DOCUMENT RESUME

ED 299 146 SE 049 671

AUTHOR Hardiman, Pamela Thibodeau
TITLE Recognizing Similarities between Fraction Word Problems.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE May 88
GRANT NSF-SED-8113323
NOTE 38p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Development; College Mathematics; Concept Formation; Educational Research; *Elementary School Mathematics; Elementary Secondary Education; *Fractions; *Mathematical Concepts; Mathematics Achievement; Mathematics Education; Mathematics Tests; *Problem Solving; Secondary School Mathematics; *Word Problems (Mathematics)
IDENTIFIERS *Mathematics Education Research

ABSTRACT

Deciding how to approach a word problem for solution is a critical stage of problem solving, and is the stage which frequently presents considerable difficulty for novices. Do novices use the same information that experts do in deciding that two problems would be solved similarly? This set of four studies indicates that novices rely more on surface feature similarity. However, when certain aspects of the problem structure were held constant, novices correctly identified problems that would be solved similarly significantly more often. The results suggest: (1) the distinction of surface and deep features may not be rich enough for describing categorization of word problems, and (2) novice problem solvers are capable of recognizing some similarities of problem structure. A third level of classification of word problems is proposed to explain these results. (Author)

Recognizing Similarities between Fraction Word Problems

Pamela Thibodeau Hardiman

Department of Physics

University of Massachusetts

Amherst, MA 01003

May, 1988

Acknowledgements

This research was supported in part by NSF grant SED-8113323. I am grateful to the members of my dissertation committee for their time and advice: Arnold Well, Alexander Pollatsek, Carolyn Mervis, and George Cobb. I would also like to thank Jose Mestre for his comments in writing this report. Arnold Well assisted in the running of Group 1, while Leslie Arriola was instrumental in the running of both Group 2 and the eighth graders.

Abstract

Deciding how to approach a word problem for solution is a critical stage of problem solving, and i. the stage which frequently presents considerable difficulty for novices. Do novices use the same information that experts do in deciding that two problems would be solved similarly? This set of four studies indicates that novices rely more on surface feature similarity. However, when certain aspects of the problem structure were held constant, novices correctly identified problems that would be solved similarly significantly more often. The results suggest: 1) the distinction of surface and deep features may not be rich enough for describing categorization of word problems, and 2) novice problem solvers are capable of recognizing some similarities of problem structure. A third level of classification of word problems is proposed to explain these results.

Deciding how to approach a word problem for solution is a critical stage of problem solving. While problem solvers often have difficulty carrying out a proposed computation correctly, planning the computation is often the most critical step in solving a word problem. For single-step fraction arithmetic word problems containing two numbers, the topic of this set of studies, the only possibilities are to add, subtract, multiply, or divide the two numbers contained in the problem. Hence, deciding on a solution approach involves a classification into one of four categories.

What information in a word problem does a problem solver use in order to make this classification, and how is that information used? One way in which this question can be approached is to ask problem solvers to decide what kinds of problems can be solved in the same way. The issue then becomes one of what types of similarities constitute a sufficient basis for deciding that two problems would be solved similarly, and whether these criteria are different for experts and novices. Studies of problem solvers in physics [1,2] have suggested that the answer to the latter part of the question is yes: novices and experts attend to different types of information when classifying problems according to solution similarity. Novices appear to classify problems mainly with respect to surface similarities, such as the objects involved in the problem, whereas experts classify problems on the basis of solution principle. Similar observations have been made with good and poor problem solvers in the same grade in school: poor problem solvers are often misled by surface structure or "pseudo" similarities [3,4]. I will refer to this claim that novices attend to the surface structure, or the details of the problem story, in contrast to the principles of solution when judging whether two problems would be solved similarly, as the <u>surface feature hypothesis</u>.

In this paper it will be argued that the surface feature hypothesis presents too limited a view of how novices judge solution similarity. The process of understanding a word problem can be viewed as a specialized form of text processing, in which the problem understander attempts to construct a model of the objects in a problem, the relations between them, and how the relations change. In the usual formulation of the surface feature hypothesis, the surface features involve literal aspects of the problem, the objects, the technical terms, and the configuration described in the problem. For example, physics novices in the Chi et al study [1] grouped togther as similar problems which contained springs, those which mentioned friction, and those with a block on an inclined plane. However, it seems reasonable that novices also consider deeper similarities of relations, noting similar types of actions, like "giving-to" or "taking from", similar formulation of questions, and related words or phrases. While noting such similarities does not necessarily indicate an understanding of the deep structure, or principles of solution of a problem, clearly more effort is involved than in simply noting that the problems contain the same objects. Hence, it might be useful to consider some intermediate level analysis of problem structure.

The studies reported here attempt to show that the surface feature hypothesis cannot offer a complete account of the judgments of experts and novices, and that aspects of the structure of the problem should be considered. In addition, several specific hypotheses (to be described later) concerning the domain of fraction word problems will also be tested. The four studies reported here concerned: 1) assessment of problem difficulty, 2) judgment of solution similarity by experts and nonexperts, 3) how judgments of solution similarity may be facilitated, and 4) replication of results with a younger population.

The Problem Domain

Arithmetic word problems containing fractions were chosen as the domain of investigation for several reasons. First, the domain is rich, yet types of problems can be well defined. A second reason for choosing this domain is that despite the fact that most people have had considerable instruction in solving problems with fractions, a large proportion of both high school students [5] and adults [6] are unable to solve fairly routine word problems involving fractions. This lack of understanding of fraction concepts is an impediment to gaining a mature understanding of rational numbers. Thus, investigation of problem solving in this domain can begin to provide the data-base needed for educational reform.

