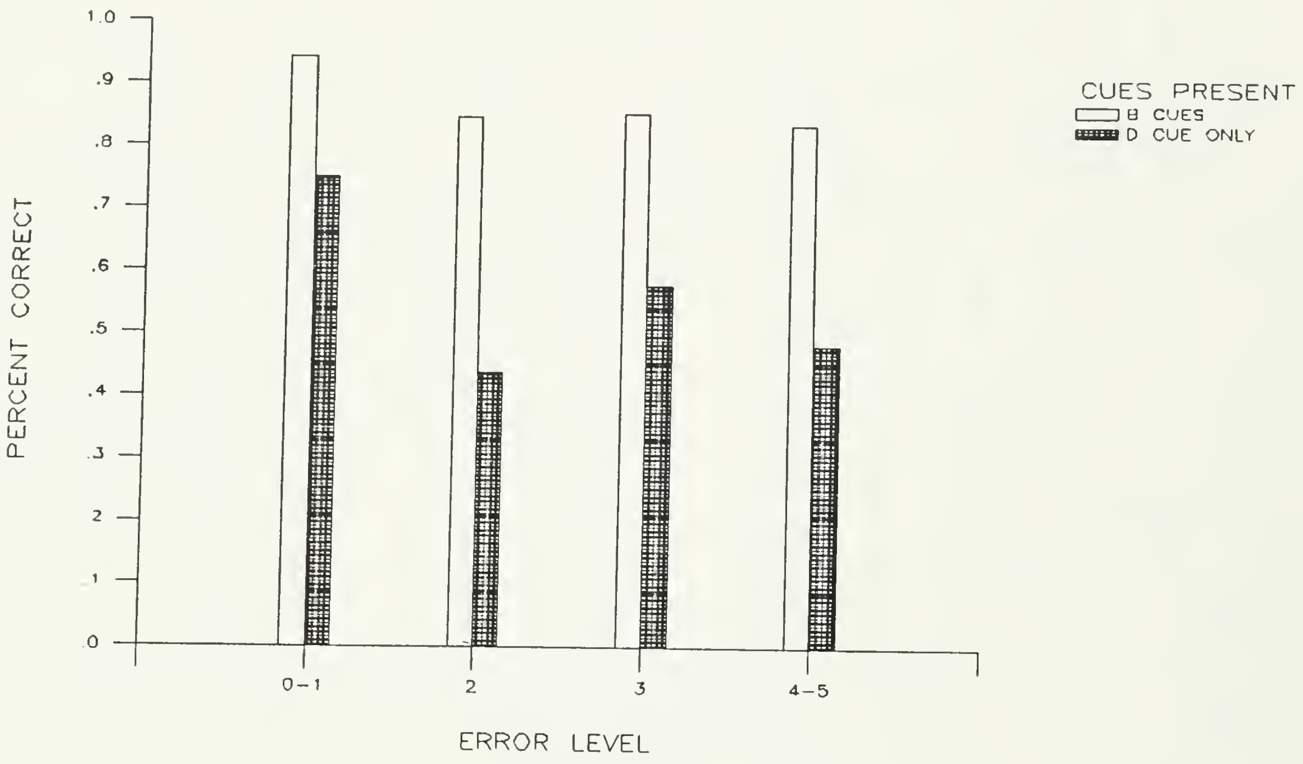
The results of Experiment I are consistent with these hypotheses. The five expert mathematicians consistently choose the B and D alternatives. Athough the subjects with 0 or 1 errors were probably less expert than the five mathematics experts, there was still a main effect of expertise, $F(3,43) = 7.17$, $p = 0.0005$. The nonexpert groups were not different from each other, $F(2,34) = 1.29$, $p = 0.289$, but overall the expert group performed marginally better than the average of the nonexpert groups, $t(43) = 1.99$, $p < 0.10$.

The main effect of Cue type was highly significant, $F(1,43) = 150.92$, $p < 0.0001$: the B alternative was chosen much more frequently than the *D* alternative (see Figure 5.2). This is predicted by the Best Examples Theory since there is less cue overlap. The Best Examples Theory also predicts an interaction between Error Level and Cue type: this interaction was significant, $F(3,43) = 3.52$, $p = 0.0229$. The expert group picked the *B* alternative slightly, although not significantly, more often than the combined nonexpert groups, $t(43) = 1.87$, $p < 0.10$, and the *D* alternative significantly more often than the nonexperts, $t(43) = 3.15$, $p < 0.01$.

Cue overlap was obviously a strong factor in making similarity judgments: 1197 of the 1536 pairs of alternatives chosen by all subjects were either *B - D* or *B - S* pairs. More importantly, of the pairs in which at least one alternative was not correct, i.e. not *B - D*, 468 of the 807 (or 58%) of the pairs were *B - S* pairs. The *B - S* pair was chosen more frequently than would be expected by chance from among the remaining five pairs of alternatives (*B - N* = 120, *D - S* = 88, *D - N* = 29, *S - N* = 102, $X^2(4) = 756.897$, $p < 0.0001$). Thus, when subjects made an incorrect similarity judgment, they tended to rely on surface structure overlap.

The effect of expertise seemed to be independent of the Operation and the Type of comparison involved, since Error Level did not interact with any variables other than Cue. Thus, experts appear to be generally more capable of

FIGURE 5.2 – SECTION I: ERROR LEVEL VS CUES PRESENT

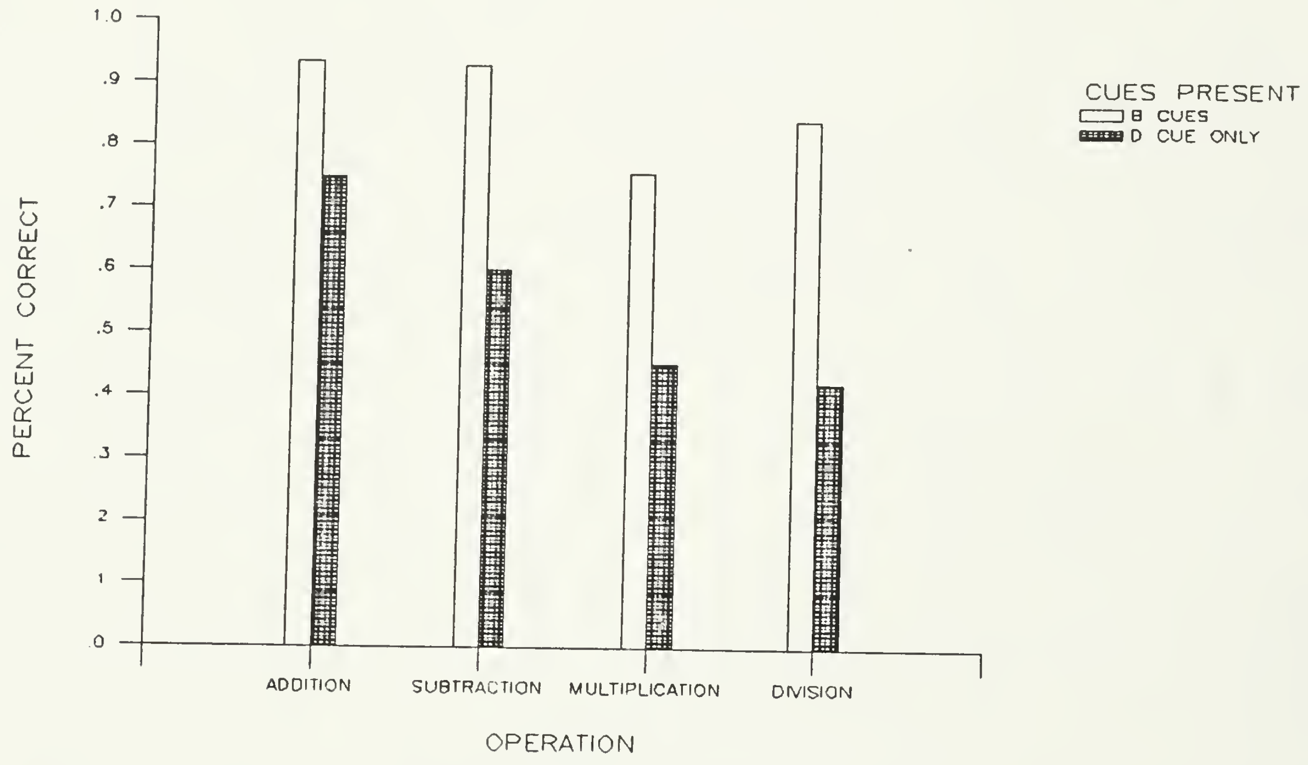


distinguishing relevant and irrelevant similarity.

*Operation.* As expected, given the results of previous studies, there was a main effect of Operation, $F(3,129) = 51.5$, $p < 0.0001$ (see Figure 5.3). Addition problems were responded to correctly more often than subtraction, $t(47) = 4.30$, $p = 0.0001$, multiplication, $t(47) = 10.51$, $p < 0.0001$, or division, $t(47) = 6.42$, $p < 0.0001$. In turn, reponses to subtraction problems were reliably better than responses to either multiplication, $t(47) = 7.52$. $p < 0.0001$, or division, $t(47) = 6.42$, $p < 0.0001$. Responses to multiplication and division problems were not significantly different. This pattern of results corresponds to the pattern of results obtained on both the NAEP and the RMP where the subjects actually computed the answers to word problems. A similar pattern was obtained in Section II as well.

Operation also interacted significantly with Cue type, $F(3,129) = 9.51$, $p < 0.0001$. Subjects chose the *B* alternative with equal frequency for addition and subtraction problems, but picked the *D* alternative more often for addition problems than subtraction problems, $t(47) = 4.15$, $p = .0001$). For multiplication and division problems, the reverse was true: the *B* alternative was chosen more often for division problems, $t(47) = -3.05$, $p = 0.0037$, but there was no difference in choosing the *D* alternative. The differences between subtraction and

FIGURE 5.3 – SECTION I: OPERATION VS CUES PRESENT



multiplication were the same for both the B and D alternatives.

Type. There were four Types of comparisons: 1) a dynamic word problem with word problem alternatives, 2) a static word problem with word problem alternatives, 3) a word problem with numeric equation alternatives, and 4) a numeric equation with word problem alternatives. The main effect of Type was marginally nonsignificant, $F(3,129) = 2.58$, $p = 0.0562$ (see Figure IIA, Appendix II). However, there was a significant difference between responses to Dynamic - Word items versus Static - Word items , with more correct responses to Dynamic - Word items, $t(47) = 4.14$, $p = 0.0001$. This difference was also observed in the results of Section II: responses stating the operation required for the solution of a word problem were generally better for Dynamic problems. Thus, there is additional evidence suggesting that all word problems which require the same operation are not responded to similarly.

Although there was no overall effect of Type, Type did interact with several variables. There was an interaction between Type and Match, $F(3,129) = 7.44$, $p = 0.0001$: the B alternative was chosen less frequently for Word - Number problems than for all other types (not significant), but the D alternative was chosen significantly more often for Word - Number problems than for the average of all other types, $t(47) = 4.67$, $p < 0.0001$. If a subject had correctly identified the operation which should be used to solve the word problem, it would be easier to chose a second alternative in which the operation was clearly indicated, as in a numeric equation.

Type also interacted with Operation, $F(9,387) = 12.65$, $p < 0.0001$ (See Appendix II for figure). Although no single format was best for all operations, for the operations of subtraction and division, the Word - Number problems were responded to correctly most often . For the operation of addition, Word - Number problems ranked second for number of correct responses, while for multiplication problems, Word - Number problems ranked last. Thus, responses to Word - Number problems were generally better than responses to other types of problems, except for multiplication problems. Considering only Word - Number problems, reponse to multiplication problems was significantly worse than performance on all other types of problems combined, $t(47) = 6.68$, $p < 0.0001$.

The pattern of results for the Cue x Operation x Type interaction, $F(9,387)$, $p = 0.0071$, further indicate difficulty with Word - Number multiplication problems. The E choice seems to be chosen at slightly above chance over the S choice (Mean = .58) and the D choice picked at chance from the remaining three alternatives (Mean = .35). See Appendix II for figures of the CT, OT, and COT interactions.

Difficulty. There was no overall effect of the Difficulty of the fractions involved in the problems. Although this result corroborated results obtained with the NAEP, it seems somewhat surprizing. However, difficulty did interact with several variables. There was an interaction of Difficulty and Operation, $\underline{F}(3,129) = 6.71$, $\underline{p} = 0.0003$: there was no difference in performance between Hard and Easy problems for all operations except subtraction, $\underline{t}(47) = 4.14$, $\underline{p} = 0.0001$. The remaining four interactions suggest difficulty may have some effect on performance, but it is not consistent over Cue Operation, or Type: Difficulty x Operation x Type, $\underline{F}(9,387) = 2.19$, $\underline{p} = 0.0219$), Cue x Difficulty x Operation, $\underline{F}(3,129) = 7.93$, $\underline{p} = 0.0001$, Cue x Difficulty x Type, $\underline{F}(3,129) = 2.83$, $\underline{p} = 0.0411$, and Cue x Difficulty x Operation x Type, $\underline{F}(9,387) = 4.12$, $\underline{p} < 0.0001$, See Appendix II for figures).

## Discussion

The results of Experiment I are consistent with a Best Examples interpretation: nearly all subjects correctly chose the B alternative, which had a high degree of cue overlap with the standard (Note that it is possible that the slight difference between the number of expert and nonexpert subjects who chose the _E_ alternative is due to a response strategy employed on some items: for the best match, subjects may have chosen one of the two alternatives that matched in surface structure, _E_ or _S_, realizing that one of these was probably the correct response. The second choice was then made from the remaining two alternatives, _D_ and _N_. This strategy would also explain why subjects sometimes chose the _N_ response.). However, nonexperts chose the _D_ alternative much less often than the more expert subjects, and instead often chose the _S_ alternative, which was similar to the standard in the arbitrary details of the story. In other words, novices made fewer errors in deciding that two problems which both required the same operation for solution were similar if the story lines were the same. Errors were commonly made by choosing the alternative that was similar only in story line, not the essential characteristic of operation required for solution.

The results of the study also suggest that certain types of problems are more difficult to understand, as indicated by low rates of success in identifying similar problems and in determining the correct operation which should be used to solve the problem. Performance was best on addition problems, followed by subtraction, and then mutiplication and division problems. There was no evidence to suggest that either the difficulty of the numbers involved in a problem or the format of the problem presentation has any systematic effect on the subjects' ability to determine that two problems require the same operation for solution.

Section I provides some suggestion, and Section II stronger evidence, that there are differences in ability to comprehend fraction word problems which require the same operation. The error rates for these problems in which the result was unknown were different for dynamic problems and static problems. For the multiplication problems in Section II, not only were the error rates different, but the most common incorrect answer was different: 63% of the subjects would have subtracted to find the answer to the Dynamic multiplication problem, while 52% would have divided to answer the Static multiplication problem. Certainly, one might expect multiplication and division to be confused, given the similarity of the solution algorithms, but it is not obvious why one would call a multiplication problem a subtraction problem. To develop some explanation for this phenomenon, let us carefully consider the wording of these two problems.

First, the dynamic multiplication problem that was considered by many to be a subtraction problem was:

Margret had 2/5 of a gallon of ice cream. She gave

1/5 of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive?

Intuitively, the wording of this problem seems consistent with the idea of whole number subtraction: Margret had some ice cream and she gave some of it away. Yet, if whole numbers are substituted into this problem ( and the wording is changed to be consistent with whole numbers) the wording is not consistent with an interpretation as either a multiplication or a subtraction problem: "Margret had 2 gallons of ice cream. She gave 1 gallon of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive?". If Margret gave Anne Marie one gallon of ice cream, that is what Anne Marie received. There is no subtraction needed.

The static multiplication problem, commonly misclassified as division, was:

5/12 of the garden was planted with flowers. 2/5 of the flowers were tulips. What fraction of the garden was planted with tulips?

In this case, when whole numbers are substituted into the problem, and liberal changes are made to make the wording consistent with whole numbers, the problem does become a division problem. "5 parts of the garden were planted with flowers. 2 of the parts were planted with tulips. What fraction of the garden was planted with tulips?". Even with these liberal wording changes, the problem cannot be interpreted as a multiplication problem.

If the exercise of substituting whole numbers into fraction word problems is repeated for dynamic and static versions of other operations, the outcome is quite revealing. One can substitute whole numbers into all addition and subtraction problems, as well as dynamic division problems with only minor wording changes (Note: For example, a dynamic division problem with fractions: Grace had 3/4 of a pound of chocolate chips. She needed 1/4 of a pound of chocolate chips to make a batch of cookies. How many batches of cookies could Grace make? One could substitute 2 for 3/4 and 1 for 1/4 and the problem still sounds sensible as a division problem.). However, one cannot substitute whole numbers into static division problems or any multiplication problems. The reverse exercise, of substituting fractions into whole number problems yields similar results.

This exercise suggests that whole number and fraction problems may have different structures. Addition and subtraction problems with fractions may be understood in the same way that whole number problems are understood, so there is little difficulty in making a transition to using fractions. However, fraction multiplication and division problems may not be understood in the same way, and hence are more difficult to understand. Thus, there is an a priori basis for predicting that these three types of problems, dynamic multiplication, static multiplication, and static division, should be more difficult. The worst performance should occur on dynamic multiplication problems, since they actively suggest an alternative operation for solution.

These data suggest that there may be some structure inherent in word problems which is intermediate between the surface structure or story line and the deep structure or operation. That structure would be shared by all problems of the same operation which had the same action cue, i.e., either all had an action cue or all did not have an action cue. If this is true, then nonexperts should be more able to say that two problems require the same operation if they also share the presence or absence of an action cue than if they are mixed. A major purpose of Experiment II will be to investigate this intermediate level of problem structure.

# CHAPTER VI

## EXPERIMENT II

### Introduction

Experiment I provided support for an account of novice and expert problem solving behavior within a Best Examples framework. The results of the experiment also raised several new issues which will be investigated in Experiment II. These include: 1) whether it is possible to predict when the surface structure only alternative is likely to be chosen over the operation only alternative, 2) the possibility that patterns of errors can be better understood by postulating a middle level of structure which is between deep structure (operation) and surface structure (story line) and 3) whether difficulties in recognizing similarities among fraction problems are related to difficulties in interpreting corresponding whole number problems. A fourth aim of this experiment was to determine whether the results could be extended to a younger age population.

#### Surface Structure Errors

In previous chapters, it was argued that nonexpert subjects may weigh surface structure similarity too heavily when they must decide whether two problems are similar.

The high frequency with which the B alternative and the S alternative were chosen support this contention. Yet most nonexpert subjects did not seem to rely exclusively on surface structure similarity, choosing incorrect alternatives less often than would be expected by chance. Since the S alternative was chosen for some items, but not others, it is of interest to investigate the conditions that may lead to correct and incorrect answers. Section III of Experiment II was designed to test whether confusions made on the operations assessment task could predict the conditions under which errors would be made.

The majority of the subjects clearly do not regard similarity in surface structure as the most important indicator that two problems require the same operation. Hence, subjects must access and correctly interpret the deep structure cues at least some of the time. If the deep structure cues were not correctly accessed or interpreted, then the S alternative might appear to have more cue overlap with the standard. Thus, the Best Examples theory would predict that the subjects would fail to choose the D alternative, which shares fewer surface cues with the standard, when the operation needed to solve the standard cannot be easily identified by the subject. This information could be gained using the results of the Operations Assessment task, which was Section II of

Experiment I.

This argument can be extended somewhat further: even if the subject cannot correctly identify the operation needed to solve a problem on the operations assessment task, the subject may not choose an incorrect alternative. If standards requiring each operation were paired with alternatives requiring every other operation, certain confusions might be very likely and others unlikely. For example, if the subject had said one should multiply when given a division problem, one might expect the subject to choose the S alternative if it required multiplication for solution. In constrast, if the problem required addition, it should be less likely that the subject would choose the S alternative, because subjects rarely confuse addition and division problems.

In order to test whether errors made on the similarity judgment task are the result of an incomplete understanding of the operations involved, a task such as the operations assessment task may be used. In this task subjects must state what operation would be appropriate to solve a problem. The errors made on this task would allow for predictions concerning which specific pairs of standards and alternatives should lead to incorrect judgments of similarity. For this purpose, the data from Experiment I could have been used. However, the operations assessment task was repeated in Experiment II to ensure that the results were replicable with different problems. The similarity judgment task was modified for Experiment II. In Experiment I, the word problem standards were paired with alternatives that seemed likely to be confusable. In order to test the current hypothesis, word problems requiring each operation for solution were paired with sets of alternatives in which the S and M alternatives required every other operation for solution.

