The Effects of Problem Representation and Problem Similarity on Analogical Transfer

by

Lori Meiskey, Gary L. Bradshaw, & Lyle E. Bourne, Jr.

Department of Psychology

University of Colorado
Campus Box 345
Boulder, CO 80309-0345

ICS Technical Report #90-7

The Effects of Problem Representation and Problem Similarity On Analogical Transfer

Lori Meiskey, Gary L. Bradshaw, and Lyle E. Bourne, Jr.

Abstract

We hypothesized that variability within instances of training problems facilitates analogical transfer to new problems. This prediction was based on an analysis of the assumed effects of problem similarity on the mental representation created by training. Specifically, variation in training problems should produce a general schema which matches new problems better than schemas particular to only a certain type of problem. Two levels of problem representation were manipulated: the text base in the form of relational terms in the problems, and the situation model in the form of the story context of the problems (see van Dijk & Kintsch, 1983). Variations in either level should produce a more generalized schema, which in turn is expected to enhance transfer. In the first experiment, no such effect was obtained for either level of problem representation. Additional training in the second experiment did produce a positive transfer effect for variability at the story context level, suggesting that subjects may require more than two structurally isomorphic problems in order to form a generalized schema adequate to aid transfer. The variable and non-variable groups differed in the amount of positive or negative transfer from the previous problem but not in the time to complete the transfer problem.

Introduction

Solving a problem from scratch is a complex task: it involves deciding which aspects of the problem are relevant, determining which operators are useful to obtain the solution, and recognizing when the solution has been reached. Each step in turn may involve extensive search, perhaps even blind trial and error. If a successful solution to an isomorphic previously-solved problem is known, these initial difficult steps may be eliminated. Indeed, solving a problem using a known method is often more efficient than resorting to general problem-solving procedures. When a student enters a new domain of problem solving, however, general heuristics may be all that he or she has available. During training, a student builds domain-relevant skills and knowledge by solving problems and/or by learning general principles of the domain (e.g., the meaning of Conservation of Energy in physics). At some later time, the student may be asked to use, or <u>transfer</u>, the relevant skill and knowledge to a novel situation. Instructors would like to maximize the overlap between what they teach and the new problems or situations students will encounter, but, as Bassok and Holyoak (1989) point out, training cannot always take account of the conditions under which a solution may be useful. Therefore, it is hoped that the skill and knowledge gained in training will generalize to new situations which were not directly taught.

Although transfer of problem solving skill is desirable, research on problem solving has indicated that people have a difficult time accessing relevant solutions when presented with a new problem (e.g., Reed, Ernst, & Banerji, 1974; Reed, Dempster, & Ettinger, 1985; Gick & Holyoak, 1980, 1983). Subjects in these experiments often failed to retrieve and use information from structurally isomorphic problems solved earlier in the experiment. Since transfer often failed to occur in these experiments between episodes separated by a few minutes, it would seem that extra-laboratory transfer must also be a rare and unusual occurrence. Yet when students learn a new domain in a more naturalistic setting, they often spontaneously retrieve old examples to help them understand the new problem (Ross, 1984; Anderson, Farrell, & Sauers, 1984). In our everyday learning experiences spontaneous analogical transfer might be promoted by factors absent from previous laboratory settings. In order to understand why subjects in these studies exhibited lack of transfer, the conditions governing access and mapping of a previous solution must be examined.

As we shall see, studies in which little transfer was found have failed to consider one or both of the following determinants of transfer. First, subjects must understand the original problem if it is to be useful as an analogue during transfer. That is, at a minimum, subjects must have encoded and clearly represented the information that is relevant to solving the problem. Second, it is important to consider the similarity of the representations of training and transfer problems.

Understanding the problem

One reason why transfer of the solution of an old problem to a new, structurally isomorphic problem may not occur is that some aspects of the old problem (e.g., the solution or particular relations in the problem) may not be well understood or remembered. For example, a study by Reed et al. (1985) found little transfer between structurally isomorphic word algebra problems. In these experiments, subjects were shown the solution to each type of word problem, yet no criteria were used to establish that subjects understood or attended to the given solutions. In order to build the equation (the solution) in these problems, subjects must understand the relations between elements in the problem. Subjects must also know which values to assign to each variable in the equation. Mayer (1982) found that some aspects of a problem are more difficult to remember, which might reflect a lack of complete understanding. Relational propositions (which indicate the relation between two variables, such as x=2y) in a problem were less well remembered than assignment propositions (which assign a value to a variable, such as x=5). Relational propositions are essential to the solution of the word problems used by Reed et al. (1985). Thus, the lack of transfer in these experiments may derive from a lack of understanding of source problems, not from failures in the transfer process. In fact, Reed et al. (1985) found better transfer when the initial problems were better explained. Further, Perfetto, Bransford, and Franks (1983) found that subjects' incorrect but self-generated solutions interfered with access of previously given correct solutions, even when a hint was given to use the previous correct information. Subjects in the studies showing lack of transfer may only have remembered their solutions, rather than the given solution.

Problem categories

Gick and Holyoak (1983) provide an alternative explanation for the lack of transfer found in some previous studies. In most studies where subjects failed to show transfer, subjects were trained on a single problem. Gick and Holyoak replicated the usual finding of lack of transfer for their word problems when subjects were trained with one example. However, when subjects solved two problems which had the same general solution, they were able to transfer this solution to the same novel problem that the one-example group failed to solve. The authors hypothesized that, by giving subjects more than one problem, an abstract representation was composed out of the similar elements from both problems.

Other investigators have similarly found evidence for abstract schemata in the form of problem types or categories (e.g., Hinsley, Hayes, & Simon, 1977). Still other researchers have found that experts but not novices can classify problems based upon a more abstract representation of the structure of the problem (e.g., in studies of physics problems, Chi, Feltovich, & Glaser, 1981). Finally, Larkin, McDermott, Simon, and Simon (1980) found that experts solved physics problems in fundamentally different ways than did novices. Novices seemed to rely on a syntactic, line-by-line translation of the problem; experts, on the other hand, transformed the problem into a representation in which inferences could be made more easily (e.g., a diagram). Hinsley et al. (1977) also reported that when subjects were not able to categorize a problem, they proceeded in solving the problem by syntactic translation. When a problem was categorized as being of a familiar type, however, subjects proceeded by reading the entire problem and then providing the relevant formula in one step (rather than generating the formula in a series of steps).

This evidence suggests that during training, people develop categorizations of problem types, then retrieve problem representations and solutions to aid in representing and solving new problems from the same category. But why would an abstract representation, or problem schema, provide for better transfer to a new problem than would an analogous solution from a single problem? Presumably factors affecting retrieval of previous information include the strength of that information (which depends on such factors as recency and amount of practice) and the similarity between the representation of the current problem and that of the old problem in memory. For the subject to retrieve a previous solution, the current problem must be seen as similar to the previously-solved problem. Similarity can exist at the level of surface features or at the level of abstract structural relations. Also, a more recent and well-practiced solution should be more easily retrieved than an earlier or less well-practiced solution.

In Dellarosa's (1985) model of schema acquisition, after an initial text comprehension phase, the representations of two successfully-solved problems are compared. Elements of the two problems that match (e.g., structural relations) are extracted and variabilized. By this process, elements which are irrelevant to the problems' solutions (e.g., surface features) are eliminated or weakened. As further problems are solved and compared, the schema comes to contain fewer irrelevant elements. This analysis of schema formation as intersection of problem representations forms the basis of Gick and Holyoak's two arguments for the superiority of schemata over single stored examples in facilitating transfer.

One of Gick and Holyoak's arguments is based upon Tversky's (1977) analysis of similarity. According to this analysis, all problems or situations are decomposed into features. Similarity between two problems depends on the amount of overlap of the features. An abstract schema would contain the common elements of several problems, and so will have many similarities and few idiosyncratic differences with the new problem. Thus there would be less discrepancy between a new problem and the relevant schema than the new problem and a specific old problem with all features intact (that is, unless the old problem happens to have all of the same features, solution-irrelevant as well as relevant, as the new problem). Gick and Holyoak's second argument for the superiority of a schema in facilitating transfer is that information relevant to the solution is highlighted by a schema, which aids in mapping the relevant features of the new problem onto the schema. In comparing two particular problems, on the other hand, many mismatches will occur, and effort is required to decide whether or not these mismatches are relevant or irrelevant to the problems' solution.

Whether a specific example or a schematic representation is accessed for any given problem may depend upon the extent of a subject's training. At first, subjects will rely on analogies to particular examples to help them; as training progresses, however, subjects depend less on particular examples (Ross, 1984; Homa, Sterling, & Trepel, 1981; Elio & Anderson, 1981).

Levels of problem representation

Nearly all current researchers agree that the representation of problems is an important factor in explaining analogical transfer, but they have differing notions about what this representation is and about the mechanics of its effects on transfer. In Gentner's (1983) theory, the structure of a problem or situation is represented as nodes (consisting of objects) connected by relations. In mapping from one situation to another, similarity between two problems at the level of higher-order relations takes precedence over similarities at the level of lower order relations (e.g., relations among objects take priority over attributes connected to objects).