Why are arithmetic word problems that contain fractions difficult to solve? First, let us consider two explanations, which on the surface appear to be cogent, but which do not hold up under closer scrutiny. Then a more plausible cognitive explanation will be offered.

Deceptive explanation 1: All word problems are more difficult to solve than the corresponding computational problems because in order to compute the answer to a word problem, the solver must first translate the words into an equation. In contrast, no extra step is required to solve a computational problem. This explanation would imply that performance in solving word problems containing fractions should be lower than performance in solving similar computational problems. However, results from the National Assessment of Educational Progress [5] for 13 year olds clearly argue against this explanation. Although it is true that performance is slightly worse for whole number addition and multiplication word problems than for comparable computation problems (5% difference), this trend is not observed for all fraction problems. For fraction addition problems, performance is actually slightly lower for computation problems than for comparable word problems (47% versus 54% correct sclutions). In contrast, for fraction multiplication problems, performance is considerably better for computation problems than for comparable word problems (69% versus 17% correct solutions). Thus, for fraction problems, these performance differences are neither constant in size nor in the same direction.

Deceptive explanation 2: Fractions are more complex numbers and add incrementally to the difficulty of a problem. Fractions, being composed of two parts, numerator and denominator, are more complex numbers than whole numbers because the magnitude of the fraction is dependent on the relation between the two whole numbers that comprise the fraction. This means that fractions do not have transparent ordering properties, making it difficult to gauge whether the answer makes sense. Hence, performance on all fraction problems should be lower than performance on all whole number problems. The size of this effect should be constant with operation.

Although these opaque ordering properties probably adversely affect the development of computational skills [7], and thereby indirectly influence performance on word problems, the mere presence of fractions in a word problem does not necessarily make it harder to understand the problem situation. In fact, the question of whether or not the presence of fractions has a direct influence on understanding a problem situation is an experimental question that is addressed in Study 1.

Cognitive explanation: The schemas that solvers have developed for solving whole number problems are appropriate for solving some types of fraction problems, but not others. Understanding a problem text requires the development of a situational model based upon the reader's understanding of the conceptual relations among quantities in the text [8]. In other words, the reader must correctly infer the deep structure. Kintsch and Greeno claim the problem solver has schemata for various types of relations among quantities, which may be cued and substantiated during the course of problem understanding. Hence, problem solvers may have difficulty if they associate certain effects rigidly with specific operations, and cue inappropiate schemata.

Fischbein and associates [9] hypothesize that arithmetic operations are intuitively linked to primitive behavioral models (or schemata) developed during initial school instruction. Such models can facilitate problem solving when appropriate, but inhibit problem solving when inappropriate. The constraints of the model can force the choice of an inadequate operation.

The argument made here is that the type of numbers in a problem, whole numbers or fractions, influences the kinds of units, or quantities, allowed, and hence the possible types of relations among quantities. It is not always possible to substitute fractions for whole numbers (and vice versa) and retain the similar relations among the problem elements. Hence, only for certain kinds of problems are the schemas that problem solvers have developed for understanding whole number problems appropriate. In cases where there is a lack of parallelism in problem structure between whole number and fraction problems, novices are likely to misinterpret the fraction problem statement because of their experience with whole number problems.

To illustrate how units influence problem structure, consider the following fraction multiplication problem:

"Margret had 4/5 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?"

It is possible to substitute a whole number for the operand, i.e., replace "4/5 of a gallon" with "4 gallons," and retain the problem structure. The relations between the problem elements are not influenced by whether Margret had 4 gallons of ice cream or 4/5 of a gallon of ice cream. However, it is not possible to replace the operator, the "1/5" in "1/5 of the ice cream," with a whole number and still have a multiplication problem; "She gave 2 of the ice cream" is not a meaningful construction. This is true in general for all multiplication problems in which the operator is a fraction.

On the other hand, if "1/5 of the ice cream" is misread as "1/5 of a gallon of ice cream," then a whole number can be substituted for the 1/5, and the statement might appear to be a subtraction problem because something is given away.

> "Margret has 4/5 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 Anne Marie receive?"

However, this statement actually contains no unknown. To be a subtraction problem, the question would have to be about Margret, not Anne Marie, as in "How much ice cream did Margret have then?" The novice may unconciously misread both the unit information and the question if he or she is actively trying to recognize the problem as one of a familiar type. In whole number word problems, the "giving away" scheme commonly appears in subtraction problems, and never in multiplication problems. Thus, the "gave" in this multiplication problem may be inappropriately interpreted as a cue for subtraction, particularly by novices.

In contrast to multiplication problems, the structure of addition and subtraction problems makes it possible to substitute fractions for whole numbers (and vice versa) freely without substantially altering the meaning, as long as the fractional units remain sensible (e.g., "2 cats + 3 dogs" makes sense, but "1/2 a of cat + 1/3 of a dog" ordinarily does not). By this reasoning, the interpretation of addition and subtraction word problems with fractions should not be substantially more difficult than the interpretation of similar problems with whole numbers. The same should be true for measurement division problems, since substitution is possible for these problems (which ask, "How many units of a given measure are contained in a given quantity?"). Therefore, novices should be able to recognize and correctly interpret addition, subtraction and measurement division problems containing whole numbers or fractions equally well. Since the interpretation of fraction multiplication problems presents unique difficulties, it is likely that novices will experience difficulties with fraction multiplication problems in deciding what operation should be performed and how to set up the problem for solution, regardless of whether or not they could actually compute the answer.