Middle Level Structure

The results of Experiment I indicated that subjects do not treat in the same way all problems which both require the same operation for solution and have the same open sentence description. In a set of items controlled for type of open sentence, there were differences in performance related to the presence or absence of an action cue. Such differences suggest subjects seem to be sensitive to a middle level of structure for fraction word problems that is intermediate between surface structure and deep structure.

What might this level of structure be like? The middle level structure of a problem is probably more related to the wording of a problem than the deep structure, which indicates what operation is appropriate, and is probably fairly abstract. For example, it is not immediately obvious how a dynamic and a static multiplication problem are alike, other than that they require the same operation for solution. However, two dynamic multiplication problems with the same open sentence description may share a common flow of action, have verbs with similar meanings, and similar question statements.

For the purposes of Experiment II, middle level structure will be defined as follows: problems with the same middle level structure require the same operation, have the same open sentence description, and have the same value on the dynamic-static dimension. The advantage of using middle level structure to aid in problem categorization is that it would allow one to make meaningful discriminations amongst problems with less effort than that required by a discrimination based on deep structure.

If this middle level of structure is useful in categorizing and solving word problems, then the following task should provide evidence for the existence of these structures. A subject is presented with a word problem standard and four alternative word problems. Each of the alternatives requires a different operation for solution. If the task is to choose the alternative which requires the same operation as the standard, it should be easier to choose this alternative if it has the same middle level structure as the standard. In other words, there should be facilitation in choosing the correct alternative if either both word problems are dynamic or both are static. A result of no difference between same comparisons and different comparisons would indicate that either there is no difference in the difficulty of making these comparisons or the subjects are not sensitive to middle level cues.

Whole Numbers Versus Fractions. There is one further way in which middle level structures might differ. The analysis based on the substitution of whole numbers into sensible fraction word problems and vice versa conducted in the last chapter indicates that middle level structures may be different for corresponding whole number and fraction word problems for at least some types of problems. Some substitutions are not sensible, indicating differences in the ways in which whole number and fraction multiplication and division problems are generally structured. If one considers only those problems in which the result is unknown, three factors may influence middle level structures: 1) the operation required to solve the problem, 2) the presence or absence of an action cue, and 3) the presence of whole numbers or fractions in the problem.

If the types of numbers in a word problem critically affect its structure, then performance on the operations assessment task should vary as a function of whether whole numbers or fractions are used. Differences should be greater for problems in which whole numbers cannot be substituted for fractions or vice versa (i.e. dynamic multiplication, static multiplication, and static division). Such differences would suggest that difficulty in understanding these three types of fraction problems is not due to a misunderstanding of the corresponding whole number problems, but is unique to the fraction problems. To test this aspect of middle level structure, the operations assessment task was expanded to include whole number as well as fraction word problems for each Operation-Action Cue combination.

A Younger Population. One of the more attractive aspects of the Best Examples theory is its potential to account for categorization behavior across a wide range of ages and domains. Theoretically, the same principles should be applicable to infants acquiring language and adults learning physics. One should find the same guiding general principles of categorization amongst nonexperts of all ages. Therefore, the results of these studies should be extendable to children who have recently learned how to solve fraction problems, and have less experience solving them. The overall levels of performance for adults and children may be different, but the patterns of data should be the same, i.e. nonexpert children should be sensitive to the same kinds of cues and make the same kinds of errors that are made by nonexpert adults. Accordingly, a sample of eighth grade students was included in Experiment II, in addition to an adult sample. Eighth graders were the youngest population available which had completed all instruction involving fractions at the time of the study.

## Method

### Subjects

Fifty-seven college students and 53 eighth graders participated in Experiment II. The adults were undergraduate students enrolled in pyschology classes at the University of Massachusetts who participated in exchange for course credit. There were 29 females and 28 males, with an average age of 20.0 years and average PPVT test stanine of 6.0.

The eighth graders were students at the Frontier Regional Junior High School in South Deerfield, Massachusetts. They were enrolled in the top three of five eighth grade mathematics classes, two of which were Algebra I classes and the other a standard eighth grade mathematics class. All the students were taught by the same mathematics teacher. The 20 males and 32 females who completed the study had an average age of 13 years, 6 months, and an average PPVT stanine of 6.4. In return for allowing the study to be conducted, the school received an honorarium of seventy-five dollars.

Materials

The problems were presented in a booklet composed of three subsections (See Appendix III). The first section contained the Match task; this task was described in the introduction to Experiment II and was intended to investigate the effects of middle level structure on the task of recognizing that two problems require the same operation for solution. The operations assessment task was contained in Section II. It was modified from Experiment I to include whole number as well as fraction problems. Section III was the similarity judgment task, modified to include pairings of standards requiring each operation with and incorrect alternatives requiring every other operation.

Section I. For the Match task in Section I, the subjects were given 16 items, each of which had a word problem standard and four word problem alternatives. The subjects were required to choose the alternative that was "the same kind of problem", with an example clearly indicating that the alternative should match the standard in operation. The alternatives were structured as follows: they all employed the same characters and had as similar a story line as possible. Each of the four alternatives required a different operation for solution, so there could be only one alternative which matched the standard in operation.

Four of the 16 standards required each operation. Half the standards and half the sets of alternatives had an action cues, while half did not, i.e. half were Dynamic and half Static. The standards were paired with the sets of alternatives so that half the items were matched for presence of action cue and half were not.

Section II. The operations assessment task for Section II in Experiment II was essentially the same as Section II in Experiment I: the subjects were given a set of word problems and were asked to state whether they would add, subtract, multiply, or divide to solve each problem. However in Experiment II, there were 16 problems rather than 8 problems, so as to include a corresponding set of whole number problems.

All of the problems were new, to ensure that the results of Experiment I were not due to the specific problems used. They included four problems requiring each operation, half of which had whole numbers and half fractions. Half the problems were Dynamic and half were Static.

Section III. Section III was a similarity jugment task with 24 items. Thus, a word problem standard was presented with four word problem alternatives: a B match, a D match, an S and an N match. The subject was told to choose the alternatives which went "the best" and "the next best" with the standard. It was made clear, as in Experiment I, that operation should be the basis for the similarity judgment.

For Experiment II, the standard and its alternatives matched on the dynamic-static dimension. There were three sets of items for each of the eight operation-action cue combinations: the distractor items for each of the three sets each required different operations for solution. Thus, standards requiring each operation were paired with alternatives requiring every other operation as alternatives.

Procedure

The instructions for each section in the experiment were provided in the written booklet immediately preceding the appropriate section (see Appendix III for copy of instructions and test items). For Section I, the subject was instructed to chose the single alternative that matched the standard in operation. In Section II, subjects were told to indicate the operation they would use with the two numbers given in the problem to solve that problem. In Section III, subjects were to choose the alternative that went the best and then the next best with the given standard. Examples clearly indicated that operation should be the basis for a similarity judgment.

The adults were run in small groups of no more than six subjects and were self-paced. There was a message at the end of Section II indicating it was time to take the PPVT. The PPVT was administered by a second experimenter in another room. Subjects then returned and finished Section III.

The children completed the experimental tasks in their mathematics classroom in two separate 45 minute sessions. In the first session, they did Sections I and II. Section III was completed a day or two later in the next class. The PPVT was administered individually by two experimenters during the two weeks following the administration of the experiment.

Design

The 16 problems in the Matching task of Section I were analyzed using an ANOVA design with two between-subjects variables (sex and age) and two within subjects factors (Match and Type). There were eight Operation-Action Cue combinations : dynamic addition, static addition, dynamic subtraction, static subtraction, dynamic multiplication, static multiplication, dynamic division, and static division (Note - Since the analysis of Experiment I indicated that subjects do not treat all problems requiring the same operation for solution in the same way, Type was the only distinction that was made was among eight types of problems.).

Section II, or the operations assessment task, was again analyzed independently of its use as an indicator of fraction problem solving ability. There were two between-subjects factors, sex and age, and two within-subjects factors, Number type (whole numbers versus fractions) and Type of problem (8 types).

The 24 problems in the similarity judgment task (Section III) were used in a correlational analysis. The errors made in Section II were used to predict the number of errors subjects would make in each pairing of standard and alternatives. Children and adults were analyzed separately and together.

## Results

### Scoring

For Section I, the alternative chosen for each problem was scored as correct or incorrect. A choice was regarded as correct if the alternative required the same operation for solution as the standard. For each problem, a correct choice was coded as a 1 and an incorrect choice as a 0.

In Section II, the operations assessment task, responses were coded in a similar manner. In addition, the incorrect choices were tabulated, resulting in a confusion matrix that was used to predict the responses for Section III. A measure of expertise was also computed as in Experiment I, using only the responses to the eight fraction problems. Unfortunately, this data did not allow for an ANOVA with Error Level as a factor, as in Experiment I, because the sizes of the error level groupings were too disparate. However, several t tests comparing novices and experts were performed in an attempt to determine if the results replicated Experiment I. In the present experiment, there were 15 relatively expert subjects who made 0 or 1 error on Section II, and 94 nonexperts who made 2 or more errors. Only four of the experts were eighth graders, which did not allow for separate tests of adults and eighth graders.

Responses to Section III, the similarity judgment task, were scored on the basis of the correctness of the pairs of alternatives chosen. The subject was considered to have responded correctly if the two choices were the B alternative and the D alternative.

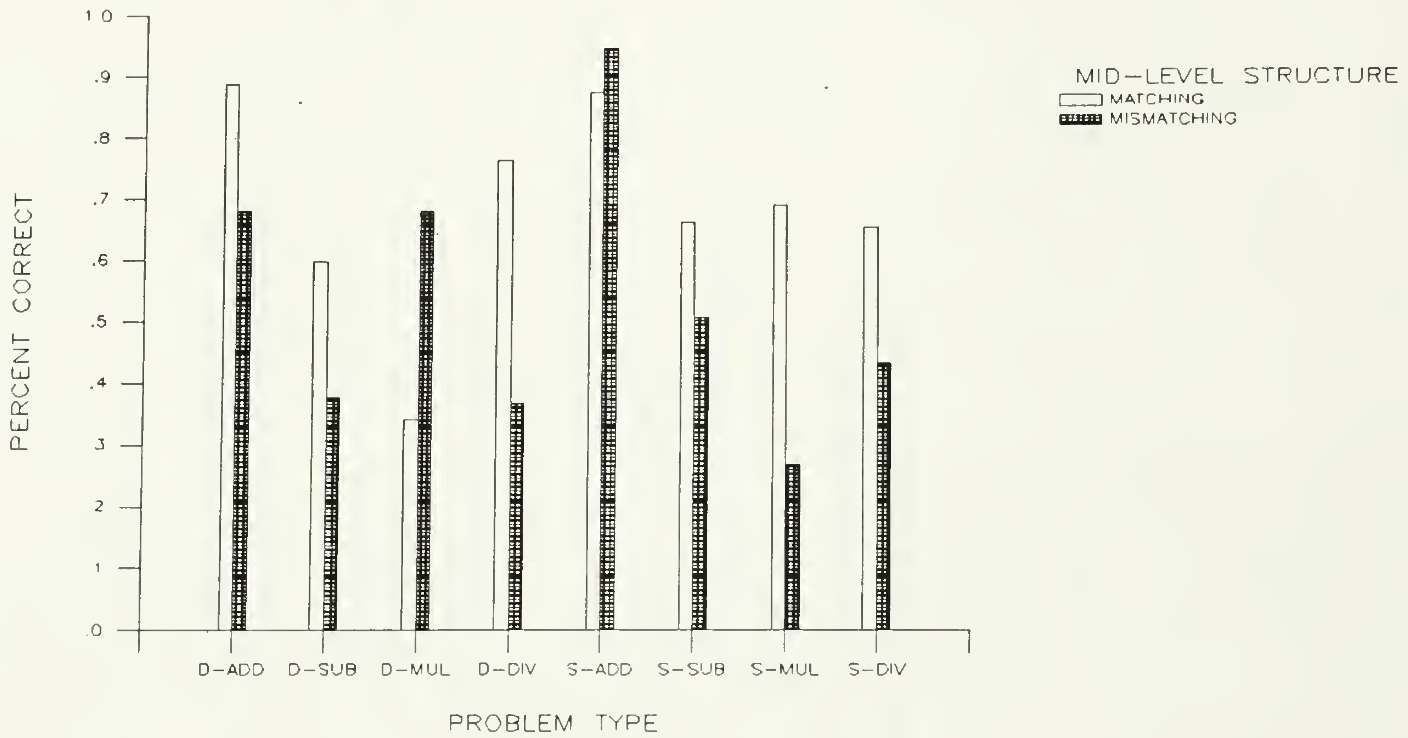
Section I

Section I was developed to determine whether shared middle level structure facilitates the judgment that two problems require the same operation for solution. A 2 (Sex) x 2 (Age group) x 2 (Match - Matching or mismatching in middle level structure) x 8 (Type) ANOVA was performed. There were no main effects of either sex or age: adults were no better than eighth graders on this task. The effects of Match and Type will now be discussed in detail.

The main effect of Match was highly significant, $F(1,105) = 29.09$, $p < 0.0001$. Subjects chose the alternative that required the same operation as the standard more often when the middle level structures were the same (See Figure 6.1). Thus, the data from Section I support the hypothesis that subjects are sensitive to middle level structure and that it is helpful in deciding that two word problems require the same operation for solution.

The main effect of Type was also quite significant, $F(7,735) = 29.09$, $p < 0.0001$, as expected, given the strong effects of Operation and presence of Action Cue observed in

FIGURE 6.1 – SECTION I: PROBLEM TYPE VS MID-LEV STRUCTURE



Experiment I. However, the pattern of errors was somewhat different from that observed in Experiment I (See Figure 6.1). Recalling the results of Section II in Experiment I, the operations needed to solve the dynamic multiplication, static multiplication, and static division problems were the most difficult to identify. In the current experiment, performance on items with dynamic addition and static addition standards was better than for the remaining types of items combined, $\underline{t}(108) = 11.48$, $\underline{p} < 0.0001$. Static addition problems were responded to correctly more often than dynamic addition problems, $\underline{t}(108) = -4.23$, $\underline{p} < 0.0001$. There were no significant differences among the six remaining types of standards.

Match and Type also interacted, $\underline{F}(7,735) = 16.63$, $\underline{p} < 0.0001$. The effect of a match in middle level structure was positive for all types of problems except two: performance on the static addition item was slightly better when the alternatives did not match in middle level structure, $\underline{t} = -2.17$, $\underline{p} = 0.0319$, while performance on the dynamic multiplication problems was significantly better in the mismatching condition, $\underline{t} = -5.19$, $\underline{p} < 0.0001$. An examination of Figure 6.1 indicates that performance on the former two problems was quite high, and any difference between them may be due to chance.

A possible explanation for the latter reversal might be found by examining the pattern of errors shown in Table 6.1: when the dynamic multiplication problem was paired with other dynamic problems, many subjects incorrectly chose the alternative that required subtraction. As noted in the last chapter, both of these types of problems involve the removal of some portion of a set, and hence it is likely that subjects would consider both problems to be subtraction problems. When the dynamic multiplication problem was paired with the static alternatives, there were no other problems that involved removal. Thus, the subject would be less likely to think the multiplication and subtraction problems required the same operation for solution.

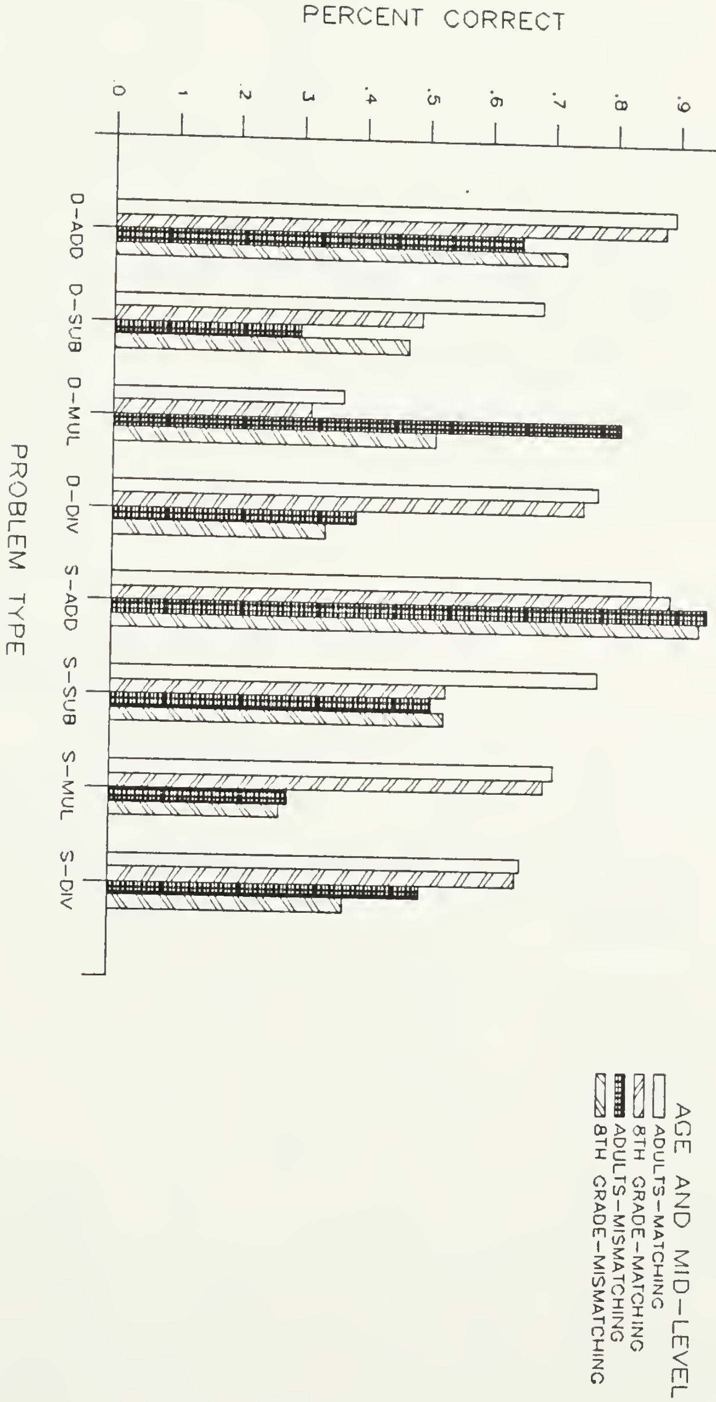
Match by Type also interacted with Age, $F(7,735) = 2.73$, $p = 0.0083$. As can be seen in Figure 6.2, the patterns of facilitation differed slightly for adults and eighth graders. The most notable difference was on the dynamic multiplication problem in the mismatching condition, in which adults chose the correct alternative more often than eighth graders. Thirteen of the eighth graders chose the subtraction alternative in this case, versus 3 adults, indicating the strength of the perception among the eighth graders that this type of problem is a subtraction problem.

TABLE 6.1 -Section 1: Responses of Adults and Eighth Graders

| Problem Type | Response | | | |
|---|---|---|---|---|
| | Addition | Subtraction | Multiplication | Division |
| Matching | | | | |
| Addition | | | | |
| Dynamic | 51!46 | 2!5 | 3!0 | 1!1 |
| Static | 49!46 | 1!0 | 5!1 | 2!5 |
| Subtraction | | | | |
| Dynamic | 3!5 | 39!26 | 11!16 | 4!5 |
| Static | 3!3 | 44!28 | 9!19 | 1!2 |
| Multiplication | | | | |
| Dynamic | 1!3 | 23!26 | 21!16 | 12!7 |
| Static | 9!8 | 1!3 | 40!35 | 7!6 |
| Division | | | | |
| Dynamic | 4!3 | 7!5 | 2!25 | 44!18 |
| Static | 5!4 | 4!8 | 11!6 | 37!34 |
| Mismatching | | | | |
| Addition | | | | |
| Dynamic | 37!37 | 8!8 | 5!4 | 7!3 |
| Static | 54!49 | 0!1 | 2!1 | 1!1 |
| Subtraction | | | | |
| Dynamic | 6!8 | 17!24 | 27!14 | 7!6 |
| Static | 3!4 | 29!26 | 16!19 | 9!3 |
| Multiplication | | | | |
| Dynamic | 5!2 | 3!13 | 46!28 | 3!7 |
| Static | 9!8 | 1!3 | 40!35 | 7!6 |
| Division | | | | |
| Dynamic | 0!4 | 8!5 | 27!25 | 22!18 |
| Static | 12!11 | 10!11 | 7!11 | 28!19 |

Note - Adults!Eighth graders

FIGURE 6.2 — SECTION I: AGE VS MID-LEV VS PROBLEM TYPE



The performance of experts and novices in Section I was compared in three ways: 1) responses on all items combined, 2) responses on only the items that matched in middle level structure, and 3) responses on items that mismatched in middle level structure. There were marginal differences between experts and novices on all problems combined, $t(20) = 2.38$, $p = 0.0274$, and on the matching middle level structure items, $t(25) = 2.31$, $p = .0291$ (Note, $p < 0.0167$ is needed for significance with a Bonferrroni $t$ test and three comparisons). Experts may be slightly more sensitive to middle level structure cues than novices.

Section II

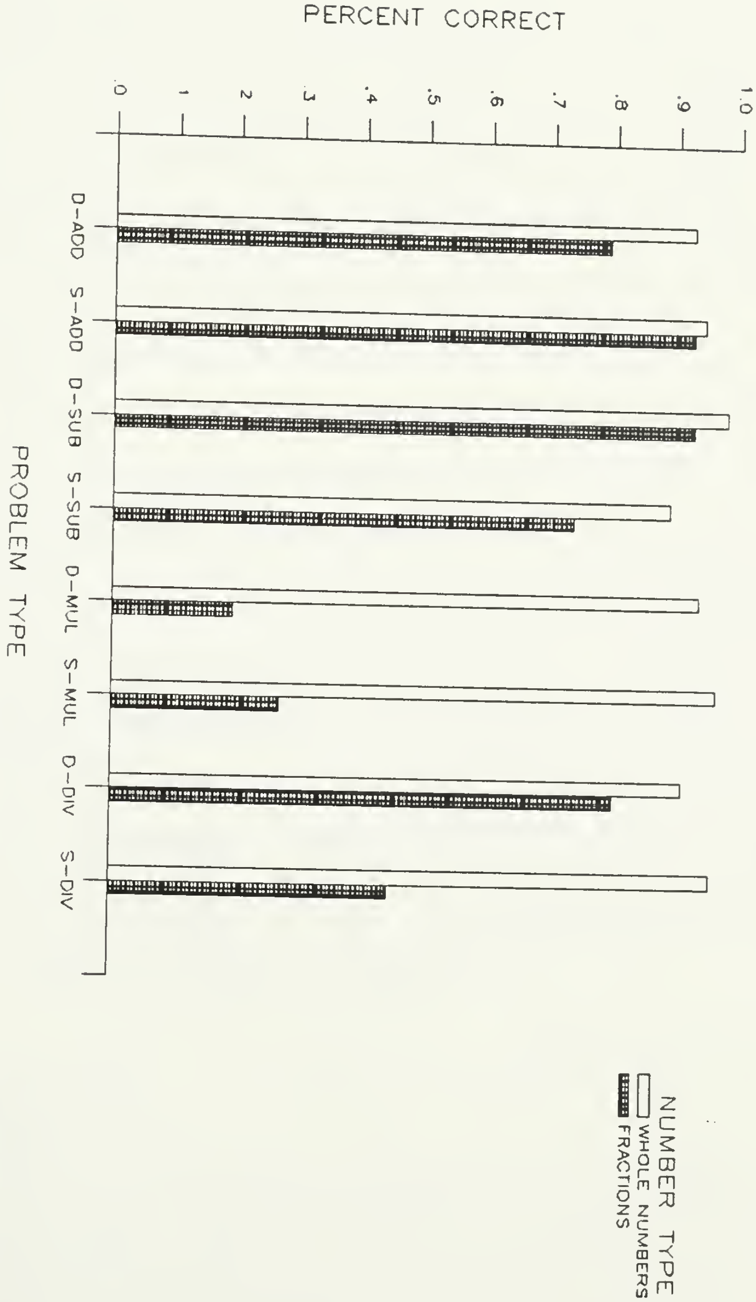
Section II was the operation assessment task of Experiment I. It was used in the current experiment for three purposes: 1) to provide an indication of expertise, 2) to replicate the results of Section II in Experiment I, and 3) to compare responses to problems with whole numbers to problems with fractions. A 2 (Sex) x 2 (Age group) x 2 (Number type - whole number or fraction) x 8 (Type) ANOVA revealed no main effect of sex and no interactions involving sex. The main effect of age was significant, $F(1,105) = 20.98$, $p < 0.0001$; adults chose the correct operation for solution more often than eighth graders. There were no significant interactions with age. Thus, although adults responded correctly more often than eighth graders, the patterns of response were the same, as predicted by the Best Examples theory.

There was a large main effect of Number type, $F(1,105) = 440.98$, $p < 0.0001$; problems with whole numbers were responded to correctly more often than problems with fractions. An examination of Figure 6.3 shows that the average rates of response to problems with whole numbers ranged from 89 to 98 percent correct. In contrast, the average solution rates for fraction problems ranged from 19 to 93 percent correct.

The main effect of Type was again significant, $F(7,735) = 45.08$, $p < 0.0001$. The patterns of responses replicated Section II of Experiment I: performance on the dynamic multiplication, static multiplication, and static division problems combined was significantly worse than performance on all other types of problems, $t(108) = 13.48$, $p < 0.0001$.

There was a significant interaction between Type and Number, $F(7,735) = 52.64$, $p < 0.0001$, which was predicted by the number substitution exercise conducted in the last chapter. Performance on all whole number problems was better than performance on corresponding fraction problems (See Figure 6.3), but the negative effect of fractions was more pronounced for certain types of problems. An

FIGURE 6.3 – SECTION II: NUMBER TYPE VS PROBLEM TYPE



examination of Figure 6.3 shows relatively small, but significant, differences in performance on dynamic addition, $t(108) = 3.26$, $p = 0.0015$, and static subtraction problems, $t(108) = 3.75$, $p = 0.0003$. Larger differences in performance were observed on the problems where performance was predicted to be lower by the number substitution exercise: dynamic multiplication, $t(108) = 17.68$, $p < 0.0001$, static multiplication, $t(108) = 15.13$, $p < 0.0001$, and static division, $t(108) = 9.97$, $p < 0.0001$ (Note, $p < 0.0063$ is needed for significance).

The performance of experts and novices was compared on the whole number problems in Section II. The mean performance for both groups was quite high: the mean for experts was 7.87 (98%) correct of 8 problems, while the mean for novices was 7.45 (.93). Experts were able to identify the operation needed to solve whole number problems signficantly more often than novices, $t(76) = 2.71$, $p = 0.0082$. Performance was not compared on the fraction problems since the responses to these problems were used to define expertise.

<u>Section III</u>

Section III was designed to investigate whether one could predict the situations in which subjects would be likely to err on the similarity judgment task. If a subject stated that division should be used to solve a problem that required multiplication, then that subject might possibly err in the similarity judgment task when a multiplication problem was presented with an S alternative that required division.

Accordingly, the performance of the group on Section II was used to predict group performance on Section III (See Table 6.2 for the confusion matrix). Performance across all subjects on Section II correlated with performance on Section III, $r = 0.389$, $t(107) = 4.36$, $p < 0.001$. For eighth graders, this correlation was significant, $r = 0.571$, $t(51) = 4.96$, $p < 0.001$, while for adults it was not, $r = 0.156$, $t(55) = 1.17$, $p > 0.10$. The correlation was significantly larger for eigth graders than adults, $Z = 2.49$, $p < .02$. Adults and eighth graders may have developed different strategies for solving these problems which contributed to the magnitude of the difference.

For eighth graders, the results support the hypothesis that subjects will tend to err on those items that present potential confounds of the type that subjects have confused in naming an operation, i.e. performance on the similarity judgment task is predicted by performance on the operations assessment task.

The performance of experts and novices on Section III essentially replicated the results of Experiment I: although there was a difference between novices and

TABLE 6.2 - Section II: Operations Assessment Task

| | Response | | | |
|---|---|---|---|---|
| | Add | Subtract | Multiply | Divide |
| **Whole Numbers** | | | | |
| Addition | | | | |
| Dynamic | 54!47 | 3!5 | 0!0 | 0!0 |
| Static | 56!47 | 0!2 | 1!3 | 0!0 |
| Subtraction | | | | |
| Dynamic | 0!1 | 57!50 | 0!0 | 0!1 |
| Static | 1!4 | 55!42 | 0!2 | 1!4 |
| Multiplication | | | | |
| Dynamic | 0!3 | 0!2 | 57!45 | 0!2 |
| Static | 0!3 | 0!1 | 57!48 | 0!0 |
| Division | | | | |
| Dynamic | 0!1 | 1!2 | 2!4 | 54!45 |
| Static | 0!2 | 0!3 | 0!0 | 57!47 |
| **Fractions** | | | | |
| Addition | | | | |
| Dynamic | 51!35 | 2!9 | 3!2 | 1!6 |
| Static | 54!47 | 1!1 | 2!0 | 0!4 |
| Subtraction | | | | |
| Dynamic | 1!2 | 53!48 | 3!0 | 0!2 |
| Static | 1!3 | 45!35 | 5!5 | 6!9 |
| Multiplication | | | | |
| Dynamic | 0!1 | 23!29 | 18!3 | 16!9 |
| Static | 1!4 | 9!20 | 22!7 | 25!21 |
| Division | | | | |
| Dynamic | 1!4 | 2!3 | 7!5 | 47!40 |
| Static | 5!4 | 6!3 | 20!23 | 26!22 |

Note: Adults!Eighth graders

experts, $t(104) = 4.01$, $p = 0.0001$, the rate of choosing the B response was quite high for both groups, Mean(Experts) = 23.8 ( of 24), Mean(Nonexperts) = 22.3). The mean difference in performance was slightly larger for the D response, Mean(Experts) = 21.9, Mean(Nonexperts) = 19.8, $t(31) = 2.80$, $p = 0.0086$.

## Discussion

The results of Experiment II are readily interpretable in terms of a Best Examples framework, as were the results of Experiment I. In addition, the experiment helps to further define the sources of information that are used in categorizing problems that require the same operation for solution.

First, subjects are more often correct in judging that two problems require the same operation for solution if they both have the same middle level structure, i.e. match in both the operation required to solve the problem and in the presence of action cues. Thus, similarity of middle level structures must provide some recognizable cue overlap that aids in making a valid judgment of similarity. It remains to be determined exactly what these middle level structure cues are.

Middle level structure is presumably more closely related to the surface wording of the problem than is deep structure, but does not include such specific characteristics as story line or actors. Problems which have the same middle level structure may have a common flow of action, such as "giving-to" or "-from" action (see Kintsch and Greeno, 1985), similar formulation of questions, and use some of the same words, such as "give" or "altogether."

It should be noted that common words alone do not seem sufficient to predict the same operation should be used to solve two problems. The question sentences for three of the four alternatives for item 16 all began with "How much", yet 44 of the 57 adults responded correctly. Many of the problems used in Experiment II that matched in middle level structure did not share any key words. Therefore, it is unlikely that common words alone could account for the results attained in Experiment II.

Second, as suggested by the Best Examples theory, patterns of responses for younger subjects were similar to those of adults. Performance of eighth graders and adults was quite similar on Section I, differing only in the magnitude of the response decrement for a few problems in the mismatching middle level structure condition. On Section II, adults were able to identify the operation needed to solve a problem more often, but there was no difference in the patterns of responses. For both Sections I and II, the common incorrect responses appear similar as well.

Third, the pattern of differences between nonexperts and experts was replicated in Section III. Both experts and nonexperts chose the B alternative for nearly every problem. Larger differences between novices and experts were again observed in the choice of the D alternative. Thus, for nonexperts, the common surface structure of the standard and the B alternative has a facilitory effect beyond the common middle and deep structures possessed by the standard and the D alternatives.

Together, the three sections of the experiment indicate three levels of structure which appear to be used by subjects in categorizing problems: 1) the deep structure, which contains abstract information indicating what operation should be used to solve to problem , 2) the middle level structure, which indicates how the quantities in the problem are related, and 3) the surface structure, which contains the detailed information about the characters and context. Casual inspection of the data suggests there is a large advantage gained in categorization performance with common middle level structure, and a smaller advantage gained with common story lines. The relative advantages of these cues should be tested within a single paradigm for two reasons.

First, the results of Experiment II indicate that common middle level structure seems to provide a substantial advantage over operation alone. If this advantage is larger than any additional advantage gained by common surface structure, it would support an argument that the middle level of structure provides a more fundamental basis for categorization than either the operation alone or the surface structure.

Rosch et.al.(1976) have referred to this most fundamental level as the "basic level". According to Mervis (1980, pg. 292), the basic level is "the most general level at which categories are formed according to large naturally occurring attribute clusters. Categories at this level are more differentiated from each other than are categories at any other level." As discussed previously, problems that have a common middle level structure seem to have many common attributes. Problems that only require the same operation for solution seem much more different from each other than do problems that only differ in surface structure. Thus, there may be reason to believe that the categorization of arithmetic problems is like the categorization of objects in a way other than cue overlap, i.e. there may be an optimal level of categorization.

Second, a strong advantage of middle level structure over deep structure would have important educational implications. Given that most children have difficulty making categorization decisions based on operation alone, such results would suggest children should be taught to make discriminations among problems based on the middle level of structure rather than the more abstract level of operation.

Finally, the results of Section II provide convincing evidence that the middle level structures for comparable types of whole number and fraction problems are not the same, for at least the operations of multiplication and static division: 93%, 96%, and 95% of the subjects classified the whole number dynamic multiplication, static multiplication, and static division problems, respectively. However, only 19%, 26%, and 44% of the subjects correctly categorized the corresponding fraction problems. This evidence corroborates the number substitution exercise and indicates that the difficulty of recognizing fraction word problems is not strictly due to a misunderstanding of whole number multiplication. Rather, it may be difficult to appropriately generalize the concept of multiplication so that it includes fractional multiplication. Such results suggest that a teaching strategy based more directly on the utilization of common middle level structure may produce more positive results.

# CHAPTER VII

## GENERAL DISCUSSION

The intent of the present research was to develop a definition of expertise that potentially could apply to a broad range of domains, as well as be able to account for the development of a novice into an expert. A guiding assumption of this research was that nonexperts, as well as experts, have consistent bases for problem categorization. Nonexperts are probably different from experts in at least one crucial way, i.e. they rely on different kinds of cues to inform them that two problems are similar.

In addition to these theoretical concerns, these experiments were conducted with an eye toward future applications to education. By concentrating on the fairly circumscribed domain of fraction word problems, it has been possible to gain specific information concerning the relative difficulty of different types of problems which should be immediately useful to educators. The results of the experiments also have much potential to help develop more efficient methods of teaching students to solve problems with fractions. Conclusions relevant to both theoretical and practical concerns will be discussed in this chapter.

## The Best Examples Theory

Although novices and experts may differ in many ways, it has been argued in this dissertation that one of the primary ways in which they differ is in their organization of concepts. Thus, one would expect to find differences in the categorization of problems by experts and novices. The Best Examples theory was chosen as a framework in which expert and novice categorization could be interpreted. The Best Examples theory has been developed to account for the categorization of naturally occurring objects, as opposed to the abstractly defined artificial categories that were studied previously.

Although the domains to which the Best Examples theory had been previously applied were limited to object categories, the principles are potentially applicable to abstract categories as well. The theory has the desirable quality of being able to account for the seemingly disparate categorization behaviors of young children and adults.

Thus, the principles of the Best Examples theory were extended to include the more abstract categories of arithmetic operations. According to the Best Examples formulation, nonexperts may both underextend and overextend arithmetic categories because they rely on nonessental characteristics in the wording of the problem for categorization.

The present research supports this extension of the Best Examples theory in several ways. It has demonstrated that 1) surface structure overlap is important to nonexperts in categorizing problems, 2) younger nonexperts and older nonexperts respond similarly, and 3) the results using this approach are replicable. In addition, this approach has made it possible to show that not all problems which require the same operation for solution are equally difficult. In addition, it has provided an indication of which types of problems may be more difficult to understand and solve.

On the similarity judgment task, all subjects in Experiments I and II tended to choose the alternative that had the highest degree of cue overlap, i.e., matched in both story line and operation. Nonexperts differed from experts by often choosing alternatives which matched the standard in story line alone. These results are typical of those commonly obtained within the Best Examples framework: children learning language tend to agree with adults on the central members of a category, which have a high degree of cue overlap, but disagree when the exemplar shares features with a contrast category (Mervis and Rosch, 1981).

The surface structure overlap of problems that required different operations for solution had a stronger negative effect on the categorization of some problems than others. The results of Experiment II suggest that subjects often choose incorrect alternatives when they are unable to identify what operation should be used to solve a problem of the same type as the standard. These data do not provide information that would distinguish whether subjects actually believe an incorrect alternative requires the same operation for solution or whether they consider themselves to be guessing.

The patterns of correct answers, as well as the types of incorrect responses were quite similar for younger and older nonexpert subjects. The performances of eighth graders and adults were quantitatively, but not qualitatively different. In an experiment comparing the categorization of five- and eight-year olds with adults, Nelson (1974) found that both groups of children responded similarly by including items adults failed to include and excluding items included by adults. Thus, in both cases, nonexpert subjects of different ages respond in a consistent, if incorrect, manner.

In accordance with this point, the results of both experiments appear to be quite stable. The results of Experiment II were consistent with those of Experiment I on both the similarity judgment task and the operations assessment task. In addition, the results of the eighth graders were similar to those of adults. The stability of these results indicates this approach may be fruitful for investigating categorization in many other abstract domains.

## Levels of Structure

The present results suggest problem solvers are sensitive to at least three levels of structure in fraction word problems: 1) surface structure, which includes the characters in the problem, the objects, and the storyline, 2) middle level structure, which includes the order in which information is presented, the position of the unknown in the problem, the pattern of actions, and the relationship among problem elements, and 3) deep structure, or the operation needed to solve the problem. Related research (Nesher, 1982) with children solving whole number addition and subtraction problems also suggests a similar taxonomy of structure.

Nesher (1982) has made distinctions among the syntactic level, the semantic level, and the logical level of word problems. The syntactic component is most dissimilar to the present characterization of levels, and includes variables such as number of sentences, location of the question, and number of words. Nesher's (1982) review indicates these variables have relatively little effect on problem solving.

The semantic component is comparable to what has been termed "middle level structure" in this paper. It includes dependencies due to relational terms, such as "more," as well as general type of text. Nesher (1982) distinguishes three types of addition and subtraction texts: 1) dynamic texts, in which the relevant information in the text is embedded in a time sequence (the present definition of dynamic is the same as Nesher's definition), 2) comparison texts, which ask questions concerning the relations between quantities, and 3) static texts, which involve no relevant actions or comparisons (In the present experiment, "static" is a combination of Nesher's static and comparison categories).

The logical structure of a problem refers to the operation needed to solve the problem. Nesher (1982) argues that all addition and subtraction problems must fulfill certain logical conditions. They must consist of at least three strings: two strings with an information component and one string with a question component. For addition problems, the information should indicate that two sets of objects exist and are disjoint sets; the question refers to the union of the two sets. For subtraction problems, the information component must refer to the union of the two sets and the question to one of the subsets.

Nesher (1982) has found that when the logical structure of a problem is held constant, the semantic structure has an important influence on how easily a problem is solved.

The results of Experiment II are in close accord with Nesher's (1982) finding. Subjects were clearly responsive to middle level cues and demonstrated this by recognizing that two problems required the same operation for solution more often when the middle level structures matched. Given the evidence that at least three levels of structure are important for arithmetic word problems, the issue of whether any of these levels is more fundamental than the others becomes important.

Research within the Best Examples framework has shown that objects are commonly labeled in a way that reflects one particular level of categorization. This level has been termed the "basic" level (Rosch et. al., 1976) and allows subjects to make important distinctions between objects with as little cognitive effort as possible. Finer distinctions may be made on the subordinate level, while broader classifications of objects are made on the superordinate level. For example, a dog (basic level) is an animal (superordinate level). My dog (specific exemplar) is an Australian shepard (subordinate level).

The results of the current experiments suggest one level of categorization may be more fundamental than the others for fraction word problems, and that level is the middle level. The operations assessment task provided strong evidence that subjects do not view all problems that require the same operation in the same way. Thus, the level of deep structure cannot be the basic level, even though this is the level at which children are taught problem solving in school.

The data from Experiment II suggest that the advantage of matching middle structures over matching deep structures may be somewhat greater than the added advantage of matching surface structure, although this hypothesis must be tested within a single experimental paradigm. One might test this hypothesis by adding another condition to the Match task in which for half the problems the story line of the standard would be as close as possible to the story line of the alternatives, while for the other half of the problems the standard would have a clearly different story line. If this hypothesis were true, it would predict that the advantage in performance of matching middle level structure and operation over matching operation only should be larger than the advantage of matching story line in addition to middle level structure and operation versus the middle level structure and operation match. This result would support the argument that the middle level of structure is more basic than either the deep structure or the surface structure.

One could also demonstrate that the middle level of structure is more fundamental in another way: one could teach a group of subjects to make discriminations between problems based upon middle structure only. In other words, the subjects would be taught to recognize dynamic and static multiplication problems, noting that these distinctions would not cause any loss of information concerning the operation needed for solution. This group of subjects should be more capable of correctly sorting problems according to operation than a control group that has not learned middle level structure distinctions, since the exemplars in middle level structure categories have more obvious similarities.