Holyoak and Koh (1987) argue that Gentner's syntactic rule that gives higher order predicates priority is insufficient to explain transfer. They claim that the relevant aspects of the problem, based on its goals, affect the processes involved in analogical transfer. Goal-relevant features of a problem form the "structure" of the problem; non-goal relevant features are termed the "surface" features. Holyoak and Koh predict and confirm that similarities among problems at the surface and structural levels will affect retrieval of an analog, but only similarities at the structural level will affect mapping of the retrieved problem onto the new problem, and hence assist in problem solution. Note that these researchers retain Gentner's notion of representation--features or objects are connected by relations. However, they add an element to the representation which they believe is vital in determining transfer, that is, goals.

Ross (1987) adds another level of structure to word problems. At one level are problem elements and their interconnections, a level of structure assumed also by both Gentner and Holyoak and Koh. But at another level is the meaning of the problem elements and interconnections. Ross notes that problem solvers may use this meaning level, as well as the level of superficial similarities, in order to make an analogy.

Although Holyoak and Koh and Ross all examined analogical transfer in the domain of word problems, they did not consider models of how the texts in which the problems are embedded might be represented by subjects (e.g., Larkin, 1988; Larkin et al., 1980; van Dijk & Kintsch, 1983; Schank &

Abelson, 1977), and how that form of representation might interact with or substitute for problem elements and their meaning. In fact, the systems of Holyoak and of Gentner seem to work independently of how a particular subject represents the problem. Both researchers assume that all subjects represent all information in the problem in an optimal manner. Larkin (1988) found that novices and experts in physics represented given problems quite differently. The novices based their representations purely on the physical elements of the problem (e.g., inclined plane, blocks on a pulley), but experts represented the problem in terms of theoretical, non-physical entities (e.g., forces). Thus, we might expect people with varying degrees of knowledge in other domains form different representations of problems in those domains. We suggest that an examination of extant text comprehension models might clarify the variables of problem representation and thereby provide some added insight into analogical problem solving and transfer.

One model of text comprehension and representation which is supported by and has produced a great deal of empirical evidence is that of van Dijk and Kintsch (1983). In this model, the representation of a text can be viewed as consisting of two levels of structure: the text base and the situation model. The text base can further be decomposed into two sub-levels of representation: the microstructure and the macrostructure. The microstructure represents propositions derived directly from the text. The macrostructure is derived from and is superordinate to the microstructure and consists of the gist or recurring ideas in the text. The situation model is the interpretation of the text based on the reader's world knowledge; that is, it represents what the text means. Ross' study seems to conform to this level of analysis, but his distinction between these levels is not entirely clear, and he does not relate this distinction to other work. However, Ross' results indicate that these two levels of representation (text base, or his surface level; and situation model, or his meaning level) can have an impact on analogical transfer. In the current study, we bring together notions from two research areas, text comprehension and schema acquisition and use, in an attempt to further elucidate analogical problem transfer. We examine the effects of these two levels of representation on the formation of problem schemas as well as on their use in solving new problems.

Framework for the study

The literature supports the following statements. Transfer depends on the strength and completeness of the representation of the original training problem(s) in memory, and on the similarity between the transfer problem and the training problem(s). Similarity depends upon the overlap of elements of the analogous problem or schema and the novel problem. Similarity can exist at multiple levels of problem representation. More general elements within an analogue schema are more likely to match specific elements in the novel problem. We examine certain implications of these statements in the domain of algebra word problems.

The main reason for choosing this domain was that algebra word problems contain a rich structure with both text base and situation model levels of semantic representation. Problem schemas are organized around an underlying equation which highlights important information in the problem. But this information can come from both the text base and situation model. The situation model, including general knowledge about mathematical relations, constrains the information in the text base which is relevant to solution and adds other information not directly contained in the problem. Algebra word problems were also chosen because they have been used in previous research on analogical problem transfer.

In the current experiments, the generality of a learned schema is examined by manipulating the variability of training problems at the levels of the text base and situation model independently. We test the hypothesis that varying problems at the text base and/or situation model level will result in a more general problem schema which will benefit transfer to a new problem. As described earlier, the slots of a schema may be derived from elements in particular problems. If the problems are similar, then the slots will be filled by only a narrow range of elements. On the other hand, if the problems are dissimilar, the slots of the abstracted schema will cover a wider range of elements. A more general schema will provide a better match to the elements of a novel problem than a schema whose slots fit only specific elements (which are not present in the new problem). We test the hypothesis that similarity between problems at either the text base or situation model level (or both) will benefit transfer. The final hypothesis, that subjects in earlier experiments may not have understood the original problem, contributing to lack of transfer, was not tested directly in these experiments. However, we attempted to control for this factor by having subjects progress through a series of hints which pointed out the essential information (especially the relations) in the training problems. That is, the hints were used to aid subjects' understanding of how an equation was constructed from information in the problem. Hints were given when a subject could not complete a problem within a set time limit.

Subjects in these experiments solved a set of training problems followed by a transfer problem. The training problems were designed so that subjects would induce either a general schema from problems with varied text bases and/or situation models, or a specific schema from a uniform set of problems. The text base manipulation consisted of varying the mathematical relational statements in the problem (e.g., "increased...over" in the sentence "Fred increased his speed 20 miles per hour over his normal speed"). Relational statements could either be identical or different from problem to problem. At the situation model level, the manipulation consisted of varying the story context of the problems (e.g., travel problems, mixture problems). Again, the story context could either be the same across problems (i.e., no variability) or varied. Similarity among problems at the level of story context or relational terms should enhance transfer to similar problems, relative to problems which are dissimilar at one or both of these levels. However, when the transfer problem requires that the subject solve a problem with both new relational terms and story context, then having solved training problems with variability of either the story context or the relational terms should enhance transfer to this new problem. Variability at both the story context and relational terms levels should maximize transfer under these circumstances.

The main dependent measure in these studies was the time required to reach the correct equation, a successful solution to the problem. Admittedly, this measure does not give a detailed account of the many processes which occur in attempting to solve a problem. The subject must read and understand the problem, identify the relevant aspects of the problem, determine how all of the relevant pieces fit together, and instantiate the values for the variables for each part of the equation. Successfully making an analogical transfer from old problems should, however, speed up all of these processes (strong analogy will facilitate retrieval of the previous solution, representation of the new problem, and mapping of the new problem onto the old problem), so that time to solution gives a consistent overall performance measure. Other dependent measures were also examined. One was whether or not a subject solved a problem without hints, the typical measure used in studies of analogical transfer. The final measure was whether or not a subject said that a problem was similar to a previous problem or problems.

In summary, the hypothesis to be examined in these experiments is that similarity, represented at either a text base or situation model level, facilitates transfer. Similarity may come from either of two sources: a) the similarity is intrinsic to two problems, as when the two problems share the same elements (e.g., Fred, a car) or story contexts; and b) through variability of instances and the induction of a general schema.

Experiment 1

Method

Subjects solved two training problems followed by a transfer problem. The training problems contained similar or different story contexts and same or different relational terms. Subjects exposed to different story contexts and/or relational terms were expected to form schemas which generalized across the varying stories or terms. Whether the generality of the induced schema would aid in transfer to a novel problem was tested by presenting a transfer problem with a new story context and/or relational terms.

Subjects

Sixty students enrolled in an introductory psychology course at the University of Colorado at Boulder in the Spring semester of 1988 participated in the study in partial fulfillment of a course requirement.

Materials

Algebra word problems were constructed which represented the general equation $a1(x+b)=a2(x-b)$, a variation on the formula distance equals rate multiplied by time. Here, $a1(x+b)$ and $a2(x-b)$ represent the distance, a1 and a2 represent times, and x and b represent different rates. The quantities for a1, a2, and b were specified in the problem; the variable x was the unknown in the problem. This equation was instantiated in three different story contexts: Travel, Work, and Mixture. See Table 1 for example problems from each of these categories.

Problems in the travel category involved someone traveling for a given time (a1 and a2) at an unknown rate (x) that was either increased or decreased by a constant (b). The distance traveled could be found by

Table 1

Example problems used in Experiments 1 and 2

Story context: Travel
Relational terms: Increase over/Decrease from
Equation: 1(x+20)=2(x-20)

Fred and his family were driving to Pueblo for vacation. When Fred goes downhill, he increases his speed 20 miles per hour over his normal speed. But when he goes uphill, he decreases his speed 20 miles per hour from his normal speed. Fred liked to try to figure out exactly how long his trips would take. Fred thought the trip was all uphill, and he figured they would get to Pueblo in 2 hours. But actually the trip was all downhill, and they arrived 1 hour after they left.
(How fast does Fred's car travel at its normal speed?)