Problem Types Used in the Present Studies

In order to make a systematic study of types of word problems, the current research focuses on the subset of fraction word problems that commonly appear in textbooks and that were included in the NAEP evaluation; each problem contained two numbers, could be solved with one step, and had the result of the operation as the unknown. Addition, subtraction, multiplication and division problems were all included.

Since, numerous studies of children solving addition and subtraction word problems have shown considerable variability in problem difficulty for problems that are solved using the same operation [10 - 13] two types of word problems requiring each operation were used. An attempt was made to choose types that were similar in certain respects across operation, although no comparisons were actually made across operations. For the operations of addition, subtraction, and multiplication, one problem, termed an Active problem, involved an action which was integral to the problem storyline. For example, in the following subtraction problem, a portion of an original quantity is disposed of:

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

The second of the two problems contained no actions, but described and asked about a relationship between the two problem elements. This was termed a Passive problem, as in:

> Ernest had 3/5 of a box of typing paper. George had 4/5 of a box of typing paper. How much more paper did George have than Ernest?

Here, two parties are compared and neither party loses any of the original quantity of paper that he possessed.

Unfortunately, the Active-Passive distinction is less applicable to division problems, which can be classified as measurement and partitive problems. For ease in discussion, measurement division problems, such as the example below, will be referred to the active problem for division:

> 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?

Partitive problems were referred to as Passive division:

> Arlen mixed up 2/3 of a bucket of birdseed. He found he had enough to fill 1/3 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders?

A third dimension of the problems that was manipulated in this set of studies was whether the problems contained fractions or whole numbers. No problem mixed fractions and whole numbers.

General Description of Subjects

Two groups (Group 1 and Group 2) of undergraduate students enrolled in psychology classes at the University of Massachusetts and one group of eighth grade students (Eighth graders) enrolled at the Frontier Regional Junior High School in South Deerfield, Massachusetts participated in this set of studies. All subjects were given the Peabody Picture-Vocabulary Test (PPVT) to ascertain that any poor performances observed could not be attributed merely to poor verbal skills.

Group 1 was composed of 47 subjects, 16 males and 31 females, who had stanine scores on the PPVT of 6 or higher (of 9). These subjects participated in Studies 1 and 2.

The 29 females and 28 males in Group 2 had an average PPVT score of 5.0. These subjects participated in studies 1 and 3.

The eighth graders were enrolled in the top three of the five eighth grade mathematics classes in the school; two of the classes were Algebra I classes, while the third was a standard eighth grade mathematics class. The students were all taught by the same mathematics teacher. There were 20 males and 32 females, whose average age was 13 years 6 months and average PPVT score was 6.4. The eighth graders participated in study 4.

In the preliminary analyses for each study, sex was included as a variable. In no case was there a main effect or major interaction involving sex, so the sex variable was dropped in further analyses.

Study 1: Assessment of Problem Type Difficulty

The general purpose of Study 1 was to assess differences in levels of understanding for the set of fraction problem types among college students. In a replication of Study 1 with Group 2, whole number problems were included as well, in order to test whether the complexity of fractional numbers (deceptive explanation 2) could be ruled out as an explanation for the poor performance displayed on fraction problems.

Subjects in Group 1 were presented with an active and a passive word problem that contained two fractions for each of the four arithmetic operations, for a total of eight problems. The subject was instructed to indicate which one of the four arithmetic operations should be performed on the two fractions given in the problem to solve it. Given the NAEP performance on word problems cited earlier, it was expected that the college students would have particular difficulty with the multiplication problems, and possibly the division problems (which were not included in the NAEP). Group 2 was given the same task with a different set of eight fraction problems, as well as a set of eight problems that each contained two whole numbers.

Procedure. The subjects were run in groups of no more than six people. This task was presented as the second section of the problem booklet that each subject received. The subjects were to indicate whether they would add, subtract, multiply, or divide the two numbers given in the problem to obtain the correct answer. They were told they could actually compute the answer if they wished, but it was not required.

Results and Discussion. The results for Groups 1 and 2 will be reported separately, since Group 2 was given a different set of eight fraction word problems than Group 1. However, note that the overall mean percent correct for Groups 1 and 2 on the fraction problems were similar: 67% correct versus 69% correct. As will be shown presently, the general pattern of results is similar as well. Thus, these results are replicable over different sets of problems and are not limited to the particular problems used.

Insert Table 1 about here

The 4 (operations) x 2 (activeness/operation) ANOVA indicated subjects were not equally good at choosing the operation needed to solve problems requiring each of the four operations, $F(3,141) = 45.95$, $p < .0001$, and $F(3,168) = 40.78$, $p < .0001$. The mean percent correct scores were (from highest to lowest): addition (Group 1 -93%, Group 2 -92%), subtraction (85%, 86%), division (57%, 60%), and multiplication (33%, 35%). For both groups, all pairwise differences between operations were significant with a Bonferroni test ($n = 6$, $p < .008$), except that between addition and subtraction where performance was fairly high in both cases.

Note that the ordering of performance on multiplication and division items is opposite to the order in which these two operations are taught in school. In fact, a small informal survey of middle school teachers suggests this result is contrary to their expectations; they believed division with fractions to be the hardest operation to understand. Certainly, the computational algorithm for fractional division is more complex than that for multiplication, bu. this factor does not seem to make fractional division word problems harder to understand.

In addition to sizable differences between operations, there were considerable differences within operations as well. Subjects do not understand equally well all problems which require the same operation, as indicated by the significant Activeness within Operation effect, $F(4,188) = 9.62$, $p < .0001$ and $F(4,224) = 8.87$, $p < .0001$. Although there was no difference between the active and passive addition problems (93% versus 91%, Groups 1 and 2 combined), for subtraction, performance was better on the active problems than on the passive problems (94% versus 77%). The same was true for division (80% active versus 45% passive). In contrast, for multiplication, performance on active problems was not as good as performance on passive problems (26% versus 42%). Thus, the better performances were generally, but not uniformly, associated with the active problems.

In general, performance on the fraction problems was rather low. One possible explanation for this such poor performance is that the students do not really understand the whole number operations, and the complexity of the fractional numbers adds to the diffculty of a problem (deceptive explanation 2). If this explanation is adequate, then one should observe a slightly higher and parallel pattern of means for the whole number items.

The performance of Group 2 on the whole number problems clearly rules out this explanation. The mean percent correct on the whole number problems was 98% (versus 69% for fractions), with scores ranging from 95% to 100% correct for the eight problem types. The difference between performances on whole number and fraction problems was overwhelmingly significant, $F(1,56) = 234.39$, $p < 0.0001$, as would be expected. However, explanation 2 is ruled out by the fact that the performances are clearly not parallel, as seen in the interaction of Number type and Operation, $F(3,168) = 39.04$, $p < .0001$, and of Number type with Activeness/Operation, $F(4,224) = 9.09$, $p < .0001$. Clearly, the pattern of results for problems with whole numbers is dissimilar to that for problems with fractions. Hence, although the complexity of fractions may may involved in the explanation, it cannot be the whole explanation.

The rejection of deceptive explanation 2 leaves us free to consider the cognitive explanation; subjects have difficulty with fraction multiplication and division problems because the structure of fraction problems is not appropriate for whole number schemata. Subjects' actual responses to the multiplication and division fraction problems (see Table 1) offer further support for this notion. Comparing the performances on the active and passive problems, it can be seen that not only do the rates of correct performance differ, but the typical wrong answers differ as well. For multiplication, 53 of the 105 subjects (Groups 1 and 2 together) chose subtraction as the operation needed to solve the active multiplication problem, while for the passive multiplication problem, 50 of the 105 indicated division should be used. For division, most (82 of 105) subjects correctly identified the operation needed to solve the active problem, while many (42 of 105) thought the passive problem should be solved with multiplication. The active and passive forms are clearly not treated by subjects in the same way, suggesting structural differences related to Activeness, as well as to the presence of whole numbers. These results agree with the conclusions of [14], who suggested that students be taught different kinds of multiplicative structures, not just computational algorithms.

Study 2: Judgment of Solution Similarity

The primary purpose of Study 2 was to test the surface feature hypothesis: Do poor problem solvers consistently categorize problems on the basis of surface features (objects in the problem), while good problem solvers consistently sort on the basis of deep structure (operation)? If not, then do poor problem solvers tend to sort on the basis of surface features when they do not understand the deep structure? Study 1 provides the basis for predicting which types of problems adults are likely to have difficulty constructing a correct deep structure representation, and hence would be likely to misperceive deep structure similarity.

The task used in study 2 was a variation on the oddity task used by Rosch and Mervis in several of their studies [15-17], in which subjects are given several objects and must decide which object does not belong to the set. In the present study, subjects were given a standard problem with four alternatives. There were equal numbers of addition, subtraction, multiplication, and division standards. The subjects were to determine which two of the four alternatives should be solved similarly to the standard. For each standard problem, there was one alternative matching in: 1) both surface structure and operation required (B), 2) only the operation required (O), 3) only the surface structure (S), and 4) neither dimension (N). The Surface Feature hypothesis predicts that there should be considerable agreement among good problem solvers and poor problem solvers in judging similarity of solution when two problems both require the same operation for solution and have a similar storyline, so all subjects should consistently choose the B alternative. It is for the second choice that the hypothesis makes different predictions; poor problem solvers should consistently classify problems as similar which only share surface structure characteristics, always choosing the S alternative. Good problem solvers should never make this error, but should always choose the alternative which has deep structure similarity, or the O alternative.

In addition to varying the operation, two other types of manipulations were made in the set of stimulus problems: 1) the difficulty of performing the actual computations with the specific numbers in the problems - in easy problems, both numbers had the same denominator, while in hard problems they had different denominators which did not have an easily computed common denominator, and 2) the Activeness of the standard problem, active or passive.

Procedure The problems were presented in a booklet with wide margins for making any computations, if desired. They were intermixed with other items that required the same type of judgment. The subject was instructed to mark the two alternatives which would be solved similarly to the standard. They were told they did not have to solve the problems, but could do do if they felt this would help. There were 16 problems altogether: 4 operations by 2 activeness types by 2 levels of difficulty of computation.

The subjects were run in small groups of no more than six people. All subjects had a break to do the PPVT at the midpoint of the task. At the end, they did the task described in study 1 to determine expertise level; here, they were to state the operation needed to solve each of 8 fraction word problems. On the basis of the scores from this task, subjects were placed in either the good problem solver (0 - 1 errors) or one of the three poor problem solver (2, 3, and 4-5 errors) groups for analysis. A random eight subjects were dropped, in order to make four equally sized groups.