A similar approach has been advocated by Mayer (1981) for algebra word problem solving. Mayer (1981) found many high school students have considerable difficulty translating algebra word problems to meaningful equations, and in addition, have poor memories of these problems. He suggested that one way of making initial learning more efficient might be to provide students with explicit instructions and practice in recognizing problem templates. If, one views middle level structures as problem templates as Kintsch and Greeno (1985) do, this approach may be equally valid for fraction word problems.

## Current Educational Implications

Thus far in this chapter, several possibilities have been raised for further research based upon other aspects of the Best Examples approach, i.e. levels of structure, which may be of future interest to educators. However, the results of the present study by itself have important implications for education.

Probably the single most dramatic finding of present research is the extremely poor performance of eighth graders on the operations assessment task in identifying multiplication problems as multiplication problems. Only 3 of 48 students said they would mutiply the two numbers in the problem to solve the dynamic multiplication problem, while only 7 said they would multiply to solve the static multiplication problem. An informal survey of junior high school mathematics teachers suggests most teachers are not aware of the extent to which students do not understand multiplication problems. However, the results of several studies, including the NAEP, the RNP, and the present study, imply that most students do not have well defined notions about the kinds of situations that require the multiplication of fractions.

The static division problems also were reasonably difficult for the students, as the teachers recognized.

However, a much higher proportion (22 of 48) of students recognized the static division problem as a division problem on the operations assessment task. For this problem, most of the wrong answers are multiplication (23 of 48), suggesting that most students at least recognized this as a multiplicative rather than an additive problem. This was not the case for the multiplication problems, where many students responded with subtraction rather than division.

These results suggest that word problems should be made more central to mathematics education, particularly in the junior high school. Even though all students may not be able to solve all simple computational problems, it is clear that students need to develop better ideas of the kinds of situations that require each operation. The ability to recognize a problem requiring multiplication of fractions and perform that operation is critical for the learning of many types of mathematics, including algebra, calculus, probability, and business mathematics. One might even argue that students would benefit from the process of categorizing, rather than solving, word problems. In fact, one might investigate whether this were so by providing students with different numbers of problems to categorize and then comparing performance on the operations assessment task.

## Summary

The Best Examples theory has proven to be a valuable tool for the investigation of the categorization of abstract concepts, such as arithmetic operations. The data are reliable, being consistent within the present study and replicating prior results with object categories.

This approach has shown that nonexperts are systematic in their categorization of problems, but do not rely on appropriate cues in all cases. It was suggested that problem categorization might be improved if subjects were: 1) taught to use middle level structure cues to categorize problems, and/or 2) given considerable practice and feedback in categorizing single step fraction word problems.

# REFERENCES


Adelson, B. (1981) Problem solving and the development of abstract categories in programming languages. Memory and Cognition, 9, 422-433.

Anderson, J.R. (1982) Acquisition of a cognitive skill. Psychological Review, 85, 355-362.

Anderson, J.R., Greeno, J.G., Kline, P.J., and Neves, D.M. (1981) Acquisition of a problem solving skill. In J.R. Anderson (Ed.) Cognitive Skills and their Acquisition, Hillsdale, N.J.: Erlbaum, 191-230.

Anzai, Y. and Simon, H.A. (1979) The theory of learning by doing. Psychological Review, 86, 124-140.

Behr, M.J., Lesh, R., Post, T.R., and Silver, E.A. (1983) Rational number concepts. In Lesh and Landau (Eds.) Acquisition of Mathematics Concepts and Processes. New York: Academic Press, 91-126.

Bhaskar, R. and Simon, H.A. (1977) Problem solving in semantically rich domains: An example from engineering thermodynamics. Cognitive Science, 1, 193-215.

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

Brown, J.S. and Burton, R.R. (1978) Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, *2*, 155-192.

Bourne, L.E. (1968) *Human Conceptual Behavior*, Boston: Allyn & Bacon.

Brown, J.S. and VanLehn, K. (1980) Repair Theory: A generative theory of bugs in procedural skills. *Cognitive Science*, *4*, 379-426.

Bowerman, M. (1980) The structure and origin of semantic categories in the language-learning child. In M.L. Foster and S. Brandes *Symbol as Sense*, New York: Academic Press, 277-299.

Carpenter, T.P., Corbitt, M.K., Kepner, Jr., H.S., Lindquist, M.M., and Reys, R.E. (1981) *Results from the Second Mathematics Assessment of the National Assessment of Educational Progress*, Reston, VA: National Council of Teachers of Mathematics.

Carpenter, T.P. and Moser, J.M. (1982) The development of addition and subtraction problem-solving skills. In T.P. Carpenter, J.M. Moser, and T.A. Romberg (Eds.) *Addition and Subtraction: A Cognitive Perspective*, Hillsdale, N.J.: Erlbaum.

Chase, W.G. and Simon, H.A. (1973) Perseption in chess. *Cognitive Psychology*, *4*, 55-81.

Chi, M.T.H., Feltovich, P.J., and Glaser, R. (1981) Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Clark, E.V. (1973) What's in a word? On the child's acquisition of semantics in his first language. In T. Moore (Ed.) Cognitive Development and the Acquisition of Language, New York: Academic Press, 65-110.

Clement, J. (1981) Analogy generation in scientific problem solving. Proceedings of the Third Annual Meeting of the Cognitive Science Society, University of California at Berkeley, 137-140.

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

Hays, J.R. (1981) The Complete Problem Solver, Philadelphia, PA.: The Franklin Institute Press.

Hinsley, D.A., Hayes, J.R., and Simon, H.A. (1977) From words to equations: Meaning and representation in algebra word problems. In P. Carpenter and M. Just (Eds.) Cognitive Processes in Comprehension, Hillsdale, N.J.: Erlbaum.

Heider, E.R. (1973) Natural categories. Cognitive Psychology, 4, 328-350.

Hunting, R.P. (1984) Understanding equivalent fractions. *Journal of Science and Mathematics Education in South East Asia*, *VII*, 26-33.

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

Langley, P. and Simon, H.A. (1981) The central role of learning in cognition. In J.R. Anderson (Ed.) *Cognitive Skills and Their Acquisition*, Hillsdale, N.J.: Erlbaum, 361-380.

Larkin, J.H., McDermott, J., Simon, D.P., and Simon, H.A. (1980) Models of competence in solving physics problems. *Cognitive Science*, *4*, 317-345.

Lesh, R., Landau, M., and Hamilton, E. (1983) Conceptual models and applied mathematical problem solving research. In Lesh and Landau (Eds.) *Acquisition of Mathematics Concepts and Processes*, Academic Press, 263-343.

Lewis, M.W. and Anderson, J.R. (1985) Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology*, *17*, 26-65.

Mayer, R.E. (1981) Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science*, *10*, 135 - 175.

McKeithen, K.B., Reitman, J.S., Rueter, H.H., and Hirtle, S.C. (1981) Knowledge organization and skill differences in computer programmers. Cognitive Psychology, 13, 307-325.

Mervis, C.B. (1980) Category structure and the development of categorization. In R.J. Spriro, B.C. Bruce, and W.F. Brewer (Eds.). Theoretical Issues is Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education. Hillsdale, N.J.: Lawrence Erlbaum, 279-307.

Mervis, C.B., and Crisafi, M.A. (1982) Order of acquisition of subordinate-, basic-, and superordinate-level categories. Child Develeopment, 53, 258-266.

Mervis, C.B. and Mervis, C.A. (1982) Leopards are kitty-cats: Object labeling by mothers for their thirteen-month-olds. Child Development, 53, 267-273.

Mervis, C.B. and Rosch, E. (1981) Categorization of natural objects. Annual Review of Psychology, 32, 89-115.

Michener, E.R. (1978) Understanding understanding mathematics, Cognitive Science, 2, 361-383.

Nesher, P. (1982) Levels of description in the analysis of addtion and subtraction word problems. In T. P. Carpenter, J.M. Moser, and T.A. Romberg (Eds) Addition and Subtraction: A Cognitive Perspective, Hillsdale, N.J.: Erlbaum, 25-38.

Newell, A., Shaw, J.C., and Simon, H.A. (1960) A variety of intelligent learning in a general problem solver. In M.C. Youvits and S. Cameron (Eds.) Self-organizing Systems: Proceedings of an Interdisciplinary Conference, New York: Pergamon Press.

Newell, A. and Simon, H.A. (1972) Human Problem Solving, Englewood Cliffs, N.J.: Prentice Hall.

Peck, D.M. and Jenks, S.M. (1981) Conceptual issues in the teaching and learning of fractions, Journal for Research in Mathematics Education, 12, 339-348.

Reed, S.K., Ernst, G.W., and Banerji, R. (1974) The role of analogy in transfer between similar problem states. Cognitive Psychology, 6, 436-450.

Reitman, J.S. (1976) Skilled perception in Go: Deducing memory structures from inter-response times. Cognitive Psychology, 8., 336-356.

Riley, M.S, Greeno, J.G., and Heller, J.I. (1983) Development of children's problem solving ability in arithmetic. In H. Ginsberg (Ed.) The Development of Mathematical Thinking, New York: Academic Press, 153-196.

Rosch, E. and Mervis, C.B. (1975) Family resemblances: Studies in the internal structure of categories. _Cognitive Psychology, 8_, 573-605.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Braem, P. (1976) Basic objects in natural categories. _Cognitive Psychology, 8_, 382-439.

Rosch, E., Simpson, C. and Miller, R.S. (1976) Structural bases of typicality effects, _Journal of Experimental Psychology: Human Perception and Performance, 2_, 491-502.

Simon, H.A. and Hayes, J.R. (1976) The understanding process: Problem isomorphs. _Cognitive Psychology, 8_, 165-190.

Smith, E.E., Shoben, E.J., and Rips, L.J. (1974) Structure and process in semantic memory: A feature model for semantic decisions. _Psychological Review_, 214-241.

Silver, E.A. (1979) Student perceptions of relatedness among mathematical verbal problems. _Journal for Research in Mathematics Education, 10_.

Silver, E.A. (1981) Recall of mathematical problem information: Solving related problems. _Journal for Research in Mathematics Education, 12_, 54-64.

Watson, W.H. (1980) Results and implications of an arithmetic test. _Mathematics in School, 9_, 9-11.

Webster's Third New International Dictionary (1976), Springfield, MA: Merriam.

# APPENDIX I

## Materials For Experiment I

## Instructions - Part 1

Please read or examine carefully each numbered problem. Four choices are given with each problem. Please pick from among the four choices the one which you think goes the best with the problem and mark a 1 next to that letter on your problem sheet. Next look at the three choices which remain. Please pick from these three choices the one which you think goes the next best with the problem and mark your problem sheet with a 2.

Please read and answer the following two example problems:

1 Billy had 28 marbles in his collection. He won 9 marbles in his next game of marbles. How many marbles does Billy have now?

____ A) 15 + 17 ____ B) 28 - 9

____ C) 28 + 9 ____ D) 42 - 15

Most people would say that C) or 28 + 9 is the best answer, since the story describes a situation where one should add the numbers and the numbers are the same ones that are given in the story. A) or 15 + 17 would be a good second best choice, since those numbers are also added together.

2 10.0

____ A) 11.0 ____ B) 10

____ C) $\frac{100}{10}$ ____ D) 238

Most people would say that B) or 10 is the best answer, since the value of 10 and the value of 10.0 are the same and they look fairly similar. C) or 100/10 is a good second best choice because the value of 100/10 is also 10.

The actual problems you will do are similar to these, although some problems may seem more difficult. Please do your best with all the problems. Answer them in the same way, marking your first choice with a 1 and your second choice with a 2. Please give two answers for each problem and do not skip any problems.

Do you have any questions?

1.  :

___ A) 　___ B) 　___ C) 　___ D) 

2. $\frac{3}{4} \times \frac{1}{4}$ :

___ A) Ralph's goat gave 3/4 of a gallon of milk this morning. Ralph gave 1/4 of the milk to his friend, Andy. How much milk did Andy get?

___ B) Christine still had 3/5 of her reading assignment left to do. She had been able to complete 1/5 of the assignment each hour. At that rate, how long should it take Christine to finish her assignment?

___ C) Cheryl won some money with a lottery ticket and was allowed to keep 2/3 of it after taxes. She put 1/3 of what she received in the bank. What fraction of the total amount of money that Cheryl won went into the bank?

___ D) Ralph's goat gave 3/4 of a gallon of milk this morning. Ralph gave 1/4 of a gallon of milk to his friend, Andy. How much milk did Ralph have then?

3. Janet works 1 day a week as an actress and 2 days a week as a clerk. What fraction of the time does Janet work as an actress?

___ A) Janet works 2 days a week as an actress and 2 days a week as a clerk. What fraction of the time does Janet work as an actress?

___ B) The Wells invited 6 friends to dinner. 2 friends could not come. What fraction of the invited guests could not come?

___ C) Janet works 2 days a week as an actress and 4 days a week as a clerk. What fraction of the time does Janet work as an actress?

___ D) Simon cut 4 cords of wood for the winter. In the spring he had 2 cords of wood left. What fraction of the wood was left?

4. Jane and Clay spent 4 nights of their vacation at a hotel and [illegible] nights at a campground. What fraction of the time did they spend at the hotel?

___ A) 　___ B) 　___ C) 　___ D) 

5. Joey bought 17/20 of a mile of kite string so he could fly his kites. He wanted each kite string to be 9/33 of a mile long. How many kites does Joey have string for?

___ A) $\frac{7}{15} \div \frac{13}{30}$　___ B) $\frac{17}{20} \div \frac{9}{33}$

___ C) $\frac{25}{48} \times \frac{11}{36}$　___ D) $\frac{17}{20} \times \frac{9}{33}$

6. The box of cereal was 5/8 full. George poured 1/8 of the box into his bowl. How much cereal was then in the box?

___ A) Kate had 5/6 of a pound of rice. Jeff used 1/6 of a pound of the rice for dinner. How much rice did Kate have then?

___ B) The box of cereal was 5/8 full. George poured 1/8 of a leftover [illegible] into the box of cereal. How much cereal was then in the box?

___ C) The box of cereal was 4/6 full. George poured 1/6 of the box into his bowl. How much cereal was then in the box?

___ D) Tony picked 1/4 of a pound of lettuce. He put it with the 2/4 of a pound of lettuce he already had. How much lettuce did Tony have then?

7. $\frac{3}{4}$ :

____ A) Gordon found 3 scallop shells and 4 clam shells at the beach. What fraction of Gordon's shells were scallop shells?

____ B) Theresa's wedding bouquet had 32 roses. 24 of them were pink roses and 6 of them were white roses. What fraction of the roses were pink?

____ C) In Louisa's collection of model animals there are 3 brown horses and 1 white horse. What fraction of the horses are brown?

____ D) Beth and Frank had a job cutting lawns for 12 families. Yesterday they cut lawns for 3 families. What fraction of all the lawns were cut yesterday?

8. $\frac{6}{8}$ :

____ A) 6 ____ B) $\frac{3}{4}$

____ C) $\frac{7}{9}$ ____ D) $\frac{9}{12}$

9. $\frac{31}{36} \div \frac{15}{18}$ :

____ A) The Williams family bought 24/50 of a ton of lawn fertilizer. They needed 10/33 of a ton of fertilizer to feed the whole lawn. How many times could they feed the lawn?

____ B) Joni and Paul were laying out a model railroad track. They had 31/36 of a yard of straight track. They used 15/18 of it in building the track. How much straight track did they use?

____ C) Joni and Paul were laying out a model railroad track. They had 31/36 of a yard of straight track. This was enough to cover 15/18 of the length they wanted to cover. How much track would they need to cover the whole length?

____ D) Naomi had 12/16 of an ounce of perfume. She gave 3/12 of it to her friend Claire. How much perfume did Claire receive?

10. Albert collected 3/8 of a pound of aluminum. His friend gave him another 1/8 of a pound of aluminum. How much aluminum does Albert have?

____ A) $\frac{3}{8} + \frac{1}{8}$ ____ B) $\frac{3}{7} - \frac{2}{7}$

____ C) $\frac{2}{5} + \frac{1}{5}$ ____ D) $\frac{3}{8} - \frac{1}{8}$

11.  :

____ A) Maribeth took careful records of the temperature for 4 days. On _ days the temperature went above 70 degrees. What fraction of the days did the temperature go above 70 degrees?

____ B) In the Smith family, 1 of the children has red hair and the other 2 children have brown hair. What fraction of the Smith children have red hair?

____ C) Mark baked a pumpkin pie for dinner and cut it into 6 pieces. When he came home he found his brother had eaten 2 pieces of the pie. What fraction of pie did Mark's brother eat?

____ D) The cook mixed 2 cups of white flour with 6 cups of whole wheat flour for the bread. What fraction of the flour mixture was white flour?

12. After a day of panning for gold, Melinda had collected 13/20 of an ounce. Barry had collected 11/36 of an ounce. How much more gold did Melinda collect than Barry?

____ A) The day the seed sprouted it grew 7/20 of an inch. The next day it grew 3/14 of an inch. How much did it grow in those two days?

____ B) After a day of panning for gold, Melinda had collected 15/32 of an ounce and Barry had collected 9/25 of an ounce. How much more gold did Melinda collect than Barry?

____ C) After two hours, Matt had 9/26 of the marathon to go. Fifteen minutes later he had 14/52 of the route left to go. What fraction of the route had he run in those fifteen minutes?

____ D) After a day of panning for gold, Melinda had collected 13/20 of an ounce. Barry had collected 11/36 of an ounce. How much gold did they have together?

13.  :

____ A) At Leslie s dinner party there was 1 cake and 3 pies for dessert. What fraction of the desserts were cakes?

____ B) Lucy had 20 pieces of candy. She gave 5 pieces to Charlie Brown and kept 15 pieces for herself. What fraction of the candy did Lucy give away?

____ C) The Clark house has 3 paintings in the living room. 1 painting is a watercolor. What fraction of the paintings are watercolors?

____ D) Pepper had 1 white puppy and 4 spotted puppies in her litter. What fraction of the puppies were white?

14. $\frac{1}{3}$ :

____ A)  ____ B)  ____ C)  ____ D) 

15. Ms. Gray reserved 4/8 of the seats in the theater. There were enough seats for 5/8 of the students in the school. What fraction of the theater was needed to fit all the students?

____ A) Ms. Gray reserved 4/8 of the seats in the theater. She found enough students to fill 5/8 of the seats she had reserved. What fraction of the theater was filled by the students?

____ B) Ms. Gray reserved 5/10 of the seats in the theater. There were enough seats for 7/10 of the students in the school. What fraction of the theater was needed to fit all the students?

____ C) Matthew had 3/4 of a pound of radish seed. 1/4 of the seed did not sprout. How much seed did not sprout?

____ D) Nicole had 2/5 of a pound of popcorn. She used 1/5 of a pound of popcorn for each popcorn string she made. How many popcorn strings could Nicole make?

16. Aunt Helen had 11/40 pieces of cheesecake left after the Christmas party. The next day the family ate 15/80 pices of the cheesecake. How much cheesecake did Aunt Helen have then?

____ A) Greg measured his feet and hand prints and found they covered 19 [illegible] of a square foot. His hands alone covered 17/72 square feet. How much area did his feet alone cover?

____ B) Marie Davis owns 17/33 parts of a real estate business. Her husband, Dick owns 10/66 parts of the business. What fraction of the business do Marie and Dick own together?

____ C) Aunt Helen had 11/40 pieces of cheesecake after the Christmas party. The next day the family ate 5/8 of what was left. How much cheesecake did Aunt Helen have left?

____ D) Aunt Helen had 7/24 pieces of cheesecake left after the Christmas party. The next day the family ate 11/48 pieces of cheesecake. How much cheesecake did Aunt Helen have then?

17. Jeff had 5/6 of a pack of gum. His sister gave Jeff another 1/8 of a pack of gum. How much gum did Jeff have then?

____ A) Allan had 5/8 of the bookshelves. He gave Debbie 1/8 of the shelves. What fraction of the shelves did he have left?

____ B) Jeff had 5/8 of a pack of gum. He gave his sister 1/8 of a pack of gum. How much gum did Jeff have then?

____ C) Jeff had 3/5 of a pack of gum. His sister gave Jeff another 1/5 of a pack of gum. How much gum did Jeff have then?

____ D) Chris ate 2/6 of the cake and Pam ate 1/6 of the cake. How much cake did they eat together?

18. Greg got 3 new shirts for his birthday. 1 of them was blue. What fraction of the shirts was blue?

____ A)  ____ B)  ____ C)  ____ D) 

19. Fresh laundry detergent comes in 8/16 pound boxes. The Ables use 2/16 of a pound of detergent each week. How long will one box of Fresh last?

____ A) Fresh laundry detergent comes in 8/16 pound boxes. The Ables used 2/16 of a pound of detergent last week. How much detergent was left?

____ B) Fresh laundry detergent comes in 5/10 pound boxes. The Ables use 1/10 of a pound of detergent each week. How long will one box of Fresh last?

____ C) 2/3 of all households have one car. 1/3 of all cars are black. What fraction of all households has a black car?

____ D) The peanut man had 3/4 of a pound of peanuts. However, this was enough for only 1/4 of the people who wanted to buy peanuts. How many pounds of peanuts would the peanut man need to have enough for all the people?

20. $\frac{3}{9}$ :

____ A) Paul made 9 cups of an orange drink by mixing 3 cups of orange juice with 6 cups of soda. What fraction of the mixture was orange juice?

____ B) Irene's afghan had 18 red squares and 36 blue squares. What fraction of the afghan squares were red?

____ C) Gail's class had 9 girls and 3 boys. What fraction of Gail's class were boys?

____ D) There were 6 trophies on Steve's shelf. 3 of them were for first place. What fraction of the trophies were for first place?

21. $\frac{3}{5} - \frac{1}{5}$ :

____ A) Meg bought 3/5 of a pound of colored sugar at the store. A leak in the bag caused her to lose 1/5 of a pound of the sugar. How much sugar did Meg have then?

____ B) Cupcake the cat eats 1/3 of a can of catfood a day. The can was 1/3 full this morning. How much will be left at the end of the day?

____ C) Meg bought 3/5 of a pound of colored sugar at the store. When she got home she found she already had 1/5 of a pound of sugar. How much sugar did Meg have then?

____ D) Martha had painted 3/6 of the wall. Emily came over and painted 2/6 of the wall for her. How much of the wall was painted?

22. $\frac{2}{4} \div \frac{1}{4}$ :

____ A) Sandy had 2/4 of a pound of stuffing for the chicken. It was enough to fill only 1/4 of the bird. How much stuffing does Sandy need to fill the chicken?

____ B) Nicki had 4/10 of a pound of coffee. Her coffee machine used 1/10 of pound of coffee for one pot. How many pots of coffee could Nicki make?

____ C) Sandy had 2/4 of a pound of stuffing for the chicken. He used 1/4 of it to fill the bird. How much stuffing was used to fill the chicken?

____ D) 1/3 of the grapes grown in New York state are green. 2/3 of the green grapes are made into wine. What fraction of all the grapes in New York state are green grapes used for wine?

23. $\frac{1}{2}$ :

____ A) 

____ B) 

____ C) 

____ D) 

24. Bob had 6 free weekends during the summer when he planned to go to the beach. On 3 of the weekends it rained. What fraction of Bob's free weekends were spoiled by rain?

____ A) 6 of the children in the nursery school class are girls and 3 of the children are boys. What fraction of the children are boys?

____ B) Jennifer bought 2 of the chairs in the kitchen set and Teddy bought 2 more chairs for the kitchen set. What fraction of the chairs did Jennifer buy?

____ C) Bob had 9 free weekends during the summer when he planned to go to the beach. On 3 of the weekends it rained. What fraction of Bob's free weekends were spoiled by rain?

____ D) Bob had 2 free weekends during the summer when he planned to go to the beach. On 1 of the weekends it rained. What fraction of Bob's free weekends were spoiled by rain?

25. A woman left 5/8 of her total fortune to her husband. He spent 1/8 of what he received on a trip to China. How much of the total fortune was spent on the trip?

____ A) $\frac{4}{6} \times \frac{1}{6}$ ____ B) $\frac{5}{8} - \frac{1}{8}$

____ C) $\frac{5}{8} \times \frac{1}{8}$ ____ D) $\frac{7}{8} \div \frac{1}{8}$

26. $\frac{4}{6}$ :

____ A)  ____ B)  ____ C)  ____ D) 

27. $\frac{2}{5} + \frac{1}{5}$ :

____ A) John had 3/5 of a quart of milk. His cat drank 1/5 of a quart of the milk. How much milk does John have now?

____ B) Peggy owns 2/5 of an acre of land. She bought another 1/5 of an acre. How much land does Peggy have?

____ C) Karen bought 5/8 of a yard of red fabric. Her friend bought 3/8 of a yard of the same color cloth. How much red fabric do they have together?

____ D) Peggy owns 2/5 of an acre of land. She sold 1/5 of an acre of the land. How much land does Peggy have?

Please STOP here. You are halfway through with the written phase of the study. Please tell the experimenter that you have arrived at this point.

28. $\frac{10}{26} + \frac{14}{52}$ :

____ A) In 1983, Alabama produced 49/75 million gallons of orange juice. Georgia produced 25/53 million gallons of orange juice. How much orange juice did these two states produce together?

____ B) Wheat was selling for 7/24 dollars a pound in July. In September the price was 11/54 dollars a pound. How much did the price per pound drop?

____ C) Mac was carrying 10/26 tons in his truck. In Cincinnati he dropped off 14/52 tons. How much weight was Mac's truck carrying then?

____ D) Mac was carrying 10/26 tons in his truck. In Cincinnati he picked up another 14/52 tons. How much weight was Mac's truck carrying then?

29. 4/5 of the Rogers family went to a reunion party. 3/5 of the family members at the party ate cheesecake. What fraction of the whole family ate cheesecake?

____ A) 5/6 of the Rogers family went to a reunion party. 2/6 of the family members at the party ate cheesecake. What fraction of the whole family ate cheesecake?

____ B) Ernie made 3/4 of the recipe for cookie dough. He had enough dough to make 2/4 of the number of cookies that he wanted. What fraction of a recipe of dough would Ernie have to make to have enough dough?

____ C) Maureen had 2/3 of a pound of turkey left from Thanksgiving. She used 1/3 of it to make some soup. How much turkey did Maureen use?

____ D) 4/5 of the Rogers family went to a reunion party. 3/5 of the whole family ate cheesecake. What fraction of the whole family was at the party but did not have cheesecake?

30. $\frac{21}{24} - \frac{3}{20}$ :

____ A The Japanese restaraunt had 21/24 of a gallon of soy sauce. After filling a small bottle for each table, there was 3/20 of a gallon left. How much soy sauce was used to fill the jars?

____ B) The hen sat on her eggs for 17/60 of an hour. She got some food and came back to sit on her eggs for another 10/25 of an hour. How long did the hen sit on her eggs?

____ C) On December 1, the ice on the pond was **16/30** of a foot thick. On January 1, the ice was **11/18** of a foot thick. How much thicker was the ice in January than in December?

____ D) The Japanese restarant had 21/24 of a gallon of soy sauce. 3/20 of a gallon could fit into one small jar. How many jars could be filled?

31. Amanda had 11/16 of a pound of coffee. She gave 7/32 of it to her mother for a party. How much coffee did Amanda give away?

____ A) Amanda had 11/16 of a pound of coffee. She gave 7/32 of a pound of coffee to her mother for a party. How much coffee did Amanda have left?

____ B) Amanda had 19/24 of a pound of coffee. She gave 12/15 of it to her mother for a party. How much coffee did Amanda give away?

____ C) Lassie has 5/12 of a bag of dog food. If she eats 1/18 of a bag each day, how long will the dog food last?

____ D) Bart had saved 19/21 of the money necessary to buy a computer. He lent his sister 4/13 of the money he had saved. What fraction of the total cost of the computer did Bart lend to his sister?

32. Jenny and Sarah raked lawns this fall. One day after the end of one hour of working, Jenny had raked 2/6 of a lawn while Sarah had raked 1/6 of the ... n. How much of the lawn had they raked together?

____ A) Jenny and Sarah raked lawns this fall. One day after the end of one hour of working, Jenny had raked 3/8 of a lawn, while Sarah had raked 2/8 of the lawn. How much of the lawn had they raked together?

____ B) Mrs. Young owned 2/5 of an acre of land. She decided to buy a piece of land next to hers, which was 1/5 of an acre. How much land c... Mrs. Young own then?

____ C) Jenny and Sarah raked lawns this fall. One day after the end of one hour of working, Jenny had raked 2/6 of a lawn, while Sarah had raked 1/6 of the lawn. How much more had Jenny raked?

____ D) Mr. Skinner had 2/3 of a basket of apples. He gave 1/3 of a basket of apples to his neighbors. What fraction of the basket did he have left?

33.  :

____ A) $\frac{1}{3}$ ____ B) 1

____ C) $\frac{2}{6}$ ____ D) $\frac{1}{2}$

34.  :

___ A)     ___ B)     ___ C)     ___ D) 

35. 10/13 of the children in the kindergarten class wanted to be in the spring pagent. There were enough children to fill 20/25 roles. What fraction of the class is needed to fill all the roles?

___ A) There was 3/16 of a gallon of orange juice for breakfast. One juice glass holds 1/20 of a gallon of juice. How many juice glasses can be filled?

___ B) 9/40 of the pieces in the quilt are square. 1/3 of the squares are large. What fraction of the pieces in the quilt are large squares?

___ C) 10/13 of the children in the kindergarten class wanted to be in the spring pagent. There were enough roles for 20/25 children who wanted them. What fraction of the class could be in the play?

___ D) 5/7 of the children in the kindergarten class wanted to be in the spring pagent. There were enough children for 28/32 of the roles. What fraction of the class is needed to fill all the roles?

36. $\frac{3}{5}$ :

___ A) 3    ___ B) $\frac{4}{5}$

___ C) $\frac{6}{10}$    ___ D) $\frac{27}{45}$

37. A restaurant mixed 6 pints of milk with 4 pints of cream to get 10 pints of creamer for coffee. What fraction of the mixture was milk?

___ A) $\frac{3}{5}$    ___ B) $\frac{6}{4}$

___ C) $\frac{6}{10}$    ___ D) 6

38. Mr. Fortini had 30/32 of a bushel of tomatoes to can. 1/10 of a bushel of tomatoes will fit into one jar. How many jars can he fill?

___ A) Greta was making a down jacket and had 5/16 of a pound of down. This was enough to fill 17/20 of the jacket. How much down was needed to fill the whole jacket?

___ B) Yesterday 5/24 of a foot of snow fell. This morning the rain washed 8/15 of it away. How much snow was left on the ground?

___ C) Mr. Fortini had 30/32 of a bushel of tomatoes to can. He put 1/10 of a bushel of tomatoes into a jar. What fraction of a bushel of tomatoes does Mr. Fortini have left?

___ D) Mr. Fortini had 30/48 of a bushel of tomatoes to can. 3/15 of a bushel of tomatoes will fit into one jar. How many jars can he fill?

39. Bill had collected 29/50 of the goal for the fund drive. His next donation was for 11/25 of the goal. What fraction of the goal had Bill collected then?

___ A) When he is nervous, Mr. DiCarli's heart rate is 31/39 beats per second. When he is calm, his heart rate is 44/65 beats per second. How much faster is Mr. DiCarli's heart rate when he is nervous than when he is calm?

___ B) Bill had collected 29/50 of the goal for the fund drive. He dropped his can and lost 11/25 of the goal. What fraction of the fund drive had Bill collected then?

___ C) Bill had collected 38/75 of the goal for the fund drive. His next donation was for 11/50 of the goal. What fraction of the goal had Bill collected then?

___ D) Joe had written 19/36 of the stories, while Pat had written 5/18 of the stories. What fraction of the stories have been written?

40. Carol had 3/4 of a pound of silver. When the price got high, she sold 1/4 of the silver. How much silver did Carol sell?

____ A) Carol had 3/4 of a pound of silver. When the price got high she sold 1/4 of a pound of the silver. How much silver did Carol have then?

____ B) Lori had 4/6 of a foot of cherry wood. She wanted pieces 2/6 of a foot long to make carvings. How many carvings could Lori make?

____ C) Carol had 2/3 of a pound of silver. When the price got high, she sold 1//3 of the silver. How much silver did Carol sell?

____ D) 5/8 of the ornaments on the Christmas tree were made of glass. The cat knocked the tree down and 2/8 of the glass ornaments broke. What fraction of all the ornaments broke?

41. In 2/3 of an hour the donut machine can make 1/3 of the donuts that are needed for the day. How long will it take to make all the donuts?

____ A) $\frac{2}{3} \div \frac{1}{3}$ ____ B) $\frac{2}{3} \times 3$

____ C) $\frac{2}{3} - \frac{1}{3}$ ____ D) $\frac{2}{3} \times \frac{1}{3}$

42. 9/24 of the garden is planted with trees. 2/10 of the trees have pink blossoms. What fraction of the garden is planted with trees that have pink blossoms?

____ A) Gabe had 9/16 of a pound of rice. He needs 1/7 of a pound of rice to make a batch of rice pudding. How many batches of rice pudding can Gabe make?

____ B) 9/24 of the garden is planted with trees. 2/10 of the garden is planted with trees that have pink blossoms. What fraction of the garden is planted with trees that don't have pink blossoms?

____ C) Mom had 5/12 of a pecan pie left after the dinner. She gave 2/5 of what was left to her nephew John. What fraction of the total pie did John receive?

____ D) 7/18 of the garden is planted with trees. 9/15 of the trees have pink blossoms. What fraction of the garden is planted with trees that have pink blossoms?

43. $\frac{2}{8}$ :

____ A)  ____ B)  ____ C)  ____ D) 

44. The Shay's car uses 9/50 of a gallon of gas to go one mile in the city. On the highway, the car uses 11/75 of a gallon of gas to go one mile. How much more gas does the car use to travel one mile in the city than on the highway?

____ A) $\frac{11}{75} + \frac{9}{50}$ ____ B) $\frac{15}{22} + \frac{8}{19}$

____ C $\frac{9}{50} - \frac{11}{75}$ ____ D) $\frac{21}{30} - \frac{10}{81}$

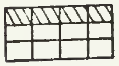
45. Diane and Carol are knitting identical scarves. Diane has knit 4/6 of her scarf, while Carol has knit 1/6 of her scarf. How much more of her scarf was Diane knit than Carol?

____ A) Diane and Carol are knitting identical scarves. Diane has knit 3/4 of her scarf, while Carol has knit 1/4 of her scarf. How much more of her scarf has Diane knit than Carol?

____ B) Diane and Carol are knitting identical scarves. Diane has knit 4/6 of her scarf, while Carol has knit 1/6 of her scarf. What fraction of a scarf have they knit together?

____ C) Cynthia had 5/8 of a pound of yarn for her project. Greg gave her 2/8 of a pound more of the yarn. How much yarn did Cynthia have?

____ D) Rob had 3/5 of a pound of sand in his leg weight. After he came back from a run, there was only 2/5 of a pound of sand in the leg weight. How much sand did Rob lose?

46.  :

____ A) $\frac{1}{3}$ ____ B) $\frac{1}{4}$

____ C) 4 ____ D) $\frac{4}{12}$

47. Caroline and John hiked 11/34 of the east coast trail. They covered another 3/16 of the trail last week. What fraction of the trail have they covered now?

___ A) $\frac{11}{34} - \frac{3}{16}$ ___ B) $\frac{15}{18} + \frac{4}{63}$

___ C) $\frac{11}{34} + \frac{3}{16}$ ___ D) $\frac{7}{23} - \frac{1}{11}$

48. A box contains 2 red marbles and 6 blue marbles. What fraction of the marbles are red?

___ A) 2 ___ B) $\frac{2}{6}$

___ C) $\frac{2}{8}$ ___ D) $\frac{1}{4}$

49.  :

___ A) $\frac{1}{3}$ ___ B) 1

___ C) $\frac{1}{4}$ ___ D) $\frac{2}{8}$

50. $\frac{19}{25} \times \frac{10}{38}$ :

___ A) Jessica had 19/25 of a package of beads. She gave 10/38 of her beads to her friend Anne. What fraction of the whole package did Anne receive?

___ B) Michael's waterbed was 30/48 full. He drained 19/30 of the remaining water before lunch. What fraction of the total capacity of the water bed did Michael drain before lunch?

___ C) Maureen had 7/12 of a bag of potting soil. One of her planters uses 2/15 of a bag of soil. How many planters can Maureen fill?

___ D) Jessica had 19/25 of a package of beads. She gave 10/38 of a package of beads to her friend Anne. What fraction of a package of beads did Jessica have then?

51. There was 3/4 of a chocolate cake sitting on the porch. The cat ruined … of the cake by stepping on it. How much of the cake was still edible?

___ A) $\frac{3}{4} + \frac{1}{4}$ ___ B) $\frac{5}{6} - \frac{3}{6}$

___ C) $\frac{3}{4} - \frac{1}{4}$ ___ D) $\frac{3}{7} + \frac{2}{7}$

52.  :

___ A) $\frac{3}{9}$ ___ B) $\frac{3}{6}$

___ C) 3 ___ D) $\frac{1}{3}$

53. Megan had 9/13 of her stuffed animals spread out on the floor. She moved … 9 of the animals that were on the floor to the bed. What fraction of all the stuffed animals were on the bed?

___ A) $\frac{10}{18} \div \frac{2}{7}$ ___ B) $\frac{14}{17} \times \frac{3}{12}$

___ C) $\frac{9}{13} \times \frac{4}{9}$ ___ D) $\frac{9}{13} \div \frac{4}{9}$

54. Auguste examined 8/23 of the patients on the ward, while Sam examined 24/46 of the patients on the ward. What fraction of the patients on the ward ha… Auguste and Sam examined together?

___ A) Auguste examined 6/17 of the patients on the ward, while Sam examined 18/24 of the patients on the ward. What fraction of the patients on the ward have Auguste and Sam examined together?

___ B) When Anna weighed all her jewelry, she found she had 15/19 of an ounce of gold. Her mother gave her another piece of gold jewelry weighing 1/11 of an ounce. How much gold did Anna have?

___ C) Auguste examined 8/23 of the patients on the ward, while Sam examined 24/46 of the patients on the ward. How many more patients has Sam examined than Auguste?

___ D) Frank used to eat ice cream 1/48 of the time for dinner. Since the new ice cream shop opened across the street, Frank has been eating ice cream 20/31 of the time for dinner. How much more often does Frank eat ice cream now?

Instructions for Part 2:

In this set of exercises you will be given a word problem and asked to decide whether you think you should add, subtract, multiply, or divide the two numbers to get the answer to the problem. You do not have to actually work out the answer to the problem. Please read each problem carefully and mark one and only one choice for your answer.

1. Hansel began the trip with 3/4 of a pound of bread. He used 1/4 of a pound of bread to mark the trail. How much bread did Hansel have then?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

2. Rachel tested 33/48 of the lab animals, while Harry tested 5/24 of the lab animals. What fraction of the lab animals have they tested together?

____ A) Addition ____ B) Substraction

____ C) Multiplication ____ D) Division

3. Arlen mixed up 5/12 of a bucket of birdseed. He found he had enough birdseed to fill 7/13 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

4. Grace had 3/4 of a pound of chocolate bits. She needed 1/4 of a pound of chocolate bits to make a batch of cookies. How many batches of cookies could Grace make?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

5. Charlie had 3/6 of a can of cake frosting. His neighbor, Ms. Field, gave him another 1/6 of a can of frosting. How much frosting did Charlie have then?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

6. Margret had 2/5 of a gallon of ice cream. She gave 1/5 of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

7. Hemingway used 47/80 of a box of typing paper last week, while Orwell used 34/40 of a box of typing paper. How much more paper did Orwell use than Hemingway?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

8. 7/20 of the beds in the garden were planted with flowers. 2/7 of the flowers were tulips. What fraction of the garden was planted with tulips?

____ A) Addition ____ B) Subtraction

____ C) Multiplication ____ D) Division

Please fill in the following:

Age ________ Sex _________

Year of graduation __________

Major ___________________________________

The last high school math class you took was __________

____________________________________

Please list any college mathematics courses you have taken

Do you anticipate taking more mathematics courses in college?

______yes _____ no

If yes,
What courses do you plan to take?

# APPENDIX II

Figures From Experiment I

FIGURE IIA – SECTION I: CUES VS TYPE

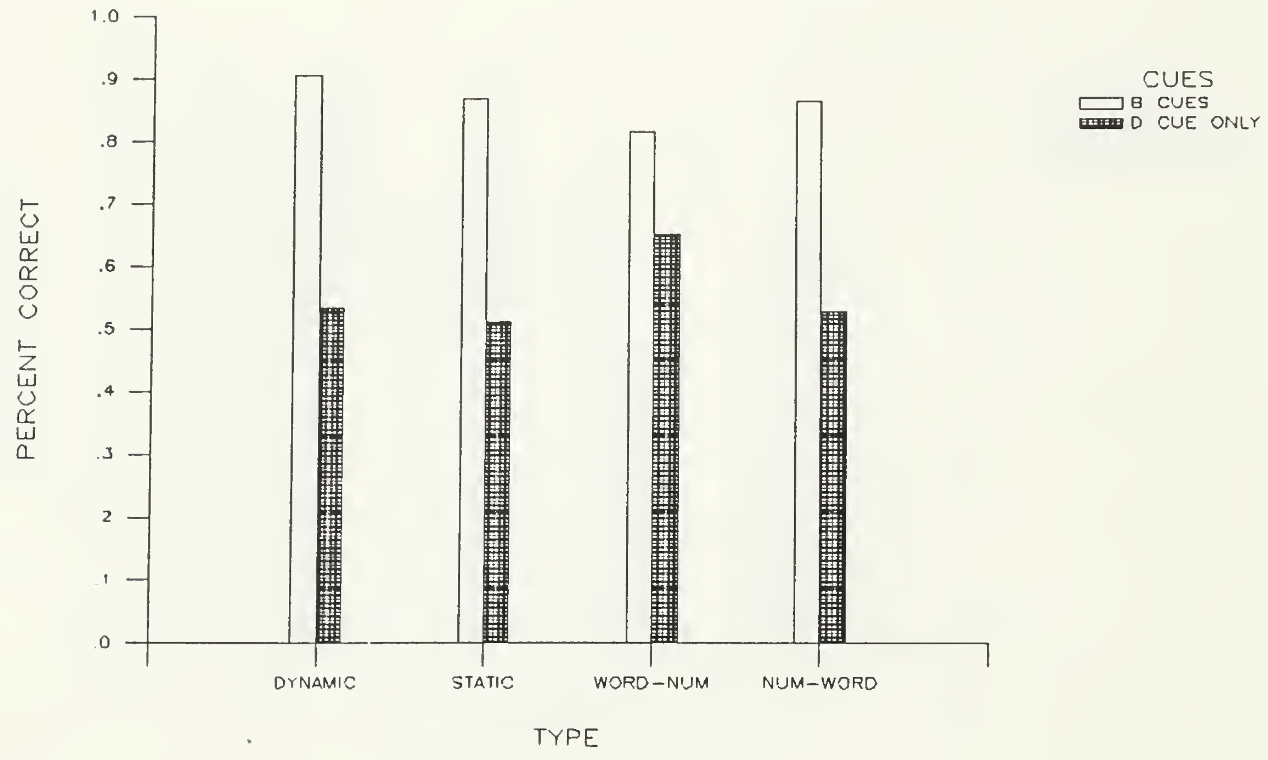


FIGURE IIB – SECTION I: TYPE VS OPERATION

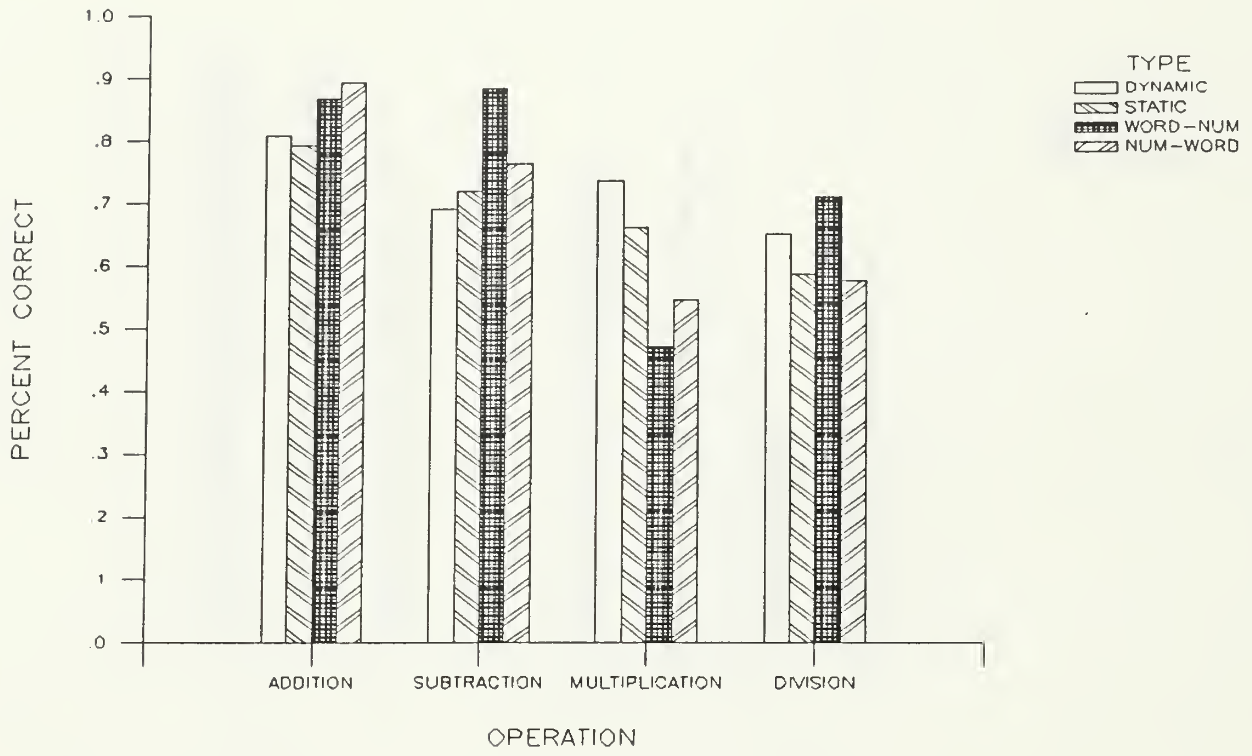


FIGURE IIC – SECTION I: CUES VS OPERATION VS TYPE

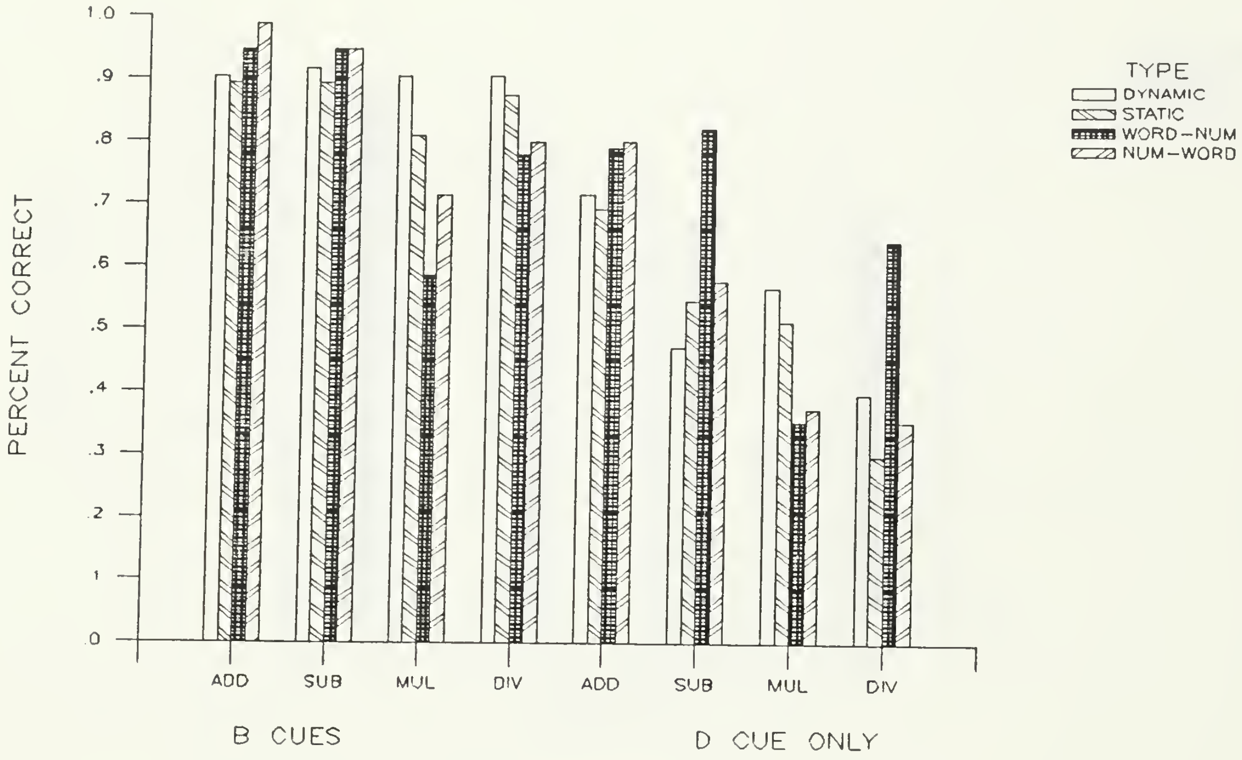


FIGURE IID — SECTION I: DIFFICULTY VS OPERATION

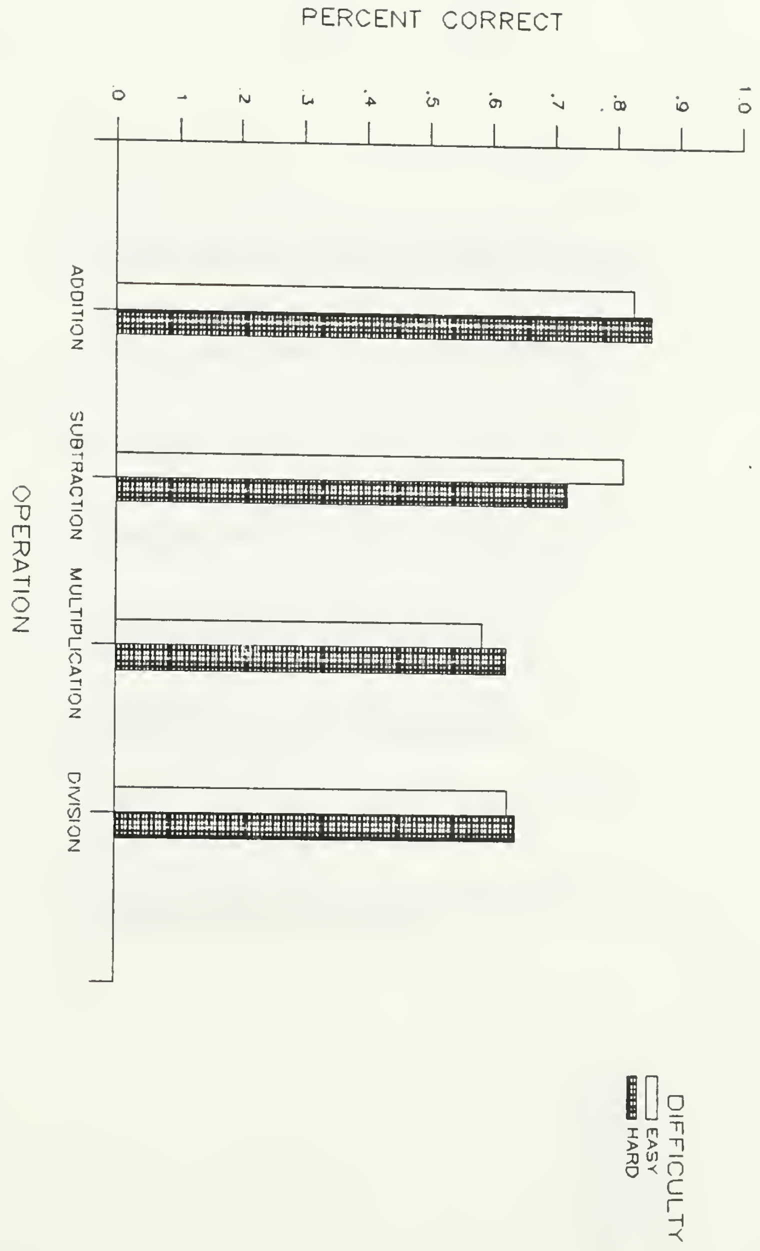


FIGURE IIE – SECTION I: DIFFICULTY VS OPERATION VS TYPE

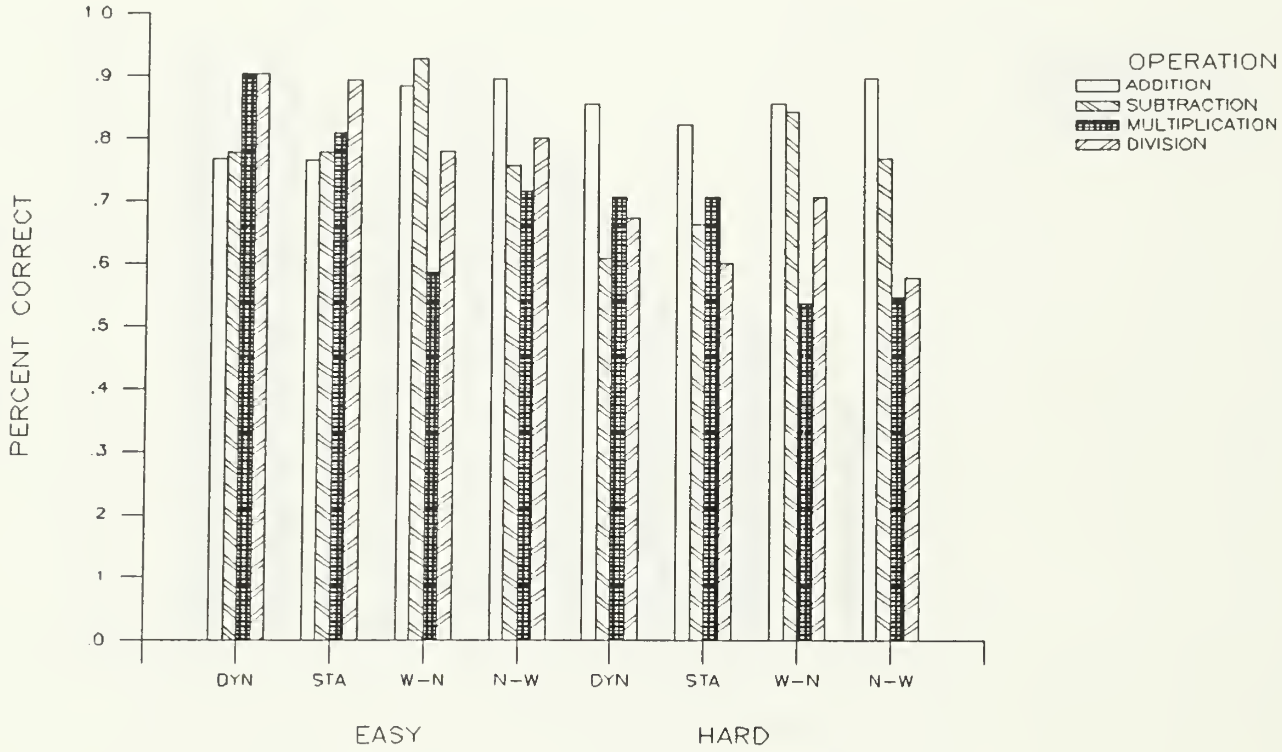


FIGURE IIF – SECTION I: CUES VS DIFFICULTY VS OPERATION

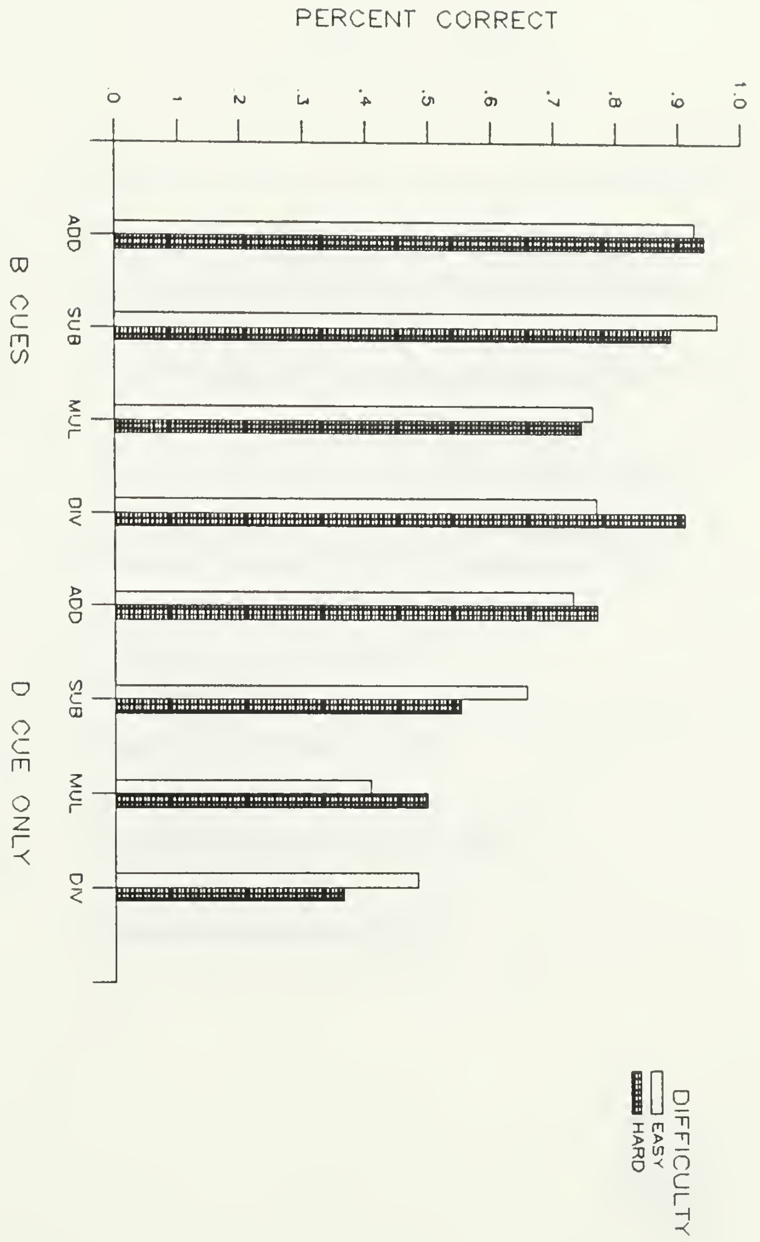


FIGURE IIG – SECTION I: CUES VS DIFFICULTY VS TYPE

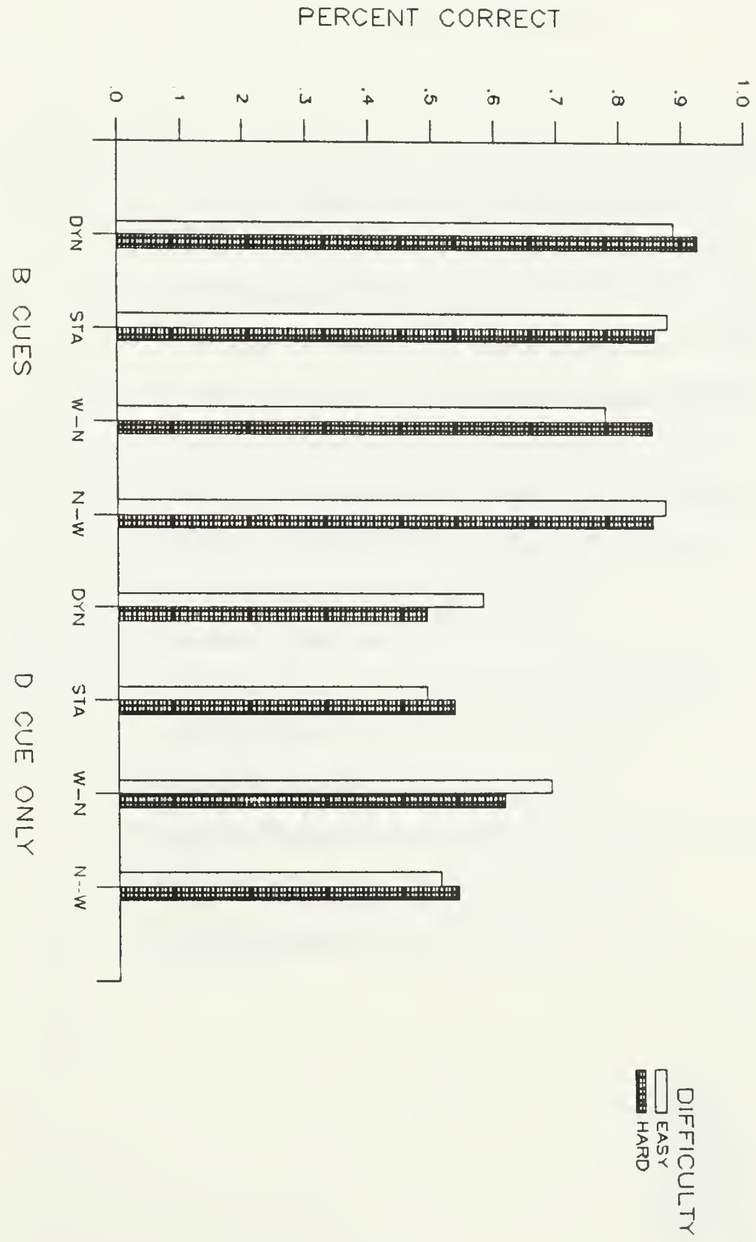
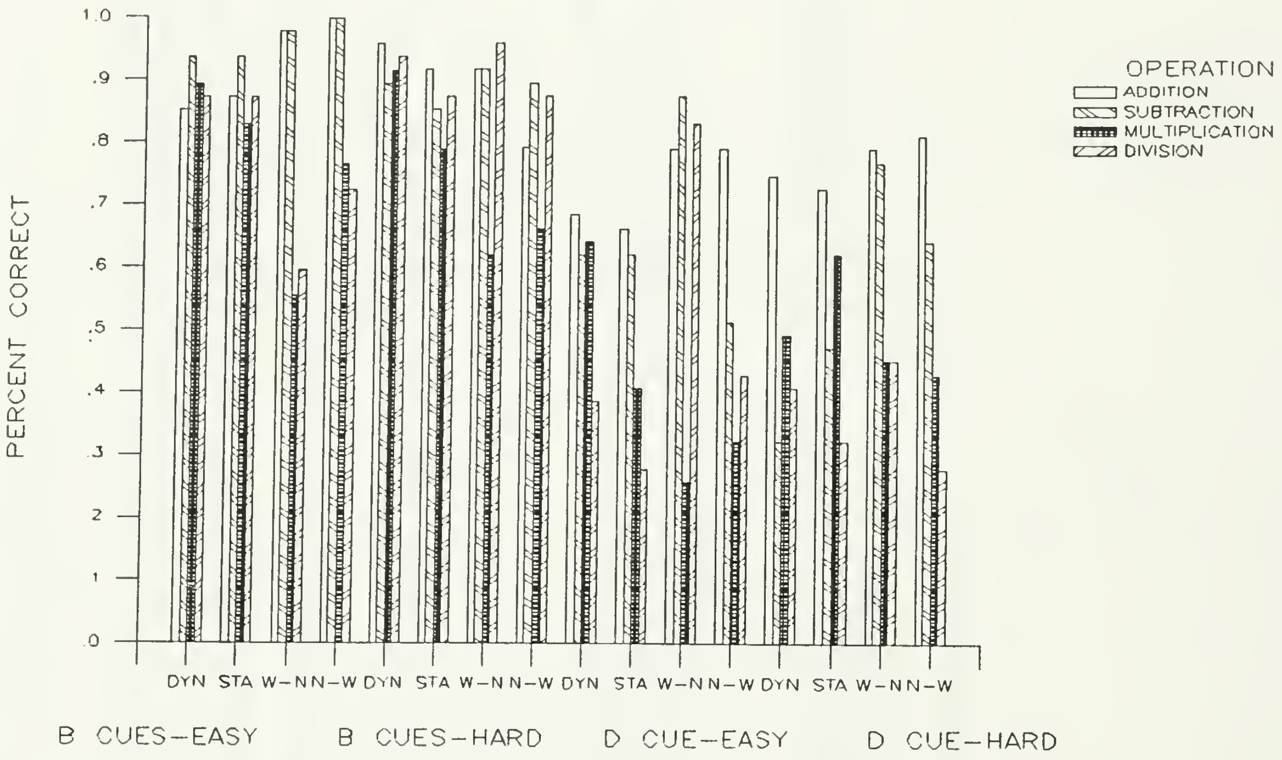


FIGURE IIH – SECTION I: CUES VS DIFFICULTY VS OPERATION VS TYPE



Instructions - Part A

Please read each numbered problem carefully. Four choices are given with each problem. Pick from among the four choices the one which you think is the same kind of problem as the numbered problem and mark the space next to your choice with a check. It is not necessary to actually solve the problems. If you would like to do any writing, you may do so in the margins.

Please read and answer the following sample problem:

Sample

A shoe store had 8 pairs of white sandals and 3 pairs of brown sandals left over from the summer season. How many pairs of sandals did the store have left over altogether?

___ A) 4 girlfriends were planning to go to the movies. They invited 2 younger children to go with them. How many people were planning to go to the movies then?

___ B) 4 girlfriends were planning to go to the movies. 2 girls could fit into each seat on the bus. How many seats on the bus did they need to use?

___ C) 4 girlfriends were planning to go to the movies. Each girl invited 2 younger children to come with her. How many children did they invite altogether?

___ D) 4 girlfriends were planning to go to the movies. Then 2 of the friends got sick and could not go. How many people were planning to go to the movies then?

The sample problem gives a situation in which you should add the two numbers together to get the answer to the question that is asked. Therefore, you should pick the choice which also describes a situation in which it is necessary to add. You would add the numbers together to get an answer for choice A. Therefore, A is the best answer.

The actual problems you will do are similar to these, although some problems may seem more difficult. Please do your best in answering all the problems. Do not skip any problems.

Do you have any questions?

# APPENDIX III

## Materials For Experiment II

1) Steve hopped around 1/3 of the track on one foot. Then he hopped on both feet around another 1/4 of the track. What fraction of the track has Steve gone around altogether?

___ a) Matthew had 3/5 of a pound of peanut butter. He wanted to make a big batch of cookies, so he bought another 1/2 pound of peanut butter. How much peanut butter did Matthew have then?

___ b) Matthew had 3/5 of a pound of peanut butter. His friend wanted to make a batch of cookies, so he gave her 1/2 pound of his peanut butter. How much peanut butter did Matthew have then?

___ c) Matthew had 3/5 of a pound of peanut butter. He needed 1/2 a pound of peanut butter to make a batch of cookies. How many batches of cookies could Matthew make?

___ d) Matthew had 3/5 of a pound of peanut butter. His friend wanted to make a batch of cookies, so he gave her 1/2 of his peanut butter. How much peanut butter did Matthew have then?

2) Riding his bike, it has taken Teddy 2/3 of an hour to make it 3/4 of the way home. How long should it take Teddy to make the whole trip?

___ a) Mrs. Young owned 5/8 of an acre of land. She bought an adjoining 1/4 of an acre plot of land. How much land did she have then?

___ b) Mrs. Young owned 5/8 of an acre of land. She sold 1/4 of the land. How much land did she sell?

___ c) Mrs. Young owned 5/8 of an acre of land. She wanted to make garden plots that were 1/4 of an acre in size. How many plots could she make?

___ d) Mrs. Young owned 5/8 of an acre of land. She sold 1/4 of an acre of the land. How much land did she have then?



3) After the shower, David checked the rain guage and found 3/8 of an inch of rain had fallen. It rained again later and David found 1/4 of an inch of rain fell then. How much rain fell during the two showers?

___ a) The Skinners picked 9/10 of a bushel of corn to freeze. 1/12 of a bushel of corn will fill one freezer container. How many containers can they fill?

___ b) The Skinners picked 9/10 of a bushel of corn to freeze. 1/12 of the corn would fit into one freezer container. How much corn fit into one container?

___ c) The Skinners picked 9/10 of a bushel of corn to freeze. 1/12 of a bushel of corn will fit into one freezer container. How much corn would be left after filling one container?

___ d) The Skinners picked 9/10 of a bushel of corn to freeze. They already had 1/12 of a bushel of corn. How much corn did they have then?

4) The tortoise and the hare were having a race. The tortoise had traveled 7/8 of the route so far, while the hare was 3/5 of the way along the route. How much further was the tortoise than the hare?

___ a) Carol had 5/8 of a pound of silver. When the price of silver got high, she sold 1/4 of her silver. How much silver did Carol sell?

___ b) Carol had 5/8 of a pound of silver. When the price of silver got high, she sold 1/4 of a pound of her silver. How much silver did Carol have then?

___ c) Carol had 5/8 of a pound of silver. If she sells 1/4 of a pound of silver, she will have enough money to buy a program for her computer. How many programs can Carol buy if she sells all the silver?

___ d) Carol had 5/8 of a pound of silver. When the price of silver was low, she bought 1/4 of a pound of silver. How much silver did Carol have then?

5) Michael and Gale were packing quarts of their homemade chocolate chip ice cream to fill an order for a store. They were able to fill 3/5 of the quart containers with 1/4 of a batch of ice cream. What fraction of a batch of ice cream do they need to fill all the containers?

___ a) 7/13 of the children in the kindergarten class wanted to play bunnies in the spring pagent. There were enough bunny roles for 5/12 of the class. How many more bunny roles than children who wanted to play bunnies were there?

___ b) 7/12 of the children in the kindergarten class wanted to play bunnies in the spring pagent. There were enough children for 5/12 of the bunny roles. What fraction of the class is needed to fill all the bunny roles?

___ c) 7/12 of the children in the kindergarten class wanted to play bunnies in the spring pagent. 5/12 of the children wanted to play birds. What fraction of the class wanted to be either a bunny or a bird?

___ d) 7/12 of the children in the kindergarten class wanted to play bunnies in the spring pagent. There were enough bunny roles for 5/12 of the children who wanted to be bunnies. What fraction of the class could be bunnies in the play?

6) 1/3 of the respondents to the survey were Republicans. 3/4 of the Republicans were male. What fraction of the survey respondents were male Republicans?

___ a) Ms. Gray reserved 3/8 of the seats in the theater. She found enough students to fill 2/4 of the seats she had reserved. What fraction of the theater was filled by the students?

___ b) Ms. Gray reserved 3/8 of the seats in the theater. There were enough seats for 2/4 of the students in the school. What fraction of the theater was needed to fit all the students?

___ c) Ms. Gray reserved 3/8 of the seats in the theater. Mr. Thayer reserved 2/4 of the seats in the theater. What fraction of the theater was reserved?

___ d) Ms. Gray reserved 3/8 of the seats in the theater. Mr. Thayer reserved 2/4 of the seats in the theater. How much more of the theater had Mr. Thayer reserved than Ms. Gray?



7) John had 1/6 of a container of antifreeze. He found another 1/3 of a container of antifreeze when he cleaned the cellar. How much antifreeze did John have then?

___ a) 7/20 of the beds in the garden were planted with flowers. 3/7 of the beds in the garden were planted with tomatoes. How much more of the garden was planted in flowers than tomatoes?

___ b) 7/20 of the beds in the garden were planted with flowers. 3/7 of the flowers were tulips. What fraction of the garden was planted with tulips?

___ c) 7/20 of the beds in the garden were planted with flowers. 3/7 of the beds in the garden were planted with tomatoes. What fraction of the garden was planted with either flowers or tomatoes?

___ d) 7/20 of the beds in the garden were panted with flowers. There were enough flowers to fill 3/7 of the vases for the wedding. What fraction of the garden should have been planted in flowers in order to have had enough for the wedding?

8) The plane was carrying 7/10 of a ton of supplies to the isolated villages. It delivered 1/4 of a ton of supplies to the village of Queshawee. What fraction of a ton of supplies did the plane carry then?

___ a) Fresh laundry detergent comes in 8/16 pound boxes. The Howes bought a new box of Fresh and used 2/16 of the detergent last week. How much Fresh did they use?

___ b) Fresh laundry detergent comes in 8/16 pound boxes. The Howes bought a new box of Fresh used 2/16 of a pound of the detergent last week. How much detergent did they have then?

___ c) Fresh laundry detergent comes in 8/16 pound boxes. The Howes bought a new box of Fresh when their old box of Fresh had only 2/16 of a pound of detergent in it. How much detergent did they have then?

___ d) Fresh laundry detergent comes in 8/16 pound boxes. The Howes use 2/16 of a pound of detergent each week. How long will one box of Fresh last?

9) 7/10 of the rooms in the hotel were taken last night. By ten o'clock the next morning, the hotel workers were able to clean 3/5 of the rooms that had been used. What fraction of the rooms in the hotel had they cleaned?

___ a) Charlie had 3/6 of a can of cake frosting. His neighbor, Ms. Field, gave him another 1/3 can of frosting. How much frosting did Charlie have then?

___ b) Charlie had 3/6 of a can of cake frosting. He gave his neighbor, Ms. Field, 1/3 of his frosting. How much frosting did Ms. Field receive?

___ c) Charlie had 3/6 of a can of cake frosting. He gave his neighbor, Ms. Field, 1/3 of a can of frosting. How much frosting did Charlie have then?

___ d) Charlie had 3/6 of a can of cake frosting. He needed 1/3 of a can of frosting to frost one loaf cake. How many loaf cakes could Charlie frost?

10) 3/5 of the fans at gymnastic events are women. 9/10 of the seats were filled for Thursday's competition. What fraction of the seats in the arena were filled by women?

___ a) Kristin had 1/2 a yard of fabric. She needed 1/3 of a yard of fabric to make a doll's dress. How many dresses could she make?

___ b) Kristin had 1/2 a yard of fabric. Her friend Shelley gave her 1/3 of a yard of fabric. How much fabric did Kristin have then?

___ c) Kristin had 1/2 a yard of fabric. She gave 1/3 of a yard of fabric to her friend Shelley. How much fabric did Kristin have then?

___ d) Kristin had 1/2 a yard of fabric. She gave 1/3 of the fabric away to her friend, Shelley. How much fabric did Shelley receive?



11) Chris and Tom were playing a game of trivia. Chris was able to answer 3/5 of the questions he was asked, while Tom was able to answer 2/3 of the questions he was asked. What fraction more of his questions was Tom able to answer?

___ a) Joni and Paul laid out a model railroad track. They had 31/36 of a yard of straight track. 15/18 of the track was used in the model. What fraction of a yard of straight track was used in the model?

___ b) Joni and Paul laid out a model railroad track. They had 31/36 of a yard of straight track and 15/18 of a yard of curved track. How much more straight track than curved track did they have?

___ c) Joni and Paul laid out a model railroad track. They had 31/36 of a yard of straight track and 15/18 of a yard of curved track. How much track did they have altogether?

___ d) Joni and Paul laid out a model railroad track. They had 31/36 of a yard of straight track. This was enough to cover 15/18 of the length they wanted to cover. How much track would they need to cover the whole length?

12) Julie was able to wash 1/3 of the windows in the house on Saturday morning. Scott washed 1/4 of the windows. What fraction of the windows in the house did they wash together?

___ a) Laurel had 4/9 of a pound of dried fruit. 1/3 of the fruit was dried pineapple. How much dried pineapple was there?

___ b) Laurel had 4/9 of a pound of dried fruit. She had enough fruit to make 1/3 of a recipe of fruitcake. How much fruit did Laurel need to make the whole recipe?

___ c) Laurel had 4/9 of a pound of dried fruit. Amy had 1/3 of a pound of dried fruit. How much dried fruit did they have together?

___ d) Laurel had 4/9 of a pound of dried fruit. Amy had 1/3 of a pound of dried fruit. How much more dried fruit did Laurel have than Amy?

13) Joe had 1/2 a box of screws. He needed 1/5 of a box of screws to put together a wooden crate. How many crates could Joe put together with his screws?

___ a) Sandy had 2/4 of a pound of stuffing for the turkey. 1/4 of it fit into the bird. How much stuffing fit in the turkey?

___ b) Sandy had 2/4 of a pound of stuffing for the turkey. This was enough to fill only 1/4 of the bird. How much stuffing was needed to fill the turkey?

___ c) Sandy had 2/4 of a pound of stuffing for the turkey. Harriet had 1/4 of a pound of stuffing for the mushrooms. How much stuffing did they have together?

___ d) Sandy had 2/4 of a pound of stuffing for the turkey. Harriet had 1/4 of a pound of stuffing for the mushrooms. How much more stuffing did Sandy have than Harriet?

14) Janice had 7/8 of a gallon of potato salad left after the party. She gave 1/4 of the leftover potato salad to one of the guests who was leaving. How much potato salad did the guest receive?

___ a) Ernie made 3/4 of the recipe for cookie dough. Bert made 1/3 of the recipe for cookie dough. How much more dough did Ernie have than Bert?

___ b) Ernie made 3/4 of the recipe for cookie dough. 1/3 of the dough was in the refridgerator. What fraction of the recipe was in the refridgerator?

___ c) Ernie made 3/4 of the recipe for cookie dough. Bert made 1/3 of the recipe for cookie dough. How much dough do they have together?

___ d) Ernie made 3/4 of the recipe for cookie dough. He had enough dough to make 1/3 of the number of cookies he wanted. What fraction of the recipe would Ernie have to make in order to have enough dough?



15) The box car was 5/6 full of lumber. Then the train stopped and delivered 2/5 of a carload of lumber to a lumber yard. How much lumber did the box car have then?

___ a) Diane and Carol are knitting identical scarves. Diane has knit 4/6 of her scarf while Carol has knit 1/6 of her scarf. How much more of her scarf has Diane knit than Carol?

___ b) Diane and Carol are knitting identical scarves. They each had 4/6 of an ounce of yarn. This was enough to complete 1/6 of each scarf. How much yarn would they need to make scarves the length they wanted?

___ c) Diane and Carol are knitting identical scarves. Diane has knit 4/6 of her scarf, while Carol has knit 1/6 of the amount Diane knit. What fraction of her scarf has Carol knitted?

___ d) Diane and Carol are knitting identical scarves. Diane has knit 4/6 of her scarf, while Carol has knit 1/6 of her scarf. What fraction of a whole scarf would they have if they put the two pieces together?

16) The parents had 12/16 of an ounce of medicine for the baby. The baby was supposed to receive 1/8 of an ounce of medicine each day. How long should the medicine last?

___ a) Ralph's goat gave 3/4 of a gallon of milk this morning. He poured it into a jar that already had 1/4 of a gallon of milk in it. How much milk did Ralph have in the jar?

___ b) Ralph's goat gave 3/4 of a gallon of milk this morning. Ralph poured out 1/4 of a gallon of milk to give to his friend, Andy. How much milk did Ralph have then?

___ c) Ralph's goat gave 3/4 of a gallon of milk this morning. Ralph had several 1/4 gallon jars. How many jars could he fill with the milk?

___ d) Ralph's goat gave 3/4 of a gallon of milk this morning. Ralph poured out 1/4 of the milk to give to his friend, Andy. How much milk did Andy receive?

Instructions - Part B

In this section, you will be given a word problem and asked to decide whether you should add, subtract, multiply, or divide the two numbers given in the problem in order to answer the problem. You do not have to actually work out the answer to the problem. However, if you would like to do any writing, you may do so in the margin.

Please read each problem carefully and mark one and only one choice for your answer.

1) The swimming pool in the town park is 25 yards long. The pool at the college is 2 times as long. How long is the college pool?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

2) Mac was carrying 10/26 tons of produce in his truck. He dropped off 14/52 of a ton of the produce in Cincinatti. How much produce was Mac carrying then?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

3) Jeff had 5 packs of gum. His sister gave Jeff another 2 packs of gum. How much gum did Jeff have then?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

4) In 3/4 of an hour the donut machine can make 5/8 of the donuts that are needed for the day. How long will it take to make all the donuts?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

5) Jim has a 12 gallon can full of gasoline. His motorcycle tank holds 4 gallons of gas. How many times can he fill the tank?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide



6) 5/8 of the ornaments on the Christmas tree were made of glass. The cat knocked the tree down and broke 1/3 of the glass ornaments. What fraction of the tree ornaments were broken?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

7) The Shays' car uses 3/16 of a gallon of gas to go one mile in the city. On the highway, the car uses 1/12 of a gallon of gas to go one mile. How much more gas does the car use to travel one mile in the city than one mile on the highway?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

8) Aunt Helen had 11 pieces of cheesecake left after the Christmas party. The next day the family ate 6 pieces of the cheesecake. How much cheesecake did Aunt Helen have then?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

9) Paul has driven for 2/5 of the trip, while Janna has driven for 3/10 of the trip. What fraction of the trip have they driven together?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

10) Jenny and Sarah raked lawns this fall. At the end of last Saturday, Jenny had raked 6 lawns while Sarah had raked 4 lawns. How many lawns had they raked together?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

11) The librarian received 7/16 of the new books two weeks after she placed the orders. She received another 3/8 of the books three weeks after she placed the orders. What fraction of the books had she received then?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

12) Ed bet ten dollars at the racetrack. When his horse won he received 3 times as much as he bet. How much money did Ed receive?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

13) After a day of panning for gold, Steeleye had collected 8 ounces. Yeastcake had collected 5 ounces. How much more gold did Steeleye collect than Yeastcake?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

14) The drama class had 27 people. They split up into 3 equal-sized groups. How many people were in each group?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide



15) Maureen had 5/6 of a bag of potting soil. Each of her planters holds 1/9 of a bag of soil. How many planters can Maureen fill?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

16) 3/8 of the books in the Wetherfields' house are novels. 1/10 of the novels have been on the best seller's list. What fraction of the Wetherfield's books have been on the best seller's list for novels?

___ A) Add ___ B) Subtract

___ C) Multiply ___ D) Divide

Please STOP here. Tell the experimenter you are at the midpoint of the booklet. You will take the Peabody Picture Vocabulary Test now, and then return to finish Part C.

Instructions - Part C

Please read each numbered problem carefully. Four choices are given with each problem. Pick from among the four choices the one which you think goes the best with the problem and mark a 1 next to that letter on your sheet. Next, look at the three choices which remain. Pick from among these three choices the one which you think goes the next best with the problem and mark a 2 next to that letter on your sheet. It is not necessary to actually solve the problems. If you would like to do any writing, you may do so in the margins.

Please do the following sample problem.

Sample

Billy had 28 marbles in his collection. He won 9 marbles in his next game of marbles. How many marbles does Billy have now?

___ A) Billy had 28 marbles in his collection. He lost 9 marbles in his next game of marbles. How many marbles does Billy have now?

___ B) Sarah had done 6 of her math problems during study hall. She did 5 more problems after school. How many math problems has Sarah completed?

___ C) The kids made a batch of 12 brownies. Then they ate 4 of the brownies. How many brownies did they have left?

___ D) Billy had 17 marbles in his collection. He won 10 marbles in his next game of marbles. How many marbles does Billy have now?

Most people would say that D was the best answer, since D describes a situation in which you should add, and the story is very similar to the story in the sample problem. B would be the second best answer, since it also describes a situation where you should add, but the story is different.

The actual problems you will do are similar to these, although some problems may seem more difficult. Please do your best with all the problems. Answer them in the same way, marking your first choice with a 1 and your second choice with a 2. Give two answers for each problem. Please do not skip any problems.

Do you have any questions?

1) Hansel began the trip with 3/4 of a pound of bread. He used 1/4 of a pound of bread to mark the trail. How much bread did Hansel have then?

___ a) Hansel began the trip with 5/8 of a pound of bread. He used 1/3 of a pound of bread to mark the trail. How much bread did Hansel have then?

___ b) Hansel began the trip with 3/4 of a pound of bread. He found 1/4 of a pound of bread along the trail. How much bread did Hansel have then?

___ c) The potter had 1/2 a ton of clay at the beginning of the summer. Over the summer, the potter used 3/8 of a ton of clay. How much clay did the potter have left at the end of the summer?

___ d) Michelle had collected 2/3 of a set of antique china. Her mother gave her 1/6 of a set of the china for her birthday. How much china did Michelle have then?

2) Arlen mixed up 5/12 of a bucket of birdseed. He found he had enough birdseed to fill 2/3 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders?

___ a) The bricklayers bought 3/4 of a ton of bricks. They were able to build the walls to 5/8 of the desired height. How much brick was needed for the whole building?

___ b) 1/4 of all gas stations in the city have a garage for repairs. 5/6 of all garages do engine repairs. What fraction of all gas stations do engine repairs?

___ c) Arlen mixed up 7/8 of a bucket of birdseed. He found he had enough birdseed to fill 4/6 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders?

___ d) Arlen mixed up 5/12 of a bucket of birdseed. 2/3 of the birdseed was millet. What fraction of the bucket was filled with millet?



3) Grace had 3/6 of a pound of chocolate bits. She needed 1/6 of a pound of chocolate bits to make a batch of cookies. How many batches of cookies could Grace make?

___ a) Grace had 3/6 of a pound of chocolate bits. She bought 1/6 of a pound of chocolate bits so she would have enough to make a batch of cookies. What fraction of a pound of chocolate bits did Grace have then?

___ b) Grace had 5/8 of a pound of chocolate bits. She needed 1/3 of a pound of chocolate bits to make a batch of cookies. How many batches of cookies could Grace make?

___ c) The oil truck had 5/8 of a ton of oil in it. Each house gets 1/16 of a ton of oil. How many houses could get oil from the oil truck?

___ d) The school had computers for 1/10 of its classrooms. It bought computers for another 1/4 of the classrooms. What fraction of the classrooms in the school had computers then?

4) Margret had 2/5 of a gallon of ice cream. She gave 1/5 of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive?

___ a) The scout leader took 2/3 of the scout troop on a camping trip. 1/4 of the scouts who went earned their camping badge on the trip. What fraction of the scout troop earned their camping badges on that trip?

___ b) Margret had 2/5 of a gallon of ice cream. She gave 1/5 of a gallon of the ice cream to her sister, Anne Marie. How much ice cream did Margret have then?

___ c) The couple selling onion rings at the fair brought 3/4 of a truck load of onions with them. They sold 1/2 a truck load of onions. What fraction of a truckload of onions did they have left?

___ d) Margret had 1/3 of a gallon of ice cream. She gave 1/4 of the ice cream to her sister, Anne Marie. How much ice cream did Anne Marie receive?

5) Hemingway used 4/8 of a box of typing paper last week, while Orwell used 3/4 of a box of typing paper. How much more paper did Orwell use than Hemingway?

___ a) Hemingway used 2/3 of a box of typing paper last week, while Orwell used 7/8 of a box of typing paper. How much more paper did Orwell use than Hemingway?

___ b) 1/5 of the class got A's on the first exam. 1/8 of the class got A's on the second exam. How much more of the class got A's on the first exam than on the second exam?

___ c) Peter bought 1/2 a load of sand to fill his daughter's sand box. He was able to fill 2/3 of the sand box. How much sand would he need to fill the whole sand box?

___ d) Hemingway used 4/8 of a box of typing paper last week. He was able to write 3/4 of his latest short story. How much paper would he need to write the whole short story?

6) Cheryl won some money with a lottery ticket and was allowed to keep 2/3 of it after taxes. She put 1/3 of what she received in the bank. What fraction of the total amount of money that Cheryl won went into the bank?

___ a) Cheryl won some money with a lottery ticket and was allowed to keep 2/3 of it after taxes. She had already saved an amount that was equal to 1/6 of what she had won. How much money did Cheryl have then?

___ b) 1/10 of the people in the city were told to leave their homes because of deadly chemical fumes. 3/4 of the people told to leave actually left. What fraction of the people in the city left their homes?

___ c) Cheryl won some money with a lottery ticket and was allowed to keep 5/8 of it after taxes. She put 1/4 of what she received in the bank. What fraction of the total amount of money that Cheryl won went into the bank?

___ d) The electrician had enough wire for 1/5 of the building. She bought more wire for 1/3 of the building. How much wire did she have then?



7) Joey bought 7/10 of a mile of kite string so he could fly his kites. He wanted each kite string to be 1/4 of a mile long. How many kites can Joey put string on?

___ a) Arnie had 7/8 of a box of disks for his computer. He used 1/3 of the disks to write a chapter for his new book. What fraction of the box of disks did he use?

___ b) Joey bought 1/2 a mile of string so he could fly his kites. He wanted each kite string to be 1/10 of a mile long. How many kites can Joey put string on?

___ c) Joey bought 7/10 of a mile of kite string so he could fly his kites. He put 1/4 of the string on one kite. How much string did Joey put on the kite?

___ d) The Grays had 9/16 of a pound of Halloween candy. They wanted to give each child 1/20 of a pound of candy. How many children could they give candy to?

8) The box of cereal was 5/8 full. George poured 1/8 of a box of the cereal into his bowl. How much cereal was then in the box?

___ a) The box of cereal was 5/8 full. George poured 1/8 of the cereal into his bowl. What fraction of a whole box did George pour into his bowl?

___ b) The balloon man had 1/2 a bunch of balloons at lunchtime. Someone bought 1/8 of a bunch of balloons to make a bouquet. What fraction of a bunch of balloons did the balloon man have then?

___ c) The tree still had 2/3 of its leaves yesterday morning. Then the wind blew off 1/2 of the leaves that still remained. What fraction of all its leaves did the tree lose in the wind?

___ d) The box of cereal was 9/10 full. George poured 1/5 of a box of cereal into his bowl. How much cereal was then in the box?

9) Albert collected 3/8 of a pound of aluminum cans. His friend gave him another 1/8 of a pound of aluminum. How much aluminum does Albert have now?

___ a) Albert collected 4/6 of a pound of aluminum cans. His friend gave him another 1/3 of a pound of aluminum. How much aluminum does Albert have now?

___ b) The trio brought 3/4 of a box of their latest record with them to the concert. They sold 2/5 of a box of records. What fraction of a box did they have left?

___ c) Albert collected 3/8 of a pound of aluminum cans. He gave his friend 1/8 of a pound of the aluminum. How much aluminum does Albert have now?

___ d) After the hurricane 1/3 of the houses in the town were without power. Then a flooding stream knocked out the power in another 1/6 of the houses. What fraction of the houses in the town had no power then?

10) Marie Davis owns 2/3 of a real estate business. Her husband, Dick, owns 1/6 of the business. What fraction of the business do they own together?

___ a) Marie Davis owns 2/3 of a real estate business. Her husband, Dick, owns 1/6 of the business. How much more of the business does Marie own than Dick?

___ b) Rose cleaned 1/3 of the floors in the tower. Mitch cleaned 3/8 of the floors in the tower. What fraction of the floors in the tower have been cleaned?

___ c) Marie Davis owns 5/8 of a real estate business. Her husband, Dick, owns 1/3 of the business. What fraction of the business do they own together?

___ d) 1/3 of Jenny's pants are blue jeans. 1/2 of her pants are corduroys. What fraction more of her pants are corduroys than blue jeans?



11) Carla had 5/8 of a chocolate bar to eat for lunch. Her girlfriend gave her another 1/8 of a chocolate bar at lunchtime. What fraction of a chocolate bar did Carla have to eat for lunch?

___ a) The landlord had 1/3 of his apartments vacant and for rent. By the end of the day, he had rented 1/6 of the empty apartments. What fraction of all his apartments did he rent that day?

___ b) Carla had 4/6 of a candy bar to eat for lunch. Her girlfriend gave her another 1/6 of a chocolate bar to at lunchtime. What fraction of a chocolate bar did Carla have to eat for lunch?

___ c) Morris the cat earned 1/4 of a million dollars in his first movie. He earned 1/3 of a million dollars in his second movie. What fraction of a million dollars has he earned altogether?

___ d) Carla had 5/8 of a chocolate bar to eat for lunch. She gave her girlfriend 1/8 of her chocolate. What fraction of a chocolate bar did Carla give away?

12) 2/3 of the menu items at the new restaurant are desserts. 1/4 of the desserts contain chocolate. What fraction of the menu items are chocolate desserts?

___ a) Laundry detergent comes in 3/8 pound packets at the laundromat. The Instructions say you should use 1/2 of the packet for one load of wash. What fraction of a pound of detergent do you use for one load?

___ b) 5/6 of the menu items at the new restaurant are desserts. 3/5 of the desserts contain chocolate. What fraction of the menu items are chocolate desserts?

___ c) 2/3 of the menu items at the new restaurant are desserts. 1/4 of the menu items are beverages. What fraction of the menu items are either desserts or beverages?

___ d) Laurel has saved 1/2 of a box full of soda cans. Sue has saved 1/3 of a box full of soda cans. What fraction of a box of soda cans have they saved together?

13) The peanut man had 3/4 of a carton of peanuts. However, this was enough for only 1/4 of the people who wanted to buy peanuts. How many cartons of peanuts would the peanut man need to have enough for all the people?

___ a) The peanut man had 1/2 of a carton of peanuts. However, this was enough for only 1/3 of the people who wanted to buy peanuts. How many cartons of peanuts would the peanut man need to have enough for all the people?

___ b) Jane cut 1/2 a yard of blue fabric into triangles. She had enough blue triangles for 1/3 of the blocks in her quilt. How much blue fabric would Jane need to have enough for all the blocks?

___ c) The peanut man had 3/4 of a carton of peanuts. His competitor had 1/4 of a carton of peanuts. What fraction of a carton of peanuts did they have together?

___ d) The trumpet players sold 2/5 of the candy the band had to sell, while the clarinet players sold 1/4 of the candy. What fraction of the candy did the trumpet and the clarinet players sell together?

14) Martha and Emily were painting the kitchen. Martha painted 3/6 of the kitchen and Emily painted 1/3 of the kitchen. What fraction of the kitchen did they paint together?

___ a) The red pickup truck can haul 3/5 of a ton of wood, while the brown pickup truck can haul 1/3 of a ton of wood. How much wood can they haul togther?

___ b) 1/5 of the soda sold in the corner store is root beer flavored. 3/5 of the root beer soda sold is made by Hires. What fraction of all the soda sold is Hires root beer?

___ c) Martha and Emily were painting the kitchen. Martha painted 5/8 of the kitchen and Emily painted 1/4 of the kitchen. What fraction of the kitchen did they paint together?

___ d) Martha and Emily were painting the kitchen. Martha painted 3/6 of the kitchen, while Emily painted 1/3 as much as Martha painted. What fraction of the kitchen did Emily paint?



15) 1/3 of the grapes grown in New York state are green. 2/3 of the green grapes are sold as table grapes. What fraction of all the grapes grown in New York state are green grapes sold for the table?

___ a) Chris had 3/4 of a gallon of white paint. He had enough paint for 2/3 of the porch. How much paint would he need to paint the whole porch?

___ b) 2/5 of the grapes grown in New York state are green. 7/10 of the green grapes are sold as table grapes. What fraction of all the grapes grown in New York state are green grapes sold for the table?

___ c) 1/3 of the grapes grown in New York state are green. There are enough green grapes to fill 2/3 of the demand. What fraction of the grapes grown should be green grapes in order to fill the demand?

___ d) 1/4 of the freshmen males tried out for the football team this year. 1/3 of them made the team. What fraction of all the freshmen males made the football team?

16) A woman left 5/8 of her total fortune to her husband. He spent 1/4 of what he received on a trip to China. What fraction of the total fortune was spent on the trip?

___ a) A woman left 5/8 of her total fortune to her husband. The cost for a trip to China was 1/4 of the total fortune. How many times could the man take a trip to China?

___ b) There was 3/4 of a load of cement in the cement truck. The builders used 5/8 of the cement to build the cellar wall. What fraction of a load of cement was used to build the wall?

___ c) The carpet company had 1/2 a roll of red carpet. It was enough to cover 2/3 of the living room floor. How much carpet was needed to cover the whole floor?

___ d) A woman left 2/3 of her total fortune to her husband. He spent 1/5 of what he received on a trip to China. What fraction of the total fortune was spent on the trip?

17) John had 3/4 of a quart of milk. He gave his cat 1/8 of a quart of the milk. How much milk did John have then?

___ a) John had 7/8 of a quart of milk. He gave his cat 1/4 of a quart of the milk. How much milk did John have then?

___ b) The artist had 1/2 an ounce of vermillion red paint. She used 1/3 of an ounce of the paint in her new painting. How much vermillion red paint did she have then?

___ c) Cindy made 2/3 of a gallon of apple butter for Christmas presents. She was able to put 1/8 of a gallon of apple butter into each jar. How many jars could she fill?

___ d) John had 3/4 of a quart of milk. He gives his cat 1/8 of a quart of milk a day. How many days can he give the cat milk?

18) Wheat was selling for 3/10 of a dollar per pound in July. Corn was selling for 1/5 of a dollar per pound. What fraction of a dollar more per pound was wheat than corn?

___ a) Maura and John were sealing envelopes. Maura had sealed 2/5 of the envelopes, while John had sealed 1/3 of the envelopes. What fraction of the envelopes had they sealed together?

___ b) Wheat was selling for 3/10 of a dollar per pound in July. Corn was selling for 1/5 of a dollar per pound. What fraction of a dollar would a pound of wheat and a pound of corn cost together?

___ c) Wheat was selling for 1/4 of a dollar per pound in July. Corn was selling for 3/20 of a dollar per pound. What fraction of a dollar more per pound was wheat than corn?

___ d) Robin has read 3/8 of the papers for her course. Debbie has read 1/2 of the papers for the course. How much more has Debbie read than Robin?



19) The Japanese restaurant had 13/16 of a gallon of soy sauce. 1/12 of a gallon of soy sauce could fit into a small serving container. How many containers could be filled?

___ a) The Japanese restaurant had 13/16 of a gallon of soy sauce. 1/12 of a gallon of soy sauce could fit into a small serving container. How much soy sauce would be left after filling one container?

___ b) The Japanese restaurant had 9/10 of a gallon of soy sauce. 1/8 of a gallon of soy sauce could fit into a small serving container. How many containers could be filled?

___ c) The cobbler had 3/4 of a hide of leather. He needed 1/3 of a hide to make a pair of shoes. How many pairs of shoes could he make?

___ d) The holding tank 3/4 full of milk. The dairyman sold 2/5 of a tank of milk to customers coming to the farm. How much milk was left in the tank?

20) On December 1, the ice on the pond was 3/12 of a foot thick. On January 1, the ice was 5/8 of a foot thick. How much thicker was the ice in January than in December?

___ a) 2/5 of the seats in the theater are on the balcony. 1/3 of the balcony seats have a good view. What fraction of the seats in the theater are balcony seats with a good view?

___ b) Laura has raked 3/5 of the yard, while Greg has raked 1/4 of the yard. How much more of the yard has Laura raked than Greg?

___ c) On December 1, the ice on the pond was 3/12 of a foot thick. A hole in the ice showed that 5/8 of the ice had formed overnight. What fraction of a foot of ice had formed overnight?

___ d) On December 1, the ice on the pond was 1/6 of a foot thick. On January 1, the ice was 3/4 of a foot thick. How much thicker was the ice in January than in December?

21) 5/8 of the pieces in Aunt Martha's quilt are square. 1/3 of the squares are yellow. What fraction of the pieces in the quilt are yellow squares?

___ a) 3/6 of the pieces in Aunt Martha's quilt are square. 1/5 of all the squares are yellow. What fraction of the pieces in the quilt are yellow squares?

___ b) At Jim's house, 2/3 of the long distance phone calls listed on his phone bill are made out of state. 7/8 of these calls are made at night. What fraction of the long distance calls are made out of state at night?

___ c) 5/8 of the pieces in Aunt Martha's quilt are sqaure. 1/3 of the pieces are triangles. How many more of the quilt pieces are squares than triangles?

___ d) Ramona walked 10/12 of the route for the walkathon. Leslie walked 1/3 of the route. How much further did Ramona walk than Leslie?

22) Greta was making a down jacket and had 7/8 of a pound of down. This was enough to fill 1/3 of the jacket. How much down did Greta need to fill the whole jacket?

___ a) Greta was making a down jacket and had 2/3 of a pound of down. This was enough to fill 1/5 of the jacket. How much down did Greta need to fill the whole jacket?

___ b) Mr. Baldwin had 2/3 of a container of driveway sealer. He was able to seal 1/4 of his driveway. How much sealer would he need to do the whole driveway?

___ c) Greta was making a down jacket and had 7/8 of a pound of down. She used 5/8 of a pound of down to fill the jacket. How much down did she have left?

___ d) Hayley used 1/3 of a roll of film to shoot pictures of horses and 3/6 of a roll of film to shoot pictures of the autumn leaves. How much more film did she shoot on the leaves than on the horses?



23) Bill had collected 3/4 of a ton of food for the food drive so far. His next contribution was for 1/9 a ton of food. How much food had Bill collected then?

___ a) Bill had collected 3/4 of a ton of food for the food drive so far. He wanted to give each of the stricken villages 1/9 of a ton of food. How many villages did he have food for?

___ b) Bill had collected 7/10 of a ton of food for the food drive so far. His next contribution was for 2/12 of a ton of food. How much food had Bill collected then?

___ c) Kelley and Anne found 1/2 a bushel of pinecones yesterday. Today they found 1/3 of a bushel of pinecones. What fraction of a bushel of pinecones do they have now?

___ d) The ice cream shop had 8/10 of a gallon of hot fudge sauce. If they put 1/16 of a gallon of sauce on each sundae, how many hot fudge sundaes can they make?

24) Greg measured the area his feet and hand prints covered and found his hands covered 5/12 of a square foot. His feet covered 9/16 of a square foot. What fraction of a square foot did his hands and feet cover together?

___ a) In one hour, the patrons of the cafeteria drank 3/5 of a tank of regular coffee and 1/4 of a tank of tea. How much coffee and tea did they drink altogether?

___ b) Holly bought 3/4 of a pound of hazelnuts. This was 2/3 of the amount called for by the recipe. How much should she have bought to have enough for the whole recipe?

___ c) Greg measured the area his feet and hand prints covered and found his hands covered 5/12 of a square fool. The area his hands covered was 9/16 of the area his feet covered. What fraction of a square foot did his feet cover?

___ d) Greg measured the area his feet and hand prints covered and found his hands covered 3/8 of a square foot. His feet covered 4/9 of a square foot. What fraction of a square foot did his hands and feet cover together?

Please answer the following questions:

Age ____ Sex ___

Major ______________________

Year of graduation _______________

Mathematics courses taken in college:

Name of last mathematics course in high school __________________

__________________

If you are interested in earning a second credit, we would like to interview several people on materials similar to what you have seen today. If you would like to help us learn more about how people solve fraction problems, please see the experimenter to set up an appointment for an interview before you leave.