Story context: Work
Relational terms: Raise above/Reduce below
Equation: 200(x+3)=600(x-3)

Henrietta stocks shelves at Kmart. Henrietta used to take speed. She would stock 600 items each day, and the management would reduce her pay rate to 3 cents per item stocked below her usual pay rate. After she quit taking drugs, she only stocked 200 items each day. The management felt sorry for her, so they raised her pay rate to 3 cents per item stocked above her usual pay rate. Henrietta earned the same amount of money each day after she quit taking drugs as she had earned while she was taking drugs.
(What was Henrietta's usual pay per item stocked?)

Story context: Mixture
Relational terms: Higher than/Lower than
Equation: 5(x+3)=8(x-3)

Suzy was a bartender who was known for miles around for her special drink, the "Quando Surprise." The secret ingredient in the drink was chili powder. Today, she mad 5 gallons of the drink with a 3 milligrams of chili powder per gallon of mix higher concentration than her usual mixture. The customers drank this batch so fast that Suzy made an 8 gallon batch, but this time she used a 3 milligrams of chili powder per gallon of mix lower concentration than her usual mixture. There was as much chili powder in the first batch as in the second batch.
(How many milligrams of chili powder are in each gallon of Suzy's usual mixture?)

---

multiplying the time of travel by the rate of travel (i.e., a1(x+b) or a2(x-b)), but this was not explicitly stated in the problem. The distance traveled at the lower speed was the same as that traveled at the higher speed, although this was also implicit in the problem.

In the Work category, a1 and a2 corresponded to the number of items produced, and x+b and x-b corresponded to an unknown pay rate (x) which was either increased or decreased by another, constant pay rate (b). Each problem explicitly mentioned that the amount of money earned (i.e., a1(x+b) and a2(x-b)) was the same at the lower pay rate as at the higher pay rate.

For problems with the Mixture story context, a1 and a2 corresponded to the liquid amount of a mixture which contained a dissolved solid with a higher or lower concentration than a reference mixture's unknown concentration (x+b and x-b, respectively). Each problem explicitly mentioned that the amount of the solid was the same in both the mixture with the higher concentration (a1(x+b)) and the mixture with the lower concentration (a2(x-b)).

Three pairs of relational statements were used in the problems to represent the mathematical operations of addition and subtraction (i.e., x+b and x-b, respectively): increase to/decrease from, raise above/reduce below and higher than/lower than.

Each problem was typed on a separate 8 1/2" by 11" sheet of paper. The question requesting the unknown quantity was typed directly below the problem in parentheses.

A series of hints was constructed for each problem. See Table 2 for an example of the series of hints. Each hint provided an essential relation in the problem. The first and second hints pointed out the addition and subtraction relations, respectively. The third hint told the subject that the amount (i.e., distance, pay, or amount of solid for Travel, Work, and Mixture problems, respectively) could be found by multiplying a1 by the rate or concentration (i.e., a1(x+b) and a2(x-b)). The fourth hint identified the unknown quantity. The fifth hint instructed subjects that the amounts a1(x+b) and a2(x-b) were the same, regardless of the higher or lower rate/concentration. The sixth hint was the equation with the correct quantities from the problem entered (e.g., 3(x+2)=4(x-2)). The final hint showed the equation with intervening algebraic steps necessary to reach the answer to the problem.

Design

The design is represented in Table 3. The groups and the relations among the problems are shown in Table 3, part A. For the group notation (which will be used in the remainder of the paper), the first two letters before the slash mark indicate the training conditions. The first letter indicates whether the story context was the same (S) or different (D) between the training problems. The second letter indicates whether the relational terms were the same or different between training problems. The two letters after the slash indicate the relation between the transfer problem and the training problems. The first letter after the slash indicates whether the story context of the transfer problem was the same as or different from the training problems; the second letter after the slash indicates whether the relational terms were the same as or different from the training problems. The letters (A,B,C and a,b,c) in Table 3, part A, of the table indicate a particular type of story context (e.g., Travel)

Table 2

Example of the series of hints used in Experiments 1 and 2

Hint 1 Addition relation

Fred got to Pueblo in 1 hour, because the trip was all downhill, and he increased his speed 20 miles per hour over his normal speed.

Hint 2 Subtraction relation

Fred planned to get to Pueblo in 2 hours, because he thought that the trip was all uphill, and he decreased his speed 20 miles per hour from his normal speed.

Hint 3 Multiplication relation

The distance to Pueblo is found by multiplying the car's speed by the time it takes to get to Pueblo.

Hint 4 Unknown quantity

You must find the car's normal speed.

Hint 5 Equivalence relation

The distance to Pueblo is the same whether Fred is going uphill for 2 hours, or he is going downhill for 1 hour.

Hint 6 Problem equation

$$1\ (x + 20) = 2\ (x - 20)$$

Hint 7 Answer for the unknown

$$1\ (x + 20) = 2\ (x - 20)$$

$$1x + 20 = 2x - 40$$

$$60 = x$$

Table 3

Design for Experiment 1: a) story context (context) and relational terms (relations) for each problem in each group; same letters indicate the same type of problem, numbers indicate instances of that type, b) the Variability groups and Transfer Cue groups.

a)

	Training Problems 1		Training Problems 2		Transfer Problem	
Group	Context	Relations	Context	Relations	Context	Relations
SS/DD	A1	a1	A2	a1	C3	c3
SD/DD	A1	a1	A2	b1	C3	c3
DS/DD	A1	a1	B1	a1	C3	c3
DD/DD	A1	a1	B1	b1	C3	c3
SS/DS	A1	c3	A2	c3	C3	c3
SS/SD	C1	a1	C2	a1	C3	c3

b) Variability Groups

		Context: Same	Context: Different
Relation	Same	SS/DD	DS/DD
	Different	SD/DD	DD/DD

Transfer Cue Groups

		Context: Same	Context: Different
Relation	Same	----	SS/DS
	Different	SS/SD	SS/DD

and type of relational terms (e.g., higher than/lower than). The numbers following these letters indicate that the problems are particular instances of that type of story context or relational terms.

Each subject received two "training" problems followed by a "transfer" problem. The second training problem had either the same or a different story context, and either the same or different relational terms as the first problem. Four groups of subjects were formed by crossing story context (same or different) with relational terms (same or different) for the training problems. All of the subjects in these four groups received a third (transfer) problem with both a different story context and different relational terms than those used in the training problems.

Two other groups of subjects had the same story context and the same relational terms in both training problems. For the fifth group, the relational terms in the transfer problem matched the terms in the training problems, but the story context was changed (SS/DS). The final group had a transfer problem with the same story context but different relational terms from the training problems (SS/SD).

The first four groups test the effects of variability in training (i.e., having a different story context and/or relational terms in training) on transfer to a problem where both the story context and the relational terms are new to the subject (i.e., are different from those in training). Thus, they will be termed the Variability groups.

The first, fifth and sixth groups (SS/DD, SS/DS, and SS/SD, respectively) test whether preserving the story context or relational terms between the training and transfer problems will facilitate transfer on this last problem. These groups will be termed the Transfer Cue groups. Note that group SS/DD is included in both groupings.

Procedure

Subjects were tested individually. The subject and the experimenter were seated on opposite sides of a table, approximately 3 ft apart. Subjects were randomly assigned to conditions. Subjects were told that they would be given three problems and would have 5 min to complete each problem. If they failed, a series of hints would be presented. Then the subject was handed the first problem, and the experimenter started a stop watch. The times at which the subject wrote down the correct equation and then the correct answer were recorded by the experimenter. If the subject arrived at an incorrect answer, the experimenter told him/her that the answer was incorrect. If the subject did not finish the problem within 5 min, the experimenter stopped the stop watch then explained that the subject could read the hints at his/her own pace until a hint was found which helped to solve the problem. A cover sheet was placed over the hints, and the subject was allowed to move the sheet down to expose each subsequent hint. All subjects read the full series of hints for the first problem; if a subject reached the correct answer before seeing the hints, the hints were given after the subject wrote the correct answer. After completing the first problem, the subject was given the second problem; upon completing this problem, the subject was given the final problem.

After each problem was completed, the subject was asked how he/she had solved that problem. Subjects who completed the first problem only after seeing the equation (i.e., hint 6) were asked if they understood how the equation related to the problem and were asked to explain that relationship. The experimenter wrote down their explanations but did not correct any misunderstandings or give feedback about the subject's explanation.

Results

Six subjects were excluded from the analyses: four for writing the correct numerical answer without writing an equation first, and two for recording errors.

The original time-to-equation measures (in minutes) were transformed by taking the natural log of each and analyses were performed on the transformed times, unless otherwise noted. There were nine subjects in each group.

Variability groups

First, we will present the analyses for the four groups for which story context (same/different) was crossed with relational terms (same/different) in training, and all subjects received a transfer problem with a different story context and different relational terms than those in training.

Time to equation. Our first measure of transfer is the difference in the time to reach the equation between consecutive problems. If a subject solved the second problem faster than the first problem, for example, this would indicate positive transfer. Also of interest were differences in the times to reach the equation for the transfer problem for each group. The mean times to equation (based on the log transformation) for each group and for each problem are plotted in Figure 1.

A contrast regression analysis was performed with the difference between consecutive problems as the dependent variable, and context,

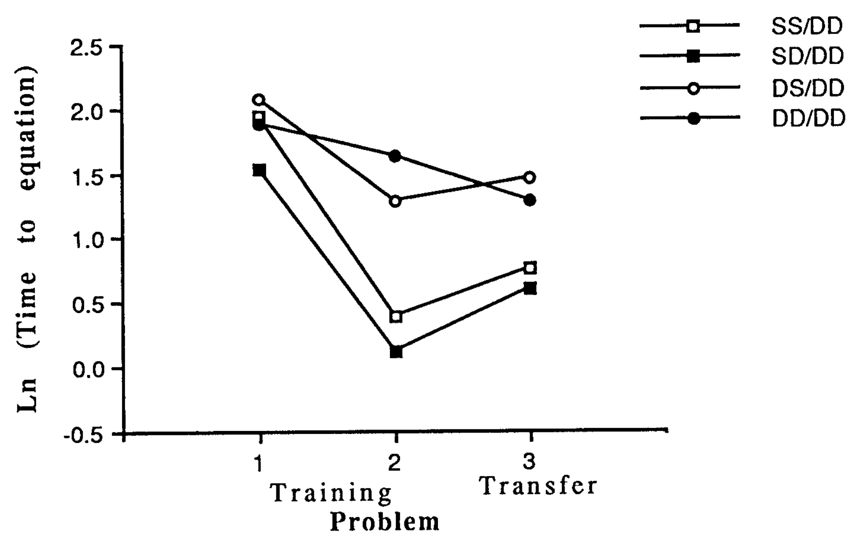

Figure 1

Mean log time to equation for Variability groups, Experiment 1 relation, and their interaction as the independent variables. The following contrast codes were used for the differences between consecutive problems (note that these contrasts are not orthogonal, but do represent the a priori questions of interest): 1 -1 0 and 0 1 -1, for the first, second, and third problems, respectively. These contrasts were regressed on the factors story context (-1=same, 1=different), relational terms (-1=same, 1=different), and their interaction (found by multiplying the story context and relational terms values). Also, the overall effect (collapsing across both contrasts) of these factors and their interaction was determined (Judd & McClelland, 1989; Rosenthal & Rosnow, 1985).

The mean log difference in time between problems for the same context groups across problems was 0.53 (1.70 minutes); the mean for the different context groups was 0.31 (1.36 minutes). The overall effect for story context was significant, $\underline{F}(1,32)=40.77$, $p<.001$, MSe=.34. The same story context groups showed a larger amount of transfer by the measure of the difference between consecutive problems than did the different context groups. The overall effects for relational terms and the interaction of context and relation were not significant, $\underline{F}(1,32)=1.62$, $p=.21$; $\underline{F}(1,32)=1.46$, $p=.24$, respectively.

The mean difference scores between the first and second problem were 1.47 (4.35 minutes) and 0.53 (1.70 minutes), for the same and different context groups, respectively, which was significant, $\underline{F}(1,32)=13.50$, $p=.001$, MSe=.29. This effect indicates that having a story context in the second problem which differed from the story context of the first problem was detrimental to transfer relative to having the same story context in both problems. The relation and context-by-relation effects were not significant, $F(1,32)=1.77$, $p=.19$; $F<1$, respectively. However, the means for the relation effect were in the predicted direction: 1.17 (3.22 minutes) for the same relational term groups and 0.83 (2.29 minutes) for subjects in the different relational term groups, indicating a facilitory effect for having the same relational terms in the second as in the first problem, relative to having different relational terms in both problems. There was also an overall improvement, across groups, from the first to the second problem, with a mean difference of 1.0 or 2.72 minutes, $F(1,32)=61.13$, $p<.001$, $MSe=.29$.

The means for the difference between the second and the transfer problem were in the predicted direction (same context mean= -0.42 or -1.52 minutes, different context mean=0.08 or 1.08 minutes). The values of these means indicated a positive effect on the difference from the second to the transfer problem with varied contexts in training (although a small difference here of 1.08 minutes), relative to having problems with a non-varying context in training, which produced negative transfer. This effect, however, was marginal, $F(1,32)=2.64$, $p=.11$, $MSe=.43$. Again, the effects of relational terms and the context by relation interaction were not significant, both F's<1. There was no significant overall improvement from the second to the transfer problem, $F(1,32)=1.24$, $p=.27$.

The means for the time to solve the last problem were 0.68 for the same story context groups (1.97 minutes) and 1.37 for the different story context groups (3.94 minutes). The groups with the same story context completed the transfer problem faster than the different story context groups. The analysis of the effects of context and relation on the speed with which subjects solved the final problem was performed by regressing the transformed time for the transfer problem on context, relation, and their interaction. Story context was a significant predictor of this time, $F(1,32)=12.22$, $p=.001$, $MSe=.35$. The effects of relation and the interaction between relation and context were not significant in predicting time to equation on the final problem (both $F<1$).

Hints. Table 4 presents the proportion of subjects completing each problem without hints in each group. Few subjects completed the first problem without hints (11% in all groups except group SD/DD). On the second problem, group DD/DD was the only group in which the majority of subjects required hints. For the transfer problem, few subjects required hints to solve the problem, except in group DD/DD where nearly half of the subjects required hints in order to solve the problem. A regression analysis similar to that for the time analysis and using the same contrasts was used to analyze the hint data. The data were coded with a 1 indicating the subject completed the problem without hints or a 0 indicating the subject required hints. The results followed the same pattern of significance as the time data, except for three differences. First, the effect of context on the first to second problem difference was not significant, $F(1,32)=1.42$, $p=.24$. Second, the effect of context on the second to the transfer problem did not approach significance ($F<1$), as it had for the time data. Finally, A significant context by relation interaction was found across problems, $F(1,32)=4.5$, $p=.04$, $MSe=0.95$; context was important only when relational terms differed during training.

Table 4

Proportion of subjects solving each problem without hints for Experiment 1, Variability groups

	Problem Number		
Group	Problem 1	Problem 2	Problem 3
SS/DD	0.11	0.78	0.89
SD/DD	0.44	1.00	0.89
DS/DD	0.11	0.67	0.78
DD/DD	0.11	0.33	0.56

Same statements. In Table 5, the proportion of subjects stating that a problem was the same as a previous problem (or the set of previous problems) is listed for each problem by group. For the groups who received a second problem with the same context as the first problem (groups SS/DD and SD/DD), all subjects indicated that the second problem was the same as the first problem. A substantially smaller proportion of subjects in the groups where the context of the second problem varied from the first said that the second problem was the same as the first (22% for group DS/DD and 33% for group DD/DD). On the transfer problem, over half the subjects in group SS/DD said that the problem was the same as the previous problems; however, note that this is a substantial drop from the 100% of subjects who reported that the second problem was the same as the first problem. The proportion of subjects who said that the transfer problem was the same as the others in group SD/DD also dropped substantially, and to a level below that of group SS/DD. The proportion of subjects who said that the transfer problem was the same as the other problems remained low for subjects in the varied story context conditions. The same statement data were also analyzed in the same way as the time data. The pattern of results was the same as for the time data, except for the difference from the second to the transfer problem. Subjects said that the given problem was similar to a previous problem or problems more often for the second problem than for the transfer problem, $\underline{F}(1,32)=10.81$, $p=.002$, $MSe=.13$. This effect interacted with context, $\underline{F}(1,32)=10.81$, $p=.002$, $MSe=.13$, with subjects in the different story context groups not differing in saying that a given problem was the same as a previous problem for the second and transfer problems, and the same story context training groups making fewer same statements for the transfer than for the second problem. Recall that no effect of context was found for the solution time differences measure from the second to the transfer problem.

Table 5

Proportion of subjects stating that the problem was the same as previous or a previous problem for Experiment 1, Variability groups

	Problem Number	
Group	Problem 2	Problem 3
SS/DD	1.00	0.56
SD/DD	1.00	0.33
DS/DD	0.22	0.22
DD/DD	0.33	0.33

Transfer Cue groups

Time to equation. The same contrasts utilized to code differences between problems used in the above analysis were used to calculate the dependent variable in a regression in which these contrast codes were regressed on contrasts for the story context (same/different) and relational terms (same/different) factors, which defined the various transfer conditions. The three groups analyzed here had identical training conditions: same context and same relational terms for the first and second problems. Since a group with the transfer condition of same context and same relational terms as the training problems was not included in the experiment, the interaction between context and relation cannot be determined. See Figure 2 for the plotted mean transformed times for each group and each problem.

The means for the context effect across problems were 0.65 for the same context groups and 0.94 for the different story groups, indicating a positive effect for having the same context between problems. This effect was marginally significant, $F(1,24)=3.16$, $p=.09$, $MSe=.33$. The overall effect of relational terms did not approach significance, $F<1$.

The mean improvement in time from the first to second problem across groups was 1.54 or 4.66 minutes, which was significant, $F=129.06$, $p<.001$, $MSe=.25$. The difference between the first and second problems was not related to either the context or relational terms factors, both F's<1. This is not surprising, since the groups did not vary in condition on the first and second problem.

When the transfer problem had the same story context as the training problems, the mean improvement in time from the second to the transfer

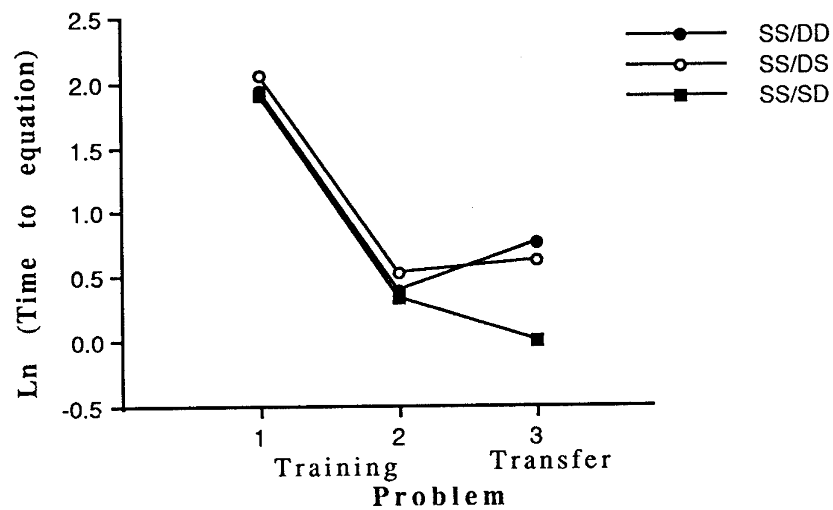

Figure 2

Mean log time to equation for Transfer Cue groups, Experiment 1 problem was 0.33 (1.39 minutes); when the story context of the transfer problem was different from the training problems, there was a mean increment in time from the second to the transfer problem of 0.24 (1.27 minutes). The difference between the second and transfer problems was marginally related to context, $F(1,24)=2.84$, $p=.11$, $MSe=.39$. The effect of relational terms did not approach significance, $F<1$.

When the transfer problem had the same context as the training problems, the mean log time to reach the equation for the transfer problem was 0.01 (1.01 minutes); when the transfer problem had a different context than the training problems, the mean log solution time was 0.69 (1.99 minutes). Story context was a significant predictor of the time to reach the equation for the transfer problem, $F(1,24)=7.27$, $p=.01$, $MSe=.35$.

Hints. Few subjects reached the equation for the first problem without hints (11% for groups SS/DD and SS/DS, and 33% for group SS/SD). For the second problem, all subjects in groups SS/DS and SS/SD completed the problem without hints; 78% of the subjects in group SS/DD finished without hints. For the transfer problem, all subjects in groups SS/DS and SS/SD and most of the subjects in group SS/DD solved the problem without any hints (see Table 6). The hint data was analyzed in the same manner as the time data and followed the same pattern of results.

Table 6

Proportion of subjects solving each problem without hints for Experiment 1, Transfer Cue groups

	Problem Number		
Group	Problem 1	Problem 2	Problem 3
SS/DD	0.11	0.78	0.89
SS/DS	0.11	1.00	1.00
SS/SD	0.33	1.00	1.00

Same statements. On the second problem, all subjects in groups SS/DD and SS/DS and 78% of the subjects in group SS/SD stated that the problem was the same as the first problem. For the transfer problem, the proportion of subjects who said that this problem was the same as the other problems was reduced relative to that of the second problem to 56% for group SS/DD and was reduced somewhat less for group SS/DS to 78%. Most subjects in group SS/SD (89%) said that the transfer problem was the same as previous problems (see Table 7). The same statement data were analyzed in the same manner as the time data and followed the same pattern of results, except where a marginal effect of context on the difference from the second to the transfer problem was found in the time data, a significant effect was found in the same statement data, $F(1,24)=5.00$, $p=.04$, $MSe=.14$. Groups whose second and transfer problem differed in story context showed a decline in same statements from the second to the transfer problem; the group with the same story context across problems showed an increase in same statements (note, however, that there is no *a* *priori* reason why group SS/SD should have differed from the other two groups in frequency of same statements on the second problem).

Table 7

Proportion of subjects stating that the problem was the same as previous or a previous problem for Experiment 1, Transfer Cue groups

	Problem Number	
Group	Problem 2	Problem 3
SS/DD	1.00	0.56
SS/DS	1.00	0.78
SS/SD	0.78	0.89

Discussion

The prediction that similarity among problem elements facilitates problem transfer was strongly supported at the story context level and weakly supported at the relational term level. Conditions of this experiment with similar story context in the first two problems showed a high degree of transfer; those with different story contexts showed a smaller degree of positive transfer. The effect of relational statements was marginal, but in the predicted direction; conditions with the same relational statements in the training problems showed a higher degree of transfer from the first to second problem than conditions with different relational terms. Subjects appeared to base their statements of similarity for the first two problems solely and strongly upon whether or not the story contexts were the same between the two problems.

In contrast to the significant early transfer obtained in some conditions of the present experiment, Gick and Holyoak (1983) suggest that subjects require at least two problems before much transfer occurs. One possible basis for the difference in results lies in the fact that the cover stories for the problems of Gick and Holyoak were quite different from each other, a condition similar to the different story context conditions in the present experiment. Under these circumstances, we also observed low transfer from the first to the second problem.

The positive effect of having similar story contexts was also supported indirectly in other conditions of the present experiment. For example, solving a problem with a different story context had a less deleterious effect on transfer after two similar problems were solved than when subjects solved only one previous problem. As in Gick and Holyoak's experiments, having two problems might allow subjects to make some abstractions relevant to any problems within the domain.

The effect of variability in story context during training leading to better transfer was marginal. In order for the hypothesized benefits of variability to occur in transfer, the subject must have used the training instances to develop a schema which generalized across those instances. Some evidence suggests that this did not occur for subjects in the variable story context groups. First, few subjects in either group DS/DD or DD/DD stated that the second problem was similar to the first problem. If subjects do not see that these problems are similar, they might not intersect the problem representations and form a general schema (depending on the degree to which such an intersect search requires controlled rather than automatic processing). Further, on the transfer problem, subjects in the varied story context training conditions were much slower at solving the problem than subjects in the non-varied story context group. On this final problem, the subjects in group DD/DD were beginning to show a decline in solution time; that is, they were beginning to show positive transfer. If these subjects had been given more training problems, they might be more likely to detect the similarities among problems and demonstrate a positive effect of the variability of training on transfer. In the second experiment, the training set was extended by one problem with the goal of enhancing transfer effects attributable to training variability.

Experiment 2

Method

Subjects

Seventy-seven undergraduates from the University of Colorado at Boulder participated in the study. Eighteen subjects were recruited through the school newspaper and posters and were paid five dollars each to participate in the study. All other subjects were students enrolled in an introductory psychology course in the summer of 1988 at the university and participated in the study for partial fulfillment of a course requirement.

Materials

The problems used in Experiment 1 were also used in this experiment, except that the values for one of the problems (Fred) were changed. One new problem was constructed for each of the Travel, Work, and Mixture categories. In this experiment, subjects solved three training problems and one transfer problem. Since subjects potentially could receive a set of problems with four different contexts (i.e., groups DS/DD and DD/DD) or relational terms (i.e., groups DS/DD, DD/DD), construction of a fourth set of problems with a new story context and relational terms was necessary. Four problems were constructed which contained the story context Map Scale. In these problems, a1 and a2 represented actual distances; b represented a map scale with map distance divided by actual distance; and x represented an unknown map scale with the same units as b. Also, a pair of relational terms were added: more than/less than. See Table 8 for an example problem from the Map Scale category.

Table 8

Example problem from Map story context and more than/less than relational terms used in Experiment 2

Story context: Map
Relational terms: more than/less than
Equation: $4(x+12)=10(x-12)$

E.Z. Munney had found two old pirates' maps. Both maps showed that there was a path straight from the site of his house to the buried treasure of Montezuma. Bluebeard's map was 12 feet per inch more than the normal map scale for the 16th century and was 4 inches long. Blackbeard's map was 12 feet per inch less than the normal map scale and was 10 inches long. E.Z. decided to use Blackbeard's map because it was bigger. He found the treasure and bought a new dishwasher.
(What was the scale in feet per inch for the normal maps of the 16th century?)

Design and Procedure

The design and procedure were identical to that utilized in the first experiment, except training consisted of three, rather than two, problems; each subject solved a total of four problems, with the fourth problem serving as the transfer problem.

Results

Five subjects were excluded from the analyses: three for not finishing any of the problems in less than 5 min, one for giving the correct numerical answer before writing an equation, and one who was given the wrong order of problems.

Variability groups

Time to equation. Times to equation were transformed with the natural log function, as in the first experiment. Statistical contrasts encoded the differences between consecutive problems (1 -1 0 0, 0 1 -1 0, and 0 0

1 -1 for the first, second, third, and fourth problems). All further analyses were performed in a manner identical to those of the first experiment, except as noted. The mean transformed times are plotted for each group and problem in Figure 3.

Across problems, the mean difference for the same story context groups was 0.69 (1.99 minutes) and for the different story context groups was 0.58 (1.79 minutes). The overall effect of context was significant, $\underline{F}(1,44)=5.37$, $p=.03$, $MSe=.76$. The relation and context by relation interaction effects were not significant (both $\underline{F}$'s<1).

The mean difference in solution time from the first to the second problem for the same story context groups was 1.20 (3.32 minutes) and for the different story context groups was 0.53 (1.70 minutes). Context was a significant predictor of the difference from the first to the second problem, $\underline{F}(1,44)=6.56$, $p=.01$, $MSe=.41$. There was a positive effect of having the same story contexts between the first two problems, relative to having different story contexts. The mean for the same relational terms groups was 1.05 (2.86 minutes) and for the different relational termsgroups was 0.68 (1.97 minutes). The effect of relational terms was marginal, but in the predicted direction, $\underline{F}(1,44)=2.01$, $p=.16$. As with context, subjects who had the same relational statements in the first two problems showed better transfer on the second problem as compared to subjects who had different relational statements in the first two problems. The context by relation interaction did not approach significance ($\underline{F}<1$). There was an overall improvement in time to the equation from the first to the second problem across groups of 0.86 (2.36 minutes), which was significant, $\underline{F}(1,44)=43.54$, $p<.001$, $MSe=.41$.

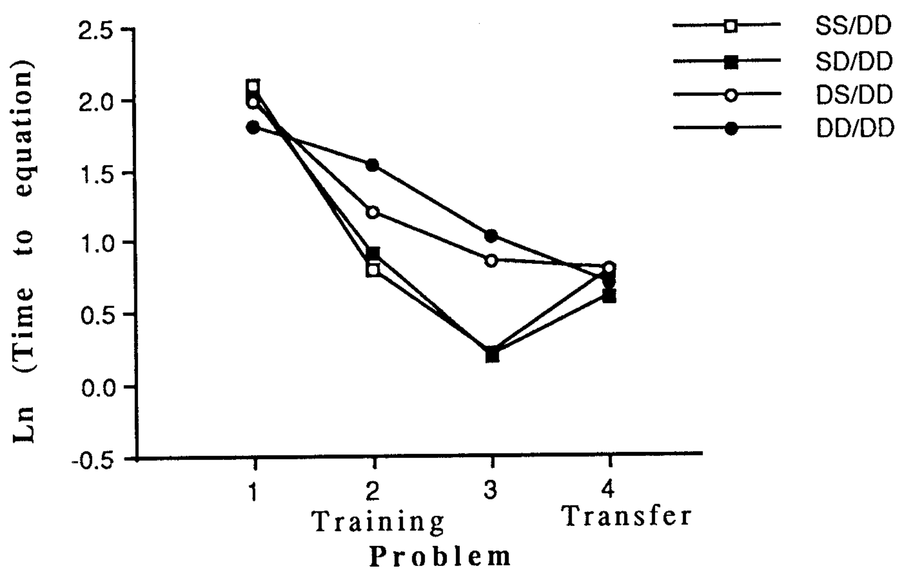

Figure 3

Mean log time to equation for Variability groups, Experiment 2

For the second contrast, the difference in time between the second and third problem, there were no significant effects of context, relational terms, or their interaction (all $F<1$). There was an overall improvement in time to the equation across groups from the second to the third problem of 0.54 (1.72 minutes), which was significant, $F(1,44)=20.93$, $p<.001$, $MSe=.33$.

For the final comparison, the difference in time between the third and transfer problems, the mean for the different story context groups was 0.20 (1.22 minutes); the mean for the same story context groups was -0.48 (an increment of 1.62 minutes from the third to the transfer problem). The effect of context was significant, $F(1,44)=9.92$, $p=.003$, $MSe=.27$. Subjects who had experienced varying contexts in training showed an advantage in transfer (from the third to fourth problems) over subjects who had not experienced variability of contexts in training. The effect of relational terms was not significant, $F(1,44)=1.05$, $p>.05$. The interaction of story context and relational terms did not approach significance ($F<1$). There was no overall improvement or decrement in performance across groups, $F(1,44)=1.75$, $p>.05$. The time to the equation for the transfer problem was regressed on story context, relational terms, and their interaction. None of these effects was significant (all F's<1).

Hints. As in the first experiment, few subjects solved the first problem without hints (between 0% and 17% across groups). On the second problem, nearly all subjects in group SS/DD solved the problem without hints (92%). A substantially smaller proportion of subjects in groups SD/DD and DS/DD solved the second problem without hints (58% for both groups), and even fewer subjects in group DD/DD reached the solution without hints (42%). On the third problem, all subjects in the same context training conditions (groups SS/DD and SD/DD) finished the problem without any hints. Most of the subjects in group DS/DD (83%) finished the problem without hints; however, group DD/DD still lagged behind the other groups, with 67% of the subjects reaching the equation without any hints. On the transfer problem, nearly all subjects in the varied context conditions (groups DS/DD and DD/DD) completed the problem without hints (92% for both groups); a somewhat smaller proportion of subjects in groups SD/DD (83%) and SS/DD (75%) reached the equation without hints. In summary, the proportion of subjects solving the problem without hints increased for groups SS/DD and SD/DD through the third problem, then dropped for the transfer problem. In contrast, the proportion of subjects requiring hints for solution for the groups with varied story context training increased with each successive problem. Group DS/DD surpassed group DD/DD through the first three problems, but the two groups had equal proportions of solution without hints on the transfer problem (see Table 9). The hint data were analyzed in the same manner as the time data. The pattern of results was the same, except there was no overall effect of context, $\underline{F}(1,44)=1.89$, $p=.18$, and the effect of context on the first to second problem difference was only marginal, $\underline{F}(1,44)=3.71$, $p=.06$.

Table 9

Proportion of subjects solving each problem without hints for Experiment 2, Variability groups

	Problem Number			
Group	Problem 1	Problem 2	Problem 3	Problem 4
SS/DD	0.08	0.92	1.00	0.75
SD/DD	0.00	0.58	1.00	0.83
DS/DD	0.08	0.58	0.83	0.92
DD/DD	0.17	0.42	0.67	0.92

Same statements. Seventy-five percent of the subjects in group SS/DD said that the second problem was the same as the first, and slightly fewer in group SD/DD (67%) made this statement. The groups with varied context in training had somewhat lower proportions of subjects making same statements (58% and 50% for groups DS/DD and DD/DD, respectively). On the third problem, nearly all subjects in the same context training conditions said that this problem was the same as previous problems (92% and 83% for groups SS/DD and SD/DD, respectively); as with the second problem, about half of the subjects in the variable context training groups made these statements (50% and 42% for groups DS/DD and DD/DD, respectively). For the transfer problem, the proportion of subjects making same statements in the same context training groups declined from problem three (75% for SS/DD and 58% for SD/DD). The proportion of subjects making this statement in the varied context training conditions remained at or near 50% (50% for DS/DD and 58% for DD/DD; see Table 10). The same statement data were analyzed in the same manner as the time data. The pattern of results was the same, except that whereas the time data showed a significant effect of context on the third to the transfer problem difference, the same statement data showed only a marginal effect, $F(1,44)=3.88$, $p=.06$. Also, the difference from the second to the third problem was not significant, $F<1$.

Table 10

Proportion of subjects stating that the problem was the same as previous or a previous problem for Experiment 2, Variability groups

	Problem Number		
Group	Problem 2	Problem 3	Problem 4
SS/DD	0.75	0.92	0.75
SD/DD	0.67	0.83	0.58
DS/DD	0.58	0.50	0.50
DD/DD	0.50	0.42	0.58

Transfer Cue Groups

Time to equation. The same contrasts utilized for the Variability group analyses were used for these analyses. As in Experiment 1, a group with the transfer condition of same story context and same relational terms was not part of the experiment, so the interaction between context and relation cannot be determined. The mean transformed times to equation for each group and problem are shown in Figure 4.

Across problems, the mean difference in times between problems was 0.51 (1.67 minutes) for the same story context groups and 0.37 (1.45

minutes) for the varied context groups. The effect of story context was marginal, $F(1,33)=3.90$, $p=.06$, $MSe=.85$. The effect of relation was not significant ($F<1$).

There was an overall improvement in time to the equation from the first to the second problem across groups of 1.10 or 3.00 minutes, which was significant, $F(1,33)=51.97$, $p<.001$, $MSe=.38$. For the difference in time between the first and second problems, neither context nor relation (F's<1.23) were significant. The mean improvement from the second to the third problem across groups was 0.68 or 1.97 minutes, which was significant, $F(1,33)=34.01$, $p<.001$, $MSe=.25$. Story context and relational terms were not significant predictors of the difference between the second and third problem (both F's<1). These results are not surprising, since all three groups had identical conditions for the first, second, and third problems.

For the difference in time from the third to the transfer problem, there was an overall decrement in performance of -0.47 or -1.60 minutes, which was significant, $F(1,33)=14.45$, $p=.001$, $MSe=.28$. Neither story context nor relational terms (F's<1.72) were significant predictors of the difference in time between the third and transfer problems. The time to the equation for the transfer problem was 0.09 or 1.09 minutes for the same story context group and 0.83 or 2.29 minutes for the different story context groups. The time to solve the transfer problem was regressed on story context and relational terms. The effect of story context was significant, $F(1,33)=4.21$, $p=.05$, $MSe=0.67$; subjects who had a similar story context in the transfer problem solved it faster than subjects with a different story context in transfer than in training. The effect of

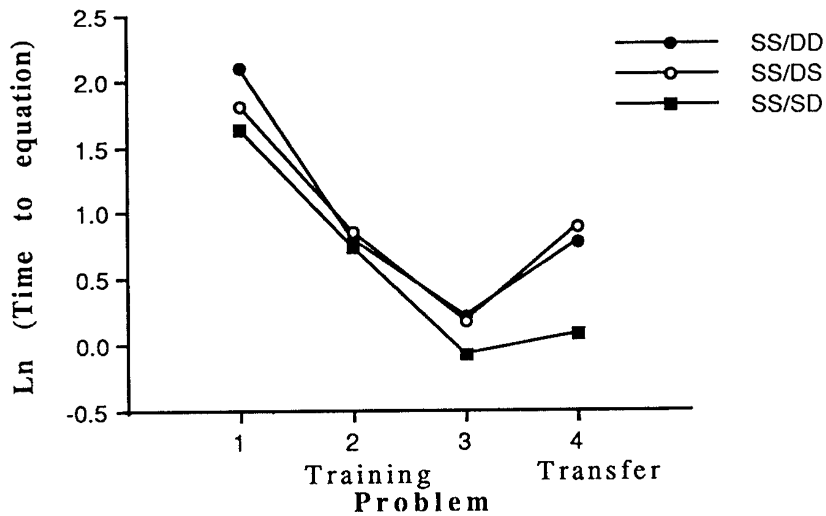

Figure 4

Mean log time to equation for Transfer Cue groups, Experiment 2 relational terms was not significant ($F<1$).

Hints. Few subjects solved the first problem without any hints (8%, 17%, and 25% for groups SS/DD, SS/DS, and SS/SD, respectively). On the second problem, nearly all subjects solved the problem without hints (92%, 83%, and 83% for groups SS/DD, SS/DS, and SS/SD, respectively); and on the third problem, all subjects in the three groups solved the problem without hints. For the transfer problem, somewhat fewer subjects solved the problem without hints (75%, 83%, and 92% for groups SS/DD, SS/DS, and SS/SD, respectively) than for the third problem. (See Table 11.) The pattern of results was the same as for the time data, except there was no overall effect of context, $F<1$.

Table 11

Proportion of subjects solving each problem without hints for Experiment 2, Transfer Cue groups

	Problem Number			
Group	Problem 1	Problem 2	Problem 3	Problem 4
SS/DD	0.08	0.92	1.00	0.75
SS/DS	0.17	0.83	1.00	0.83
SS/SD	0.25	0.83	1.00	0.92

Same statements. For the second problem, 75%, 83%, and 92% of the subjects said that the problem was the same as the first problem for groups SS/DD, SS/DS, and SS/SD, respectively. Nearly all subjects said that the third problem was the same as previous problems (92%, 83%, and

92% for groups SS/DD, SS/DS, and SS/SD, respectively). The proportion of subjects stating that the transfer problem was the same as the previous problems was reduced relative to the proportion for problem three in the groups with a transfer problem that had a different story context than the training problems (75% and 58% for groups SS/DD and SS/DS, respectively). The proportion of subjects in group SS/SD who said that the problem was the same as the other problems remained at the same high level (92%) as in the third problem (see Table 12). The pattern of results was the same as for the time data, except there was no overall effect of context, $\underline{F}(1,33)=1.13$, $p>.05$, and no significant difference between the second and third problems, $\underline{F}<1$.

Table 12

Proportion of subjects stating that the problem was the same as previous or a previous problem for Experiment 2, Transfer Cue groups

	Problem Number		
Group	Problem 2	Problem 3	Problem 4
SS/DD	0.75	0.92	0.75
SS/DS	0.83	0.83	0.58
SS/SD	0.92	0.92	0.92

Discussion

As in Experiment 1, problem similarity at the story context level aided transfer. Subjects whose first three problems contained the same story context showed a greater degree of transfer from the first to the second problem than subjects whose first three problems contained different story contexts. The same story context groups, however, did not show better transfer from the second to the third problem than the different context groups, suggesting the possibility that, after solving two problems, the different story context groups may have generalized across the problems. If so, the benefit of having a general schema was as substantial as that of having literal similarity of story contexts among training problems.

The effect of similarity at the relational term level was weak, with groups having the same relational terms in the first two problems showing greater transfer on the second problem than groups with different relational terms. Similarity of relational terms was helpful, however, only when story contexts differed (i.e., group DS/DD versus group DD/DD). When some other similarity existed to base transfer on (i.e., same story contexts), similarity of relational terms had no effect.

For the Transfer Cue groups (SS/DD, SS/DS, and SS/SD), similar story context in the transfer problem as in the first three problems produced less negative transfer than a different story context in the transfer problem (that is, group SS/SD was better than groups SS/DD and SS/DS). There seemed to be no transfer effect of having same or different relational terms in the transfer and the three training problems. However, note that all three groups (including SS/SD, where only relational terms differed in the transfer problem) showed negative transfer on the final problem.

Variability of story context enhanced transfer to a novel problem, but relational term variability did not. From the third to the transfer problem, groups with same story context training showed negative transfer, but groups with varied context training showed positive transfer. Similarity of problems at the story context level aids problem solving transfer when more problems with the same story context must be solved; however, when a subject is faced with a problem containing a novel story context, better transfer is likely if training has involved variability of story contexts.

Although subjects with same story context training showed negative transfer from the third to the final problem, they still solved this final problem in the same amount of time as the varied story context groups. Two factors are likely to contribute to this outcome. First, while the different story context groups showed continued improvement in performance throughout the four problems, they do not improve as rapidly as same context groups. It is possible that additional training problems are needed in order for subjects so trained to perform faster on the final problem than the same story context groups. Second, the relational terms in all problems were general; that is, these terms have overlapping meanings. Subjects might have been able to map the relational terms of the training problems onto those of the transfer problem, aiding transfer even in the situation where the transfer problem contained a novel story context. This may also explain why relational terms did not exhibit a powerful effect on transfer: different relational terms were easily mapped between problems, and the generality of the terms did not allow for a strong similarity versus non-similarity distinction. Subjects might have had a much more difficult time transferring to a problem with a new story context and new specific relational terms if all training problems contained relational terms specific to a particular story context (e.g., faster/slower for Travel problems). A design in which story context and relational terms are crossed as training conditions requires, however, that the manipulated relational terms apply to more than one story context; thus, the more general relational terms were required by the design of these experiments.

General Discussion

Theories of analogical transfer (e.g., Holyoak & Koh, 1987; Gick & Holyoak, 1983; Gentner, 1983) have used implicated features of a subject's mental representation of problems in their accounts of transfer. Their notions of representation are "flat" in the sense that they consist of one level of objects and their relations. Much of human cognition, however, appears to demand hierarchical representations (for example, goal structures in problem solving [Anderson, 1982; Anderson, Farrell, & Sauers, 1984], text comprehension [van Dijk & Kintsch, 1983], and natural categories [Rosch, 1975]). The representation of texts in which problems are typically embedded exists at multiple levels (e.g., situation model and text base, according to van Dijk and Kintsch); the nature of the representation at both of these levels should affect problem solving transfer. We found that similarities at the situation model (story context) level aided transfer. We also found that experiencing variability of problems at this level aided transfer to problems with a novel story context. There was some indication of a similar effect at the text base (relational term) level, but this effect was weak and, when it appeared, was secondary to the situation model effects.

In the past, researchers have focussed exclusively on the text base level in analogical transfer; they have focussed on the structure of the problem and have ignored the meaning level's influence on transfer. In particular, Gentner's (1983) theory implies that differences in analogical transfer are the result of differences in relational structure. However, although all problems in the current study had the same relational structure, we found differences in transfer due to manipulations of the story context, or meaning, of the problems. Gentner's model does not take this meaning level into account.

Holyoak and Koh's (1987) and Gick and Holyoak's (1983) analyses of transfer are based on notions of schema abstraction and thus would conform to our findings that variability of instances facilitated transfer to a novel problem. Holyoak and Koh's (1987) predictions of transfer are based upon similarity of the goals between problems. In the present studies a hierarchy of goals, not a single goal, was present. For example, the top level goal might be, "determine the equation underlying this problem." A sub-goal might then be, "find the important information", and an even lower sub-goal of, "find out what mathematical notion 'increase over' maps onto." In fact, the type of conceptual problems used by Holyoak (variants of Duncker's problem, where multiple forces converge to accomplish some task) probably also are solved using a hierarchical goal structure. For example, the top level goal might be, "find out how to accomplish the task," with a sub-goal of, "find relevant information in the problem." Holyoak and Koh are not specific about the level at which similarity of goals must exist to predict transfer. They also do not specify whether these goals are the subject's goals (as in the example above) or goals implicit in the problems. For example, in the problem of "Fred" travelling to Pueblo used in these experiments, the top level goal implicit in the problem is, "Find Fred's normal speed." Other sub-goals arise, including "find the distance travelled in two hours." Further, the situation model of a problem is not the same as the goal structure in either algebra word or conceptual problems, yet we found that this level of representation had strong effects on transfer.

Ross (1989) discussed two types of superficial similarity, story lines and object correspondences, and their effects on transfer. This corresponds roughly to the analysis in the present study, with story line being similar to story context, and object correspondences at the same level as relational terms. However, the situation model which is posited to cover the story context includes both the story line and the structure of the problem. Nevertheless, Ross found that similarity of corresponding objects enabled subjects to use the formula learned earlier; similarity of story lines affected access of the analogous problem/formula.

Another exception to the "flat" representation of problems is found in the research of Brown, Kane, and Echols (1986). They considered the representation of problems according to the van Dijk and Kintsch (1983) model and hypothesized that the situation model of a problem might be important in determining analogical transfer. Brown et al. found that when young children understood the important information in the situation model (e.g., the goals), they showed spontaneous transfer to an analogous problem. Subjects who only expressed superficial aspects of the problem did not show such transfer.

Information in the text base of a problem is also important to problem solving and analogical transfer. Our inability to find strong effects at the relational term or text base level may have been due to the generality of the terms used. Bassok and Holyoak (1989) found that subjects were able to transfer their knowledge of an arithmetic progression method learned in the context of algebra to new physics problems which could be solved with the same method; however, there was little transfer of the method from physics to algebra. The apparent reason for the asymmetry was that the algebra method was discussed in general terms (e.g., increase), but the analogous physics method was discussed in terms relevant only to motion (e.g., speed, acceleration). The relational terms used in the current experiments were relatively general and perhaps could be easily mapped between problems (even when the actual terms differed), as in Bassok and Holyoak's study.

The question we began with was, How do we aid students in transferring known solutions to novel problems? When the new and old problems are literally similar (e.g., they share the same objects or terms), we found that students showed a high degree of transfer (at least for similarity at the story context level). But transfer difficulty is greatest when the student is faced with a problem which is superficially and/or structurally unlike any they have solved before. We found that for structurally isomorphic problems, providing variability of instances during training aided transfer to a problem which differed from all of the training problems (again, this occurred mainly for variability at the story context level). Consistent with Gick and Holyoak's (1983) analysis of schema formation, our results appear to show that variability leads to the formation of a general schema with characteristics which enhance transfer to problems within the relevant domain.

The results of the present experiments are comparable to those in the categorization literature (e.g., Posner & Keele, 1968) which show that greater variability of training instances (instances were manipulated dot patterns) produces enhanced ability to categorize new instances. We suggest that particular slots in a schema may follow principles similar to those found in categorization. That is, the slots of a schema depend on the instances from which they were formed. For example, your experience with many types of dogs would cause the "color" slot in your dog schema to be general in the sense that many values fit the slot. One possibility is that a color slot contains both a prototypical dog color and the range of colors that might be expected. An object which matched the dog schema in all other ways (e.g., has fur, barks, walks on four legs) but was purple would therefore have somewhat doubtful status.

Current models of schema abstraction (Dellarosa, 1985) and generalization (e.g., Anderson, 1982) simplify this notion to two possible states: a slot contains specific information or it is variabilized. That is, there is no range of acceptable values and no prototypical value as part of the representation. The notion that slots preserve information about the instances from which they developed is supported in part by the performance of Group DD/DD in the present experiments. For subjects in this group, all problems contain a new story context and new relational terms. Subjects in this group showed continual improvement throughout the problem series. If subjects were merely variabilizing similar elements of specific problems to form the schema, we might expect performance to remain stable after the first two problems (by then, all relevant terms would be variabilized). On the other hand, if the slots are based on and retain information from particular instances, experiencing more problems with different information would further broaden the generalization of that slot and aid transfer to new problems.

This analysis is tentative, because there are several possible explanations for the continued improvement of group DD/DD beyond the second problem. For example, even after a stable schema is formed, subjects may show speed-up due to practice alone. Alternatively, as Dellarosa (1985) has proposed, the probability of noticing similarity of structure might increase with practice. The extent to which specific parts of problem schemas retain the character of the instances from which they were built remains an interesting possibility to be examined in follow-on research.

In the first experiment, only marginal effects of variability were found when two training problems were given. A stronger effect was found in the second experiment when three training problems were given. One important question raised by the current research is: How much training is necessary to show the positive effects of variability on transfer? Subjects may require either many training problems or an extensive discussion of how the solution is found in order to gain sufficient knowledge to transfer to a new problem. Additionally, subsequent problems may serve to broaden an already-formed schema, as discussed above. A detailed analysis of what a subject knows while solving problems would be the first step in further addressing this question of how much training is required for transfer.

Our results suggest that the dynamics of learning are important to understand. We need to find a delicate balance between giving students enough variation that they can apply their knowledge in new situations, but sufficient similarity to allow them to identify general principles, instead of learning each problem as a special case. In addition, the present research provides evidence that a hierarchical analysis of problems is necessary in order to explain analogical problem solving transfer fully. Variables affecting transfer (e.g., similarity, generalization) can independently affect each level of representation. Further, the meaning or the subject's model of the problem is at least as important as the information from the text base level in affecting transfer. Models which do not take world knowledge into account will not explain transfer sufficiently.

References

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J.R., Farrell, R., & Sauers, R. (1984). Learning to program in LISP. Cognitive Science, 8, 87-129.

Bassok, M. & Holyoak, K.J. (1989). Interdomain transfer between isomorphic topics in Algebra and Physics. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 153-166.

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

Dellarosa, D. (1985). Abstraction of problem-type schemata through problem comparison. Technical Report No. 146, Institute of Cognitive Science, University of Colorado.

Elio, R. & Anderson, J.R. (1981). The effects of category generalizations and instance similarity on schema abstraction. Journal of Experimental Psychology: Human Learning and Memory, 7, 397-417.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gick, M.L. & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M.L. & Holyoak, K.J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Hinsley, D.A., Hayes, J.R., & Simon, H.A. (1977). From words to equations: Meaning representation in algebra word problems. In

P. Carpenter & M. Just (Eds.), Cognitive Processes in Comprehension. Hillsdale, N.J.: Erlbaum Associates.

Holyoak, K.J. & Koh, K. (1987). Surface and structural similarity in analogical transfer. Memory & Cognition, 15, 332-340.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. Journal of Experimental Psychology: Human Learning and Memory, 7, 418-439.

Judd, C.M. & McClelland, G.H. (1989). Data Analysis: A Model Comparison Approach. San Diego: Harcourt Brace Jovanovich, Inc.

Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner & A.L. Stevens (Eds.), Mental Models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, J., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Mayer, R.E. (1982). Memory for algebra story problems. Journal of Educational Psychology, 74, 199-216.

Perfetto, G.A., Bransford, J.D., & Franks, J.J. (1983). Constraints on access in a problem solving context. Memory and Cognition, 11, 24-31.

Posner, M.I. & Keele, S.W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Reed, S.K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 106-125.

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

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104, 192-223.

Rosenthal, R. & Rosnow, R.L. (1985). Contrast Analysis: Focused Comparisons in the Analysis of Variance. Cambridge: Cambridge University Press.

Ross, B.H. (1984). Remindings and their effects in learning a cognitive skill. Cognitive Psychology, 16, 371-416.

Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 629-639.

Ross, B.H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 456-468.

Schank, R.C. & Abelson, R. (1977). Scripts, Plans, Goals, and Understanding. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

van Dijk, T.A. & Kintsch, W. (1983). Strategies of Discourse Comprehension. Orlando,Fl: Academic Press, Inc.

Author Notes

This research was supported in part by National Institute of Mental Health Training Grant Number MH14617-12 to the University of Colorado, by United States Army Research Institute Contract MDA903-86-K-0155, and United States Air Force Human Resources Laboratory Contract Number VE5744-022-001 to the Institute of Cognitive Science at the University of Colorado, and by Biomedical Research Support Grant RR07013-16 awarded to the University of Colorado by the Biomedical Research Support Grant Program (Division of Research Resources, National Institutes of Health).

Portions of this research were completed by the first author as partial fulfillment for the degree of MA from the Psychology Department at the University of Colorado, Boulder.

Correspondence concerning this article should be addressed to Lori Meiskey, Department of Psychology, University of Colorado, Campus Box 345, Boulder, CO 80309.