Memory & Cognition
1990, 18 (1), 83-98

Selecting analogous problems: Similarity versus inclusiveness

STEPHEN K. REED
San Diego State University, San Diego, California

and

CAROLYN C. ACKINCLOSE and AUDREY A. VOSS
Florida Atlantic University, Boca Raton, Florida

Students were asked to select one of two analogous problems in order to solve algebra word problems. In Experiment 1, one problem was less inclusive and the other was more inclusive than a test problem. The students judged the complexity and similarity of problems, selected analogous problems, and used the solutions to solve test problems. They performed significantly better on the test problems when given the more inclusive solutions, but used perceived similarity rather than inclusiveness to select analogous problems. The same pattern of results occurred in Experiment 2, in which isomorphic problems replaced the more inclusive problems. The results show that students are deficient in selecting good analogies, both from the same category (Experiment 1) and from a different category (Experiment 2). Students who saw the analogous solutions (Experiment 3) or were majoring in mathematics (Experiment 4) were more likely to select an isomorphic problem over a less inclusive problem, but were not more likely to select a more inclusive over a less inclusive problem.

A popular heuristic for solving problems is to use the solution to a related or analogous problem. Psychologists have used several different experimental procedures to study how effectively people can use the solution of one problem to solve another problem. One procedure is to ask people to solve two related problems to determine whether solving the first problem will result in faster solutions for the second problem (Reed, Ernst, & Banerji, 1974). Another procedure is to give students a correct solution to a problem to determine whether they will subsequently use the solution to solve an analogous problem (Gick & Holyoak, 1980). A third procedure is to allow students to refer to a detailed solution as they attempt to solve a related problem (Reed, Dempster, & Ettinger, 1985).

Each of these paradigms allows the investigators to examine how successfully subjects can use the provided solution. However, research on analogy has usually not emphasized how students would select a potentially useful solution if they were allowed to make the choice. In a recent paper on analogical problem solving, Holyoak and Koh (1987) have identified four basic steps in transferring knowledge from a source domain to a target domain: (1) construction of mental representations of the source and the target; (2) selection of the source as a potentially relevant analogue to the target, (3) mapping of the components of the source and target; and (4) extension of the mapping to generate a solution to the target. They state that the second step, selection of a source analogue, is perhaps the least understood of the four.

Our purpose in this study was to investigate how people select an analogous problem. Imagine that you are given a problem to solve and can see the solution to a related problem. But first you must choose the problem that would provide the most useful solution. Which problem would you choose?

One variable that should influence how students select an analogous problem is the perceived similarity of two problems. In several studies, sorting tasks have been used to measure how students perceive problems. Silver (1979) investigated the relationship between students' ability to classify mathematical word problems in a sorting task and their performance on tests of problem solving ability. He found that good problem solvers sorted problems according to their mathematical structure, but that poor problem solvers were more influenced by story context. Chi, Glaser, and Reese (1982) found similar results in a study of how novices and experts sort physics problems. Schoenfeld and Herrmann (1982) extended these results by looking more directly at the shift in novices' perceptions of problems after a month-long intensive course in mathematical problem solving. Students' perceptions of the

This research was supported by National Science Foundation Grant BBS-8312763 and Air Force Office of Scientific Research Grants AFOSR 88-0008 and AFOSR 89-0107 to Stephen K. Reed. Part of this article is based on Carolyn C. Ackinclose's MA thesis, which was supervised by S. K. Reed. We are grateful to Javier Sanchez for his assistance with Experiment 4, and to David Bjorklund, Edward O'Brien, Rob Nosofosky, and anonymous reviewers for their comments. Requests for reprints should be sent to Stephen K. Reed, Department of Psychology, San Diego State University, San Diego, CA 92182.

—Accepted by previous editor, Alice F. Healy

mathematical structure of problems were examined before and after the course, revealing that, following instruction, their classification of problems was more influenced by the mathematical structure.

Although relevant, these studies do not directly require that students select an analogous problem from a specified set of problems. In the present study, we examined the choice process more directly, by asking students to indicate which of two analogous problems they would prefer to use in order to solve a related problem. Several variables, including perceived similarity, were investigated to determine how they influence subjects' selections.

Inclusiveness Versus Similarity

The problems used in Experiment 1 of the present study differ from those used in the sorting studies, because students had to select from among problems with the same story context but slightly different solutions. Previous studies have usually focused on the recognition of isomorphic problems—problems that have identical solutions but different story contexts. How does one select a potentially useful solution if none of the solutions are identical to the required solution? We propose that if one solution is less inclusive than the required solution and another solution is more inclusive than the required solution, the more inclusive solution will be more useful. One solution is more inclusive than the other, according to our definition, if it contains all the information needed to solve the less inclusive problem, plus some additional information.

A more precise definition of inclusiveness can be formulated by using Gentner's (1980) concept of target exhaustiveness. According to her structure-mapping theory, concepts and relations in a base domain are mapped to concepts and relations in a target domain. *Target exhaustiveness* refers to the proportion of relational predicates in the target domain (i.e., test problem) that can be mapped back to the base (i.e., solution). All of the predicates in the test problem can be mapped onto the solution if the solution is more inclusive than the test problem, but only some of the predicates in the test problem can be mapped onto the solution if the solution is less inclusive than the test problem. In addition, a more inclusive solution will contain some predicates that are not included in the test problem. A more inclusive solution therefore provides extraneous information and a less inclusive solution provides too little information.