Results and Discussion. There was a main effect of error level among the four groups, $F(3,36) = 4.42$, $p = 0.0096$; the good problem solver subjects (0-1 errors) performed better overall (84% correct) than the less expert subjects (2, 3, and 4-5 errors, 66% correct), $t(38$, one tailed$) = 3.37$, $p < 0.001$. There were no significant differences among the three poor problem solver groups.

As predicted, both good problem solvers and poor problem solvers choose the B alternative a high proportion (89%) of the time. The B alternative was chosen more often than the $\underline{O}$ alternative (53%), $\underline{F}(1,36) = 94.67$, $\underline{p} < 0.0001$. Thus, the addition of a match in surface structure facilitated the decision that a problem would be solved similarly to the standard. There was no difference among the groups in the amount of facilitation provided by the additional match in surface features.

Cue overlap of both surface features and deep structure is clearly relevant to making judgments of similarity of solution for both good problem solvers and poor problem solvers; 600 of the 764 pairs (78%) of alternatives chosen by all subjects were pairs in which both alternatives had a feature (or features) in common with the standard (i.e. were $\underline{B}/\underline{O}$, $\underline{B}/\underline{S}$, or $\underline{O}/\underline{S}$ pairs). When subjects made one incorrect selection, 62% of the time the incorrect alternative chosen was the surface only alternative. Thus, subjects tend to err by choosing problems with obvious, but misleading characteristics of similarity.

Although the tendency to judge solution similarity on the basis of surface features is quite strong, similarity of surface features is not the only source of information that poor problem solvers use. As can be seen in Figure 1, poor problem solvers correctly chose the alternative 47% of the time that matched only in deep structure. Although the good problem solver subjects correctly chose the $\underline{O}$ alternative more frequently than the poor problem solvers (71% of the time), they did not universally judge solution similarity on the basis of deep structure, but sometimes used surface features. These results suggest that although the Surface Feature hypothesis is generally correct, it does not account for the data as well as might be hoped, suggesting that other factors may be involved in these descisions.

Insert Figure 1 about here

A more detailed look at the results reveals that the frequency with which subjects chose the correct alternatives (B and O) varied with operation, $F(3,108) = 12.90$, $p < .0001$. Overall, the pattern of differences among operations was slightly different from that in study 1; correct alternatives were chosen most often for addition (81%), followed by subtraction (70%), multiplication (70%), and division (67%). Operation interacted with Alternative choice, such that the B alternative was chosen a high proportion of the time for all operations (mean 89% correct), while rate of choosing O varied considerably over operations (addition - 71%, subtraction - 52%, multiplication - 54%, division - 34%). Thus, most of the errors were the result of failures to choose the alternative that had a different storyline, but matched the standard in operation (i.e., O) .

Contrary to most teachers' intuition, the difficulty of the numbers in a problem did not influence ability to make a judgment of solution similarity. However, this result does corroborate previous NAEP results of no consistent effect of the difficulty of numbers. From this point, it will be assumed that the important factor is whether there are fractions in a word problem, and relatively simple fractions will be used.

Study 3: The Role of Activeness in Judgments of Solution Similarity

Judgments of solution similarity can be facilitated by similarity of storyline, as indicated in Study 2. However, the source of this facilation may not be the match in superficial characteristics, such as the characters and setting; rather, similarity in the types of words, actions, and situations that can occur may be sufficient to produce such facilitation.

For example, consider the following two active multiplication problems:

Mary cooked a 3/4 pound steak for dinner. She ate 1/3 of the steak. How much steak did she eat?

Tom found 1/4 of a bottle of glue. He used 3/4 ot the glue building a birdhouse. How much glue did he use?

Although neither the characters, the objects involved, nor the actions that occur are the same in these two problems, they are similar in that the same type of action occurs. Both problems concern the size of a fractional portion of the original quantity that has been consumed. These two problems should be contrasted with a third passive multiplication problem which is dissimilar, in that no consumption of or separation from the original quantity occurs:

7/8 of the sandwiches the waitress delivered were hamburgers. 1/4 of the hamburgers were cheeseburgers. What fraction of the sandwiches served were cheeseburgers?

Note that, like the two problems above, this problem singles out and names a fractional portion of an original quantity, but the fractional portion is not "acted upon." Although the first two and the third problem are similar in that they require the same operation for solution, similarity in a condition where only the operation used to solve the problem is the same seems harder to recognize because of its abstractness than that which occurs when both the problem type and the corresponding types of actions are the same, as in problems 1 and 2. Therefore, in study 3, the hypothesis tested is that a match in Activeness facilitates the judgment of solution similarity.

In order to investigate whether a match in activeness facilitates the judgment of solution similarity, a task was designed in which subjects were to choose which one of four alternatives (requiring addition, subtraction, multiplication, and division for solution) required the same operation for solution as the specified standard. The sets of alternatives were structured in the following manner: a) all alternatives were of the same type, active or passive, b) they employed the same characters and objects, and c) they had similar story lines. For example, one set of alternatives was:

Addition

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?

Subtraction

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

Multiplication

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

Division

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

The standard had a different story line and could either match or mismatch the alternatives in activeness. Examples of standards that match and mismatch the multiplication alternative above are in activeness:

Matching

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?

Mismatching

1/3 of the people in the survey were Republican. 3/4 of Republicans were male. What fraction of the people in the study were male Republicans?

Thus, the subject read the standard, and then choose the one of the four alternatives that would be solve similarly. There were 16 items on this task: 4 operations x 2 types of standards(active or passive) x 2 matching or mismatching (the standard in activeness) sets of alternatives.

Procedure The Group 2 college students participated in this task. They were given the problems printed in a booklet, with two items on a page. There was room allowed for any computations subjects wanted to make. The subjects were to select the alternative that would be solved in the same way as the standard. They were run in small groups of no more than six people and were self-paced. This task was the first of three tasks performed by group 2.

Results and Discussion The results of Study 3 indicate that similarity of problem structure due to a match in activeness does facilitate the decision that two problems require the same operation for solution. The main effect of a match in activeness was highly significant, $F(1, 56) = 28.00$, $p < .0001$: subjects chose the alternative that required the same operation as the standard more often when the problems matched in activeness (71% correct for matching versus 55% correct for mismatching). However, the size of this effect differed with operation, $F(3,168) = 9.41$, $p < 0.0001$ (see Figure 2). A match in activeness provided the most facilitation for subtraction (73% correct for matching vs 40% correct for mismatching) and division (71% vs 44%) items, a smaller facilitation for addition items (88% vs 80%), but no facilitation for multiplication items (54% vs 54%). This pattern suggests that a match in activeness may provide for a larger facilitation effect for problems that students have some difficulty understanding, such as subtraction and division, but not extreme difficulty, such as multiplication.

Insert Figure 2 about here

A somewhat better understanding of the conditions under which facilitation occurs can be gained by considering activeness within operation, since there was an interaction of Match (in activeness) and Activeness within Operation, $F(4,224) = 17.67$, $p < 0.0001$. Facilitation occurred for all Activeness within Operation combinations except two: performance was actually somewhat better in the mismatching condition for passive addition problems (86% matching vs 95% mismatching), $t(56) = -2.17$, $p = 0.0319$. Given the high rate of correct responding for addition problems, it is probable that this difference is simply due to chance. Performance was also better for active multiplication items in the mismatching condition (37% matching vs 81% mismatching), $t(56) = -5.19$, $p < 0.0001$. An examination of the wrong selections suggests that this difference is not due to chance, since a number of these responses were the subtraction alternative (23 of 57 responses or 40%). Subjects probably chose the subtraction alternative in the matching condition because of the similarity between active subtraction and active multiplication (see Table 1). Both types of problems concern the removal or consumption of some portion of a set: in subtraction, the portion is indicated in the units of the original quantity, whereas in multiplication the portion is a fraction of the original quantity. Subjects most likely cued on this salient aspect of the problem and incorrectly chose the subtraction alternative. This confusion could not arise in the mismatching condition because passive subtraction and passive multiplication are quite dissimilar.

The results of this study suggest that problem solvers are able to utilize features other than surface similarity in order to judge that two problems would be solved similarly: a match in problem type also facilitates similarity judgments. Hence, a third level of problem structure will be proposed here to account for these results, termed intermediate structure. Problems with the same intermediate structure are solvable by the same operation, can be expressed using the same algebraic sentence, and are all active or all passive.

There are several ways in which a match in intermediate structure might facilitate the judgment that two problems would be solved similarly. Problems that match in intermediate structure share many features, including common patterns of actions, such as "giving-to" or "-from" (see [8] for other types of action patterns), similar formulation of questions, and the use of closely related words or phrases, as "gave away", "spent", and "lost." Similarity in these types of features is often a sufficient cue to correctly judge solution similarity. Note that it is not necessary that problems share "key words," as most of the standards were not paired with alternatives that had the same key words. The similarity perceived seems related to the meaning of the words, not the actual words.

Although learning to recognize problems with different intermediate structures might be of limited value to an expert problem solver, acquiring this skill should be of considerable value to less expert problem solvers. By learning what kinds of cues indicate specific subtypes of problems, they should become more able to categorize correctly on the basis of operation .

Study 4: Replicability with a Younger Population

One objection which could be raised concerning the first three studies is that college students who are poor problem solvers may be different from true novices, in that college students have had considerably more opportunity to practice and apply inappropriate problem solving stratgies. In order to determine whether this objection has any justification, the three studies were repeated with a younger population of subjects.

These subjects were eighth grade students, who were the youngest students available who had completed all instruction in fractions. In order to show that older poor problem solver college students are not simply poor problem solvers who behave neither like good problem solvers or novices, it is necessary to show that the patterns of performance are similar for eighth graders and adults. Although the overall levels of performance might be higher for the more experienced college students, the patterns of correct and incorrect responses should correspond.

Procedure The materials were printed in written booklets with large type (typewriter size) and wide margins for making any computations they wished to make. The students were told they did not have to actually solve any of the problems, but could do so if they thought it would be helpful.

The eighth graders were given the same tasks that the adults received in studies 1 and 3, and a task similar to that used in study 2. They received tasks 1 and 3 first, and were allowed one 45 minute school period for this session. In a second 45 minute session held a day or two later, they received task 2.

Results and Discussion

Task 1 - Relative difficulty of problem types. The overall proportion correct for eighth graders was significantly lower than that of the college students, $F(1, 105) = 20.98$, $p < 0.0001$. However, there were no significant interactions with age. Thus, although adults were correct more often than the eighth graders in indicating which operation is appropriate to solve a word problem, they had difficulty with the same types of problems. In addition, a comparison of Tables 1 and 2 indicates that when eighth graders err, they tend to make the same types of errors as adults, i.e., when they erred, they chose the same wrong operations as adults.

Insert Table 2 about here

These results suggest that eighth graders and adults employ similar strategies to determine what operation to perform to solve a problem. Adults' greater experience may provide some advantage in this task, but it has little effect on the types of errors that are made.

Task 2 - Judging solution similarity. The eighth graders received a similarity judgment task that was slightly modified from the task that the Group 1 adult subjects received, such that all pairs of correct and incorrect operations were represented in the alternative sets. This was done in order to determine whether it is possible to better predict the types of situations in which poor problem solver subjects would be likely to err, since they do not rely on surface structure similarity alone.

The results of Study 2 with the adults indicated that the problem type of the standard has an important influence on the correctness of similarity judgments. In addition, the types of alternatives that are contrasted with the standard may influence the decision as well. For example, given an active multiplication problem, a subject might err by chosing the surface feature only alternative if it is a subtraction problem, but not if it is an addition problem, because multipli.ation and subtraction are confused more often than multiplication and addition.

Performance on task 1, in which subjects decided which operation would be used to solve a number of problems, was used to obtain predictions of when subjects would err on the similarity judgment task. If on task 1, a subject said a problem should be solved using an incorrect operation, it was predicted that subjects would err on task 2 when the standard was of the same type as the task 1 problem and the S alternative would be solved using the incorrectly stated operation. In other words, subjects should consistently confuse operations between tasks.

For the eighth graders, performance on task 1 correlated with performance on task 2, $r = 0.571$, $t(51) = 4.96$, $p < 0.001$. Thus, eighth graders tend to be predictable in the types of confusions they make. This study suggests that subjects tend to make surface feature errors when they do not have a solid understanding of the operations with fractions.

Task 3 - Facilitating Judgments of Solution Similarity. The eighth graders were given the same matching task that the adults were given in study 3. They performed nearly as well the adults on the matching task (58% correct for eighth graders vs 62% correct for adults): this difference was not significant, $F(51) = 2.64$, $p = .1105$. As with adults, there was a main effect of match, $F(1,51) = 19.75$, $p < .0001$ (65% correct for matching versus 51% correct for mismatching), a main effect of operation, $F(3,153) = 41.07$, $p < .0001$, and an interaction of match and operation, $F(3, 153) = 5.46$, $p = .0014$. There was only one significant interaction involving age, that of age by match by operation, $F(3,153) = 4.38$, $p = 0.0055$. As can be seen by comparing Figures 2 and 3, the largest point differences occurred for matching subtraction problems (52% eighth grade vs 73% adults) and mismatching multiplication problems (39% eighth graders vs 54% adults). These differences seem mainly due to eighth graders having more difficulty than adults in distinguishing active subtraction and active multiplication problems. Thus, when poor problem solver adult performance differed from the novice eighth grade performance, the adults performed better: poor problem solver adults do not make peculiar and different kinds of errors from novices.

Insert Figure 3 about here

The results of these three replications with eighth graders suggest that the performance of younger poor problem solvers in categorization tasks does not differ appreciably from that of older poor problem solvers. Therefore, the behavior of older poor problem solvers can be considered representative of poor problem solvers in general, at least in this content domain. It is possible that older students have relatively few opportunities to practice problem solving skills with fractions, as there is usually no further direct instruction in fractions after the seventh grade. However, in either case it is interesting that the behavior of older poor problem solvers is not appreciably different from that of younger poor problem solvers.

General Discussion

The studies reported here imply that the the surface feature hypothesis does not adequately account for how poor problem solvers judge problem similarity: poor problem solvers do not consistently use similarity of surface features as a basis for a judgment of solution similarity. It is true that poor problem solvers tend to rely on surface feature similarity, particularly when they have difficulty understanding the operation, but this tendency is not consistent, as shown by the eighth graders' performance on task 2. Hence, even a weaker form of the Surface Feature hypothesis would not appear to account for the data.

In fact, the data from the matching task in study 3 suggest that classifications using surface features and deep structure may not provide a sufficiently rich scheme for understanding why people make the classifications that they do. What seems to be needed in addition is the recogntion of the importance of problem type, e.g., active or passive. Judgments of solution similarity by poor problem solvers were facilitated by a match in activeness. The alternatives shared no obvious surface features with the standard, yet a match in activeness produced better performance than a match in only operation for solution. Therefore, a match in activeness, or problem type, must provide some cue of similarity beyond that of operation alone. If one views the process of problem understanding as analogous to the process of constructing a deep structure for a single sentence, then it is possible that useful information might be derived by considering only some, but not all, of the relational information in the problem. In other words, correct conclusions might be drawn without a complete analysis of deep structure.

If it is true that construction of the deep structure is not required in order to gain some understanding of a problem, then the observations that poor problem solvers correctly categorize according to solution similarity with some frequency, and that good problem solvers occasionally err are readily interpreted. For those types of problems that poor problem solvers are not able to complete a deep structure analysis, a more superficial evaluation of the relationships may suffice as a cue that the problems are solved in the same way. However, such superficial evaluation can suggest irrelevant similarity, as in the case of active multiplication and active subtraction.

On the other hand, good problem solvers may decide to halt analysis if sufficient evidence of similarity is obtained. Such premature conclusions sometimes lead to errors.

Educational Implications The results of these studies have certain specific implications, as well as general implications for the teaching of problem solving. This set of studies, along with the NAEP and the Rational Number Project, implies that most students do not have well defined notions of the kinds of situations which require multiplication of fractions. This is true for both active multiplication and passive multiplication, as well as for passive division, which can be considered as requiring a multiplicative process. Multiplication of fractions is not analogous to multiplication with whole numbers, which may explain why students' error rates are so high when they try to interpret these problems as sensible within the context of their knowledge of whole number problems. Teachers should become more aware that all problems which require a single operation, particularly multiplication, can be quite varied in structure. It may be the case, as Bell, Fischbein, and Greer [14] suggest, that the various multiplicative structures should be explicitly taught.

It is quite possible that students would benefit from the activity of categorizing problems according to solution, and explicitly considering both the active and the passive forms of the problems, rather than blindly attempting to solve word problems. Students need to realize that there are several different types of problem structures that all require the same operation for solution. The categorization process, particularly in the solution similarity judgment tasks, may force the student to attend more to the types of situations that occur and how they are similar, rather than merely attending to the details of generating a numerical answer. In any area of mathematics or problem solving, developing a feel for the similarities and differences among the problems that one might encounter seems critical to understanding the concepts involved in solving the problems. Once the concepts are understood, it is much easier to make sensible judgments about the correctness of an answer.

REFERENCES

[1] 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.

[2] 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.

[3] Silver, E.A., 1979, Student perceptions of relatedness among mathematical verbal problems. Journal for Research in Mathematics Education, 10.

[4] Silver, E.A., 1981, Recall of mathematical problem information: Solving related problems. Journal for Research in Mathematics Education, 12, 54-64.

[5] 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.

[6] Watson, W.H., 1980, Results and implications of an arithmetic test. Mathematics in School, 9, 9-11.

[7] Davidson, P.S., Dickenson, A.M., and Tierney, C.C., 1985, Pies are hard to find out about...: An inquiry into children's understanding of the nature of fractions. Technical Report. Harvard University Fractions Group.

[8] Kintsch, W. and Greeno, J.G., 1985, Understanding and solving arithmetic word problems. Psychological Review, 92, 109-129.

[9] Fischbein, E., Deri, M., Nello, M.S., and Marino, M.S., 1985, The role of implicit models in solving verbal problems in multiplication and division. Journal for Research in Mathematics Edcuaticn, 16, 3-17.

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

[11] 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.

[12] Nesher, P., 1982, Levels of description in the analysis of addition 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.

[13] 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.

[14] Bell, A., Fischbein, E., and Greer, B., 1984, Choice of operation in verbal arithmetic problems: The effects of number size, problem structure and context. Educational Studies in Mathematics, 15, 129-147.

[15] Rosch, E. and Mervis, C.B., 1975, Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 8, 573-605.

[16] 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.

[17] Mervis, C.B., and Crisafi, M.A., 1982, Order of acquisition of subordinate-, basic-, and superordinate-level categories. Child Development, 53, 258-266.

Table 1: Adult Responses to State the Operation Task

TYPE of PROBLEM		RESPONSE: Add	Subtract	Multiply	Divide
Fractions (Group 1, N=48|Group2, N=57)					
Addition	Active	<u>47|51</u>	0|2	0|3	1|1
	Passive	<u>42|54</u>	0|1	6|2	0|0
Subtraction	Active	0|1	<u>46|53</u>	2|3	0|0
	Passive	1|1	<u>36|45</u>	1|5	10|6
Multiplication	Active	1|0	30|23	<u>10|18</u>	7|16
	Passive	0|1	1|9	<u>22|22</u>	25|25
Division	Active	1|1	1|2	11|7	<u>35|47</u>
	Passive	2|5	3|6	22|20	<u>21|22</u>
Whole Numbers (Group 2, N=57)					
Addition	Active	<u>54</u>	3	0	0
	Passive	<u>56</u>	0	1	0
Subtraction	Active	0	<u>57</u>	0	0
	Passive	1	<u>55</u>	0	1
Multiplication	Active	0	0	<u>57</u>	0
	Passive	0	0	<u>57</u>	0
Division	Active	0	1	2	<u>54</u>
	Passive	0	0	0	<u>57</u>

Table 2: Eighth Graders Responses to State the Operation Task

TYPE OF PROBLEM		RESPONSE: Add	Subtract	Multiply	Divide
Fractions					
Addition	Active	35	9	2	6
	Passive	47	1	0	4
Subtraction	Active	2	48	0	2
	Passive	3	35	5	9
Multiplication	Active	1	29	3	9
	Passive	4	20	7	21
Division	Active	4	3	5	40
	Passive	4	3	23	22
Whole Numbers (N=52)					
Addition	Active	47	5	0	0
	Passive	47	2	3	0
Subtraction	Active	1	50	0	1
	Passive	4	42	2	4
Multiplication	Active	3	2	45	2
	Passive	3	1	48	0
Division	Active	1	2	4	45
	Passive	2	3	0	47

APPENDIX A: Eight Types of Fraction Word Problems

Addition

Active 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?

Passive Rachel had 4/8 of a pound of bait for the fishing trip. Harry had 3/8 of a pound of bait. How much bait did they have altogether?

Subtraction

Active 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?

Passive Ernest had 3/5 of a box of typing paper. George had 4/5 of a box of typing paper. How much more paper did George have than Ernest?

Multiplication

Active Margret had 3/5 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?

Passive 7/10 of the beds in the garden were planted with flowers. 3/10 of the flowers were tulips. What fraction of the garden was planted with tulips?

Division

Active 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?

Passive Arlen mixed up 2/3 of a bucket of birdseed. He found he had enough to fill 1/3 of his birdfeeders. How much seed would Arlen need to fill all the birdfeeders?