Differences in inclusiveness can be illustrated by the three cost problems in Table 1. The first problem requires using only the standard formula for calculating average cost. The second problem can be solved by equating the average cost of a ticket for the two groups. The third problem requires incorporating the discounted cost into the equation. Equation 3 was created by adding new relations to Equation 2, which was created by adding new relations to Equation 1. In general, Problems 1, 2, and 3 in Table 1 were created by adding new relations to the previous problem, and they are therefore ordered from the least inclusive to the most inclusive. The three problems in each category allow us to compare how the inclusiveness of a solution (Problem 1 vs. Problem 3) influences students' ability to solve a test problem (Problem 2).

A Feature Representation of Word Problems

We will use a feature representation of word problems to show how problems differ in inclusiveness. The subject's task is to use an equation provided in a solution to construct an equation for a related problem. We therefore propose that the relevant features of a problem are the concepts represented by the numerical values that are needed to construct the equation. Table 2 shows these concepts for each of the problems in Table 1.

For example, solving the first cost problem requires using the total cost of the tickets and the price of each ticket. Solving the second problem requires using the total cost for the smaller group, the total cost for the larger group, and the number of additional people in the larger group. Solving the third problem requires using the total cost for the smaller group, the total cost for the larger group, the number of additional people in the larger group, and the discount for the larger group.

The third problem is more inclusive than the second problem in each of the four categories, because it contains more features than the second problem, and because all the features in the second problem are a subset of the features in the third problem. Notice that the first problem is not more inclusive than the second, because neither of these two conditions are met.

We will use the term *less inclusive* to describe the relation of the first problem to the second problem, although ideally, all the features of the first problem should be a subset of the second problem for it to be less inclusive. Note, however, that all the features in the first problem are used to solve the second problem, but they have to be computed. The price of a ticket, distance traveled, distance from the fulcrum, and amount of the task completed by one of the workers are quantities in the first problem that have to be computed from two quantities (or one quantity and the unknown variable) in the second problem. Solving the second problem therefore requires that students know how to relate features that are not provided in the first problem.

Although we have defined inclusiveness in terms of features rather than relations, as specified in Gentner's (1980) definition of target exhaustiveness, we believe that this is a relatively minor distinction for this study. The quantities in an equation are related to each other by the arithmetic operations of addition, subtraction, multiplication, and division. When an additional feature is added to a problem, there is a corresponding relation that specifies how the new quantity should be incorporated into the equation. In fact, constructing equations is challenging, because students must learn how to use arithmetic operations to formally relate quantities in a problem. The common features in our problems are accompanied by com-

Table 1
Problems Used in Experiment 1

Category	Problem
Cost	1. A group of people paid \$238 to purchase tickets to a play. How many people were in the group if the tickets cost \$14 each? $\$14 = \$238/n$
	2. A group of people paid \$306 to purchase theater tickets. When 7 more people joined the group, the total cost was \$425. How many people were in the original group if all tickets had the same price? $\$306/n = \$425/(n+7)$
	3. A group of people paid \$70 to watch a basketball game. When 8 more people joined the group the total cost was \$120. How many people were in the original group if the larger group received a 20% discount? $.8\times(\$70/n) = \$120/(n+8)$
Distance	1. A pilot flew 1,575 miles in 7 hours. What was his rate of travel? $1{,}575 = r\times 7$
	2. A pilot flew from City A to City B in 7 hours but returned in only 6 hours by flying 50 mph faster. What was his rate of travel to City B? $r\times 7 = (r+50)\times 6$
	3. A pilot flew his plane from Milton to Brownsville in 5 hours with a 25-mph tailwind. The return trip, against the same wind, took 1 hour longer. What was the rate of travel without any wind? $(r+25)\times 5 = (r-25)\times 6$
Fulcrum	1. Laurie weighs 60 kg and is sitting 165 cm from the fulcrum of a seesaw. Bill weighs 55 kg. How far from the fulcrum must Bill sit to balance the seesaw? $60\times 165 = 55\times d$
	2. Tina and Wilt are sitting 4 meters apart on a seesaw. Tina weights 65 kg, and Wilt weighs 80 kg. How far from the fulcrum must Tina sit to balance the seesaw? $65\times d = 80\times(4-d)$
	3. Dan and Susie are sitting 3 meters apart on a seesaw. Mary is sitting 1 meter behind Susie. Dan weighs 70 kg, Susie weighs 25 kg and Mary weighs 20 kg. How far from the fulcrum must Susie sit to balance the seesaw? $20\times(d+1) + 25\times d = 70\times(3-d)$
Work	1. Tom can mow his lawn in 1.5 hours. How long will it take him to finish mowing his lawn if his son mowed ¼ of it? $.67\times h + .25 = 1$
	2. Bill can paint a room in 3 hours and Fred can paint it in 5 hours. How long will it take them if they both work together? $.33\times h + .20\times h = 1$
	3. An expert can complete a technical task in 2 hours but a novice requires 4 hours to do the same task. When they work together, the novice works 1 hour more than the expert. How long does each work? $.50\times h + .25\times(h+1) = 1$

mon relations, and the introduction of new features in the more complex problems requires the problem solver to form additional relations.

Hypotheses

The objective of this study was to test two hypotheses concerned with the selection and application of inclusive solutions. The hypothesis regarding application is that a more inclusive solution (Problem 3) should be more useful than a less inclusive solution (Problem 1). This hypothesis is based on the finding that students have considerable difficulty in representing relations among variables (Reed & Ettinger, 1987). Generating the relations that are missing from the less inclusive solution should therefore be a difficult task that will limit students' performance on the test problem. In contrast, the use of a more inclusive solution requires that students ignore excess information. Although this is a nontrivial task, it should be easier than generating new mathematical relations.

The hypothesis regarding students' selections is that they should select the more inclusive solution in order to solve the test problem. If the first hypothesis—that more inclusive solutions are more useful—is correct, then students should choose the more inclusive solution in order to perform well. However, students' choices might be influenced by other factors, such as the perceived complexity of a problem or its similarity to the test problem.

The selections were investigated in Experiment 1 by showing students problems rather than the actual solutions. This experiment was therefore similar to other

Table 2
A Feature Representation of the Problems in Experiment 1

Category	Problem	Features						Complexity: Least	Complexity: Most
Cost		$Cost_1$	Price	$Cost_2$	Additional People	Discount			
	1	$238	$14					61	0
	2	$306		$425	7			2	3
	3	$ 70		$120	8	20%		0	60
Distance		Distance	Time	Return Time	Speed Change	Return Change			
	1	1,575	7					62	1
	2		7	6		+50		0	13
	3		5	6	+25	−25		1	49
Fulcrum		$Weight_1$	$Weight_2$	$Weight_3$	$Distance_{12}$	$Distance_{1F}$	$Distance_{23}$		
	1	60	55			1.65		26	5
	2	65	80		4			35	0
	3	70	25	20	3		1	3	59
Work		$Rate_1$		$Rate_2$	Tasks Completed	Time Difference			
	1	1.5			¼			31	19
	2	3		5				25	11
	3	2		4		1		8	34

Note—Subscript "12" = distance between weights 1 and 2; subscript "1F" = distance between weight 1 and fulcrum; subscript "23" = distance between weights 2 and 3.

studies in which students have been required to classify their problems according to inferred solutions (Chi et al., 1982; Schoenfeld & Herrmann, 1982; Silver, 1979). However, as stated previously, these other studies have examined whether students could recognize isomorphic problems, whereas the problems in Table 1 have different solutions that vary in inclusiveness. In Experiment 2, we allowed students to choose between a less inclusive problem and an isomorphic problem, in order to compare our findings with the findings of previous studies. And, finally, in Experiment 3 and 4, we investigated how mathematical experience and the opportunity to study the solutions would influence the selections.

EXPERIMENT 1

We studied the effect of inclusiveness by giving a questionnaire to students in a college algebra class. In the first part of the questionnaire, the students selected the solution they would prefer to use to solve each of a variety of test problems. This allowed us to evaluate the hypothesis that students would select the more inclusive solutions. They also rated the complexity and similarity of the problems. We collected complexity ratings, because we thought the inferred complexity of a solution might influence students' selections and because we wanted to determine whether perceived complexity corresponded to the inclusiveness of a solution. We collected similarity ratings, because similarity of the story context influences which problems students are reminded of when solving test problems (Gentner & Landers, 1985; Ross, 1984) and which problems novices group together in clustering experiments (Chi et al., 1982; Schoenfeld & Hermann, 1982; Silver, 1979).

In the second part of the experiment, the students received either a less inclusive solution or a more inclusive solution for each of the four test problems. This allowed us to evaluate the hypothesis that a more inclusive solution is more useful than a less inclusive solution.

Method

Subjects. The subjects were 64 students enrolled in three sections of a college algebra course at Florida Atlantic University. The subjects were tested in class several weeks after the beginning of the fall semester. They had studied some computational problems, such as multiplying and dividing polynomials, but they had not yet studied word problems in the course.

Procedure. The first page of the questionnaire contained the following instructions:

> The purpose of this study is to evaluate some instructional material on algebra word problems. We are interested in determining how useful the solution of one problem is for helping students solve a similar problem.
>
> In the first part you will see sets of 3 problems and will be asked questions about these problems. For example, which of the 3 problems has the simplest solution, or which solution would be the most useful? In the second part you will be asked to write an equation for solving a problem. We will then give you another opportunity, after showing you the correct equation for a similar problem.
>
> We would like you to spend 3 minutes on each page and will tell you when to turn the page. You can use the timer to pace yourself. The study takes 50 minutes and includes problems that you may encounter in this course. Your participation will help us with our research and give you a chance to practice on word problems. We will provide you with the correct answers at the end of the study.

The problems used in this experiment were from the four categories shown in Table 1. Each of the first four pages of the test booklet contained the set of three problems from one of these categories (without equations or solutions) followed by a series of six questions. The following questions were the same for all problems:

1. Assume that you do not know how to solve any of the three problems. For each problem, which of the other two solutions do you think would be most useful?
   To solve Problem A I would prefer the solution to Problem B or C
   To solve Problem B I would prefer the solution to Problem A or C
   To solve Problem C I would prefer the solution to Problem A or B
2. Which problem do you think has the most complex solution?
3. Which problem do you think has the least complex solution?
4. Which problem do you think is the most similar to Problem A?
5. Which problem do you think is the most similar to Problem B?
6. Which problem do you think is the most similar to Problem C?

The order of the three problems on each page was 1, 3, 2 for the cost problems; 2, 1, 3 for the distance problems; 2, 3, 1 for the fulcrum problems; and 3, 2, 1 for the work problems, with 1 being the least inclusive and 3 being the most inclusive. The first problem presented was labeled A; the second, B; and the third, C.

The students were allowed to work for 3 min on each of the first four pages. If they completed a page before the 3-min time limit, they were allowed to complete unanswered questions on previous pages but were not allowed to go forward.

The students were next asked to identify from among six alternatives the best description of their strategy for selecting solutions on the previous four pages. The alternatives were listed in the reverse order on half of the test booklets. The students were allowed 3 min to perform this task.

The second part of the experiment measured how well the students could utilize a solution to a related problem to construct an equation for a test problem. All students received test problems of intermediate inclusiveness (Problem 2), which were accompanied by a solution to a problem that was either less inclusive (Problem 1) or more inclusive (Problem 3) than the test problem. The students worked on the test problem for 3 min, studied the related solution for 2 min, and then worked on the test problem for an additional 3 min while they referred to the solution on the facing page of the test booklet.

The test problems were from the four categories in Table 1 and occurred in a random order. Each subject received two solutions that were less inclusive and two solutions that were more inclusive than the test problem.

The solutions contained the equations shown in Table 1, accompanied by a detailed explanation. A less inclusive solution had the same format as a more inclusive solution, including a table to summarize quantities and variables. An example of the two solutions for the distance problem is shown in the Appendix. The less inclusive solutions were more elaborate than necessary, in order to facilitate their generalization to the more inclusive test problems.

In order to determine if the students' strategies for selecting solutions had changed after completing the problem-solving phase, they were again asked to identify their preferred strategy, using their experience from working with the solutions. The strategies were listed in a different order from that for the first presentation. Three minutes were allotted for this task.

Results

The results are divided into several sections because of the large amount of data. We first report the complexity judgments and compare complexity and inclusiveness. We next report the similarity judgments and show how these judgments correlate with students' preferred solutions. The following section presents a major part of the results: an evaluation of how complexity, similarity, and inclusiveness influence students' selection of analogous solutions. We then examine students' reported strategies to determine how their reports correspond to their choices. The final section contains the data on the usefulness of solutions; how successfully students can use either a less inclusive or a more inclusive solution to solve the test problem.

Complexity judgments. Table 2 shows how many students judged each problem to be either the least complex or the most complex problem in its category. The results show a good correspondence between judged complexity and inclusiveness for the cost and distance problems, but a poor correspondence for the fulcrum and work problems. For the latter two categories, students judged Problems 1 and 2 to be about equally complex, even though the equation for Problem 2 contained more symbols than the equation for Problem 1 (remember, however, that students had to judge inferred solutions).

A good predictor of the judged complexity of a solution is the number of features in the problem. For the cost and distance categories, Problem 1 has the fewest features and Problem 3 has the most features. The data supported the prediction that students would select Problem 1 as the least complex and Problem 3 as the most complex. For the fulcrum and work categories, Problem 3 has the most features, but Problems 1 and 2 have the same number of features. The prediction that students would select Problem 3 as the most complex but be divided between Problems 1 and 2 as the least complex was also supported by the data.

Similarity judgments. Table 3 shows how many students selected each of two problems as the more similar to a third. It also shows how many students selected each

Table 3
Solution Preferences and Similarity Choices in Experiment 1

	Categories							
	Cost		Distance		Fulcrum		Work	
Selection	P	S	P	S	P	S	P	S
			Problem 1					
Problem 2	55	54	53	63	47	53	28	28
Problem 3	7	9	9	0	16	11	32	34
			Problem 2					
Problem 1	31	31	28	18	47	43	13	11
Problem 3	32	32	35	45	17	21	50	53
			Problem 3					
Problem 1	12	8	11	5	20	18	13	14
Problem 2	49	54	50	57	41	44	49	49

Note—The data show how many subjects selected each problem as providing the preferred solution (P) and how many subjects selected each problem as being more similar (S) to the specified problem.

of two problems as the preferred solution for solving the third problem.

We expected that students would usually select Problem 2 as the problem more similar to both Problems 1 and 3. The students' selections confirmed our expectations for seven of the eight cases. We did not expect that the students would consistently select either Problem 1 or Problem 3 as the problem more similar to Problem 2. The students' similarity judgments, in fact, varied across the four categories. They judged Problem 3 as the more similar problem for the distance and work categories and Problem 1 as the more similar problem for the fulcrum category, and they evenly divided their selections for the cost category.

We will present later a model that accounts for the similarity judgments, but our immediate concern is to evaluate how perceived similarity influences students' selections of analogous solutions. A comparison between solution preferences and judged similarity reveals a very close correspondence between the number of students who preferred a particular solution and the number who rated that problem as more similar ($r = 0.97$). This high correlation suggests the importance of similarity in determining preferences for solutions.

In the next section, we examine the similarity judgments and preferences for individual students, in order to provide a more direct measure of how perceived similarity influences selections. We also examine how students' complexity judgments influence their choice of preferred solutions. In order to compare both of these subjective measures with problem inclusiveness, we limit this analysis to determining how students selected solutions for Problem 2.

Selecting solutions for Problem 2. There are two reasons why the students' selection of solutions for Problem 2 are particularly relevant. First, the students had to choose between a problem that was less inclusive than Problem 2 and a problem that was more inclusive than Problem 2. We can therefore determine which variables (complexity, inclusiveness, similarity) influenced their selections. Second, the students had to use these solutions to solve Problem 2. By comparing their selections with the actual usefulness of the solutions, we can determine whether or not the students selected useful solutions.

Table 4 shows how complexity, inclusiveness, and similarity influenced the selection of solutions for solving Problem 2. The data are the number of students who selected the problem each judged as either less complex or more complex than, and as either less similar or more similar to, the test problem. The numbers of students who selected either the less inclusive or the more inclusive solution are also shown. The results show that neither complexity nor inclusiveness had a consistent effect on the students' preferences. The hypothesis that students would select the more inclusive solution was clearly not supported. They showed a significant preference only for the more inclusive work problem, which was balanced by their significant preference for the less inclusive fulcrum problem. In contrast, the students showed a consistent preference across all four categories for the problems they judged as more similar to the test problem.

An analysis of the results by students, rather than problems, yielded the same conclusions. In order to determine if there was a significant preference for solutions on the basis of complexity, inclusiveness, or similarity, a preference score for each student was calculated by subtracting the number of times he or she selected the less complex, inclusive, or similar solution from the number of times he or she selected the more complex, inclusive, or similar solution. The preference scores could range from −4 to +4 across the four problems. The mean score for complexity was −0.31, and the standard deviation was 2.28 [$t(57) = 1.02$, $p > .05$]. The mean score for inclusiveness was 0.17, and the standard deviation was 2.12 [$t(58) = 0.60$, $p > .05$]. These results show that neither complexity nor inclusiveness had a significant impact on the students' selections. In contrast, the mean score for similarity was 2.36 and the standard deviation was 1.65 [$t(59) = 11.09$, $p < .01$], indicating a significant preference for the more similar solution.

Report of strategies. After making their selections, the students were asked to choose from among six strategies the one that best described how they made their selections. Twenty-five subjects said they would choose the more similar solution regardless of its complexity, 18 subjects said they would choose the less complex solution because it would be easier to understand, 14 subjects said they would choose the more complex solution because it would more likely contain the information needed to solve the problem, and 6 subjects said they would choose the less complex solution because simpler solutions are usually taught before complex solutions. None of the subjects said that they would choose the more complex solution because it would be more challenging, or reported using an alternative strategy that differed from the listed strategies.

Table 4
Effects of Complexity, Inclusiveness, and Similarity on Selecting Solutions in Experiment 1

	Complexity			Inclusiveness			Similarity		
Category	Less	More	z score	Less	More	z score	Less	More	z score
Cost	31	32	0	33	28	.51	19	44	3.03*
Distance	27	36	1.01	28	35	.78	18	45	3.28*
Fulcrum	51	11	5.08*	47	17	3.63*	12	52	4.88*
Work	30	32	.25	13	50	4.55*	8	55	5.79*

*Significant at the $p < .01$ level.

Subjects also received the strategy questionnaire after attempting to use solutions to solve the problems, in order to determine whether their strategies would change. Essentially the same distribution of responses occurred after the subjects had the opportunity to use some of the solutions. There is a correspondence between students' selections and their verbal reports. The reported preference for similar solutions is consistent with their selections, and the divided reports for less as opposed to more complex solutions are consistent with the finding that neither inclusiveness nor complexity had a general effect on their selections.

Using solutions. The students' failure to select more inclusive solutions would be of little interest if such solutions did not facilitate problem solving. Providing a less inclusive solution resulted in an improved performance from 9% correct equations on Trial 1 to 17% correct on Trial 2. Providing a more inclusive solution resulted in an improved performance of 8% correct equations on Trial 1 to 33% correct on Trial 2. The students therefore solved about twice as many problems with a more inclusive solution than with a less inclusive solution; this was a significant difference [$t(63) = 3.27$, $p < .01$].

Because the students used similarity, rather than inclusiveness, to select solutions, it would be informative to compare whether or not the more similar solutions are more helpful than the less similar solutions. The results suggest that the perceived similarity of a problem did not influence the effectiveness of a solution. The subjects solved 26% of the test problems when they were given a solution to the more similar problem (as determined by each subject) and 24% of the problems when given the solution to the less similar problem.

However, it is difficult to make a statistical comparison by subjects, because the similarity judgments were not used to assign solutions to subjects. Thus 1 subject could receive all four solutions from the problems he or she judged as less similar to the test problems, whereas another subject could receive all four solutions from the problems he or she judged as more similar to the test problems.

We therefore evaluated the effect of problem similarity on the successful use of solutions by comparing the two levels of similarity for each of the problem categories. Table 5 shows whether students significantly improved their performances on Trial 2 when they received a solution to the problem that they judged as either the less or the more similar to the test problem. The data clearly show that having the solution to the more similar problem was not an advantage. Neither group significantly improved its performance on the cost problem, both groups significantly improved their performance on the fulcrum and distance problems, and only the students who received the solution to the less similar problem improved their performance on the work problem. In contrast, only the students who received the more inclusive solution significantly improved their performances on the cost, distance, and fulcrum problems, although the less inclusive solution was the more effective for the work problem (see Table 6).

Table 5
Effect of Problem Similarity on the Successful Use of Solutions in Experiment 1

Category	Problem Similarity	Percent Correct		z score
		Trial 1	Trial 2	
Cost	Less	17	24	1.00
	More	21	35	1.51
Distance	Less	10	23	2.24*
	More	13	33	2.24*
Fulcrum	Less	5	31	3.16*
	More	0	28	2.65*
Work	Less	0	19	2.45*
	More	0	9	1.73

*Significant at the $p < .05$ level.

Discussion

The results of this experiment demonstrated that more inclusive solutions are more useful for solving algebra word problems than less inclusive solutions are. However, students had a significant preference for the more inclusive solution only for the work problem. Ironically, this was the only problem in which the more inclusive problem was not helpful, thus revealing an evident lack of correspondence between perceived and actual usefulness of solutions.

Perceived similarity of problems controlled the selection of solutions, as is indicated by both the selections made by subjects and their reports of how they made their selections. However, unlike the more inclusive solutions, solutions that were judged as more similar to the test problem were not more effective than the less similar solutions. A practical consequence of this finding is that students need to modify how they select analogous solutions. Rather than select solutions on the basis of perceived similarity, students should use a principle such as inclusiveness as the basis for their selections.

The distinction between using similarities and using principles to make decisions has recently emerged in the categorization literature (Barsalou, 1985; Murphy & Medin, 1985). Barsalou (1985) has argued that similarity is less important in goal-derived categories than it is in common taxonomic categories. For example, the

Table 6
Effect of Solution Inclusiveness on the Successful Use of Solutions in Experiment 1

Category	Solution Inclusiveness	Percent Correct		z score
		Trial 1	Trial 2	
Cost	Less	19	19	0
	More	19	41	2.11*
Distance	Less	13	19	1.41
	More	9	34	2.83*
Fulcrum	Less	3	13	1.73
	More	3	47	3.74*
Work	Less	0	19	2.45*
	More	0	9	1.73

*Significant at the $p < .05$ level.

category *things to take on a vacation* consists of objects that may look dissimilar even though they share a common "goal."

Some theories of analogical reasoning have also emphasized the goal-directed nature of the learning process (Carroll & Mack, 1985; Holyoak, 1985). According to this view, the useful aspects of an analogy can vary, depending on a person's goals. As an example of how goals can vary, contrast the proposed principle of inclusiveness with the heuristic of trying to solve and then generalize a simpler version of the problem if one cannot solve a problem (see Schoenfeld, 1979, for an application of this heuristic). The difference is that students in our experiment were told that they would be shown an analogous solution, whereas students who initially try to solve a simplified version of a problem do not have access to any solutions. If the problem solver did not have access to solutions, he or she would not attempt to solve a more inclusive problem that would be more complex than the test problem. In this case, attempting to solve and generalize a simpler version of the problem might be a reasonable heuristic, although our results suggest that it can be very difficult to use a simple solution to solve a more complex problem.

EXPERIMENT 2

The results of Experiment 1 revealed that although more inclusive solutions were more effective than less inclusive solutions, students did not show consistent preferences for the more inclusive solutions. Instead, they chose solutions on the basis of perceived similarity to the test problems. The purpose of Experiment 2 was to determine whether the same pattern of results would occur if the more inclusive solutions were replaced by isomorphic solutions. Two solutions are isomorphic if they have different story contexts but are represented by structurally identical equations. Problems 2 and 3 are isomorphic to each other for each of the four categories in Table 7.

We chose to compare isomorphic solutions with less inclusive solutions, because we believed this comparison would produce the same pattern of results as had been obtained in Experiment 1. We expected that the solution to the isomorphic problem would be more useful, because the solution to the less inclusive problem would lack information required to solve the test problems. We could not make clear predictions for the case in which isomorphic problems are contrasted with problems that are more inclusive than the test problem. The tradeoff in this comparison is one of processing the excess information in the more inclusive solution versus finding corresponding concepts in the isomorphic solution (see Reed, 1987, for a discussion of mapping concepts across isomorphic problems).

Experiment 2 allowed us to test the hypothesis that students would select a solution on the basis of perceived similarity, which would usually be greater for the less inclusive problems that had the same story context. We therefore anticipated that, when given the problems shown in Table 7, students would prefer the solution to Problem 1, rather than the solution to Problem 3, in order to solve Problem 2. We also hypothesized that the less inclusive solutions would be less effective than the isomorphic solutions, and we therefore expected to replicate the discrepancy found in Experiment 1 between the perceived and actual usefulness of a solution.

The basis for the first prediction (that students would use story context to select problems) is that the recognition that two solutions are isomorphic usually requires a considerable expertise that is likely to be lacking among most students enrolled in college algebra classes. The research of Silver (1979), Chi et al. (1982), and Schoenfeld and Herrmann (1982) suggests that novices are likely to be more influenced by story content than by mathematical structure when judging the similarity of two problems.

The basis for the second prediction is that although students may fail to notice isomorphic relations, they can often effectively use an isomorphic solution if told of its value (Gick & Holyoak, 1980; Reed, 1987). Gick and Holyoak (1980) distinguished between students' ability to notice an analogy and their ability to apply an analogy when told of its potential usefulness. When hints to use an analogous solution to solve a problem were given, subjects were successful in generating analogous solutions; when no hints were given, the frequency of analogous solutions decreased markedly. Reed (1987) also found that students had difficulty in spontaneously noticing isomorphic relations in word problems. However, when given analogous solutions, the students did significantly better in using isomorphic solutions than in using solutions that had the same story context but only a similar mathematical procedure.

Method

Subjects. The subjects consisted of 52 undergraduates enrolled in two sections of a college algebra class at Florida Atlantic University. They were tested in class several weeks after the beginning of the spring semester, and they had not yet studied word problems in the course.

Procedure. The procedure was identical to that used in Experiment 1. The students were asked to evaluate sets of problems for perceived complexity, similarity, and usefulness of solutions; to identify their strategy for selecting solutions; and to construct equations for test problems. During the test phase, each student received two less inclusive and two isomorphic solutions. Table 7 shows the problems used in Experiment 2. Problem 1 is a less inclusive problem, Problem 2 is the test problem, and Problem 3 is an isomorphic problem. The distance and work categories, which provided opposite results in Experiment 1, were combined with two new problem categories, interest and mixture. We replaced the cost and fulcrum problems in Experiment 1, because we were unable to think of isomorphic versions of these problems. An example of an isomorphic solution (for the distance problem) is shown in the Appendix.

Results

Complexity judgments. Table 8 shows the number of students who selected each of the three problems as either the least complex or the most complex problem in that category. As indicated in Table 7, Problem 1 has the least

Table 7
Problems Used in Experiment 2

Category	Problem
Distance	1. A pilot flew 1,575 miles in 7 hours. What was his rate of travel? $1{,}575 = r \times 7$
	2. A pilot flew from City A to City B in 7 hours but returned in only 6 hours by flying 50 mph faster. What was his rate of travel to City B? $r \times 7 = (r+50) \times 6$
	3. The Williams gave their son a 5-year loan at an adjustable rate. If the interest rate increases by 2% they would receive the same amount of interest over the first 4 years as they would receive over the entire 5 years at the current rate. What is the current rate? $5 \times r = 4 \times (r+.02)$
Interest	1. Jane invested \$4,500 and received \$810 in interest payments over a 2-year period. Assuming that the interest did not accumulate in her account, what was the rate of interest? $\$4{,}500 \times r \times 2 = \$810$
	2. A charitable trust invested part of their resources for 5 years at 11% interest. At the end of the first year they discovered they were allowed to earn only 10% on their investments. What rate must they receive for the remaining 4 years to average 10% over the 5 years? $1 \times .11 + 4 \times r = 5 \times .10$
	3. John is making organic fertilizer by dripping ground seaweed into a vat at a rate of 7 oz. per hour for 9 hours. He realizes 4 hours after he starts that he mistakenly set the drip rate to 9 oz. per hour. What should the new rate be in order to let the process continue for the full 9 hours? $9 \times 4 + r \times 5 = 7 \times 9$
Mixture	1. A chemist has 10 pints of a 30% alcohol solution. How much water should she add to make a 23% alcohol solution? $.30 \times 10 = .23 \times (10+p)$
	2. A nurse has 2 quarts of 3% boric acid. How much of a 10% solution of the acid must she add to have a 4% solution? $.03 \times 2 + .10 \times q = .04 \times (2+q)$
	3. A grocer wants to add almonds selling for \$2.10 a pound to 15 pounds of peanuts selling for \$1.65 a pound. How many pounds of almonds should he add to make a mixture that sells for \$1.83 a pound? $\$1.65 \times 15 + \$2.10 \times p = \$1.83 \times (15+p)$
Work	1. Tom can mow his lawn in 2 hours. How long will it take him to finish mowing his lawn if his son mowed ¼ of it? $(1/2) \times h + 1/4 = 1$
	2. Bill can paint a room in 3 hours and Fred can paint it in 5 hours. How long will it take them if they both work together for the same number of hours? $(1/3) \times h + (1/5) \times h = 1$
	3. Pam can ride to Mary's house in 3 hours and Mary can ride to Pam's house in 2 hours. How long will it take them to meet if they both leave their house at the same time and ride toward each other? $(1/3) \times h + (1/2) \times h = 1$

complex equation (as measured by the number of symbols) and Problems 2 and 3 have equally complex equations. However, as indicated in Experiment 1, the complexity of an equation is not always a good predictor of the judged complexity of a problem. Only for the interest category did subjects select Problem 1 as the least complex problem and evenly divide their choices for the most complex problem between Problems 2 and 3.

The proposed feature model successfully predicted the judged complexity of problems in Experiment 1. Table 8 shows a feature representation of the problems used in Experiment 2. The features listed for Problem 3 are the ones isomorphic to the features listed for Problem 2. The prediction that the number of features should determine judged complexity is partially supported by the data in Table 8. The major failure of the model is that almost all of the students selected Problem 3 as the most complex problem in the distance category, perhaps because it contains more words or because interest problems are more unfamiliar than distance problems.

The model does better for the other three categories. It successfully predicts for the interest category that Problem 1 will be judged least complex and Problems 2 and 3 will be judged equally complex. However, in order

Table 8
A Feature Representation of the Problems in Experiment 2

Category	Problem	Features						Complexity Least	Complexity Most
Distance		Distance	Time	Return Time	Speed Change				
	1	1,575	7					51	0
	2		7	6	+50			1	2
	3		5	4	+.02			0	50
Interest		Investment	Interest	$Time_1$	$Rate_1$	$Time_2$	$Rate_2$		
	1	$4,500	$810	2				40	4
	2			1	.11	5	.10	3	22
	3			4	9	9	7	8	26
Mixture		$Quantity_1$	$Concentration_1$	$Concentration_2$	$Concentration_3$				
	1	10	.30	(0)	.23			15	12
	2	2	.03	.10	.04			15	17
	3	15	$1.65	$2.10	$1.83			21	20
Work		$Rate_1$	$Rate_2$	Tasks Completed					
	1	2		¼				21	25
	2	3	5					16	11
	3	3	2					15	16

to predict that students judge problems in the mixture category as equally complex, it is necessary to assume that they infer a concentration for water (0% acid). The finding that students judge the three work problems as equally complex is predicted by the feature model.

Similarity judgments. Table 9 shows how many students selected each of two problems as the one more similar to a third. It also shows how many students selected each of two problems as the preferred solution for solving the third problem.

The influence of story context on the students' similarity judgments and preferences is shown by their selections for Problem 1. The students had to choose between the two isomorphic problems, Problems 2 and 3. We expected that the students would choose Problem 2 as more similar, because it had the same story context. The data supported our expectations for each of the problem categories.

The influence of mathematical structure on the students' similarity judgments is shown by their judgments for Problem 3. Both Problems 1 and 2 differ in story context from Problem 3, but Problem 2 has the same mathematical structure. Our expectation that students would select Problem 2 as the more similar was strongly supported for 3 of the 4 categories. The students were divided in their judgments for the mixture category, and apparently they did not distinguish between adding water and adding acid to a solution.

A comparison between the similarity ratings and solution preferences in Table 9 shows the same pattern as that obtained in Experiment 1. The high correlation ($r = 0.95$) between these two variables suggests the importance of similarity in determining solution preferences.

Selecting solutions for Problem 2. Table 10 shows how complexity, inclusiveness, and similarity influenced the selection of solutions for solving Problem 2. Complexity had a significant effect on solution preferences for only one of the four categories. The students selected the less complex problem for the distance problem. For inclusiveness, the students significantly preferred the isomorphic problem (the more inclusive of the two test problems) for only the work category. Their selection of the less inclu-

Table 9
Solution Preferences and Similarity Choices in Experiment 2

Selection	Distance P	Distance S	Interest P	Interest S	Mixture P	Mixture S	Work P	Work S
			Problem 1					
Problem 2	46	46	41	42	40	39	36	40
Problem 3	5	6	10	8	11	11	16	12
			Problem 2					
Problem 1	37	25	30	21	41	41	5	4
Problem 3	15	27	22	30	11	10	47	48
			Problem 3					
Problem 1	5	7	16	5	22	21	3	7
Problem 2	46	45	36	46	30	29	49	45

Note—The data show how many subjects selected each problem as providing the preferred solution (P) and how many subjects selected each problem as being more similar (S) to the specified problem.

Table 10
Effects of Complexity, Inclusiveness, and Similarity on Selecting Solutions in Experiment 2

	Complexity			Inclusiveness			Similarity		
Category	Less	More	z score	Less	More	z score	Less	More	z score
Distance	36	15	2.91*	37	15	2.91*	16	34	2.40*
Interest	26	25	0.00	30	22	.97	18	32	1.84
Mixture	27	24	.28	41	11	4.29*	7	43	4.94*
Work	22	29	.69	5	47	5.96*	7	43	4.94*

Note—The more inclusive problem refers to the isomorphic problem, which is the more inclusive of the two solutions. *Significant at the $p < .01$ level.

sive solutions in both the distance and the mixture categories was also significant.

The data supported the prediction that perceived similarity would influence students' selection preferences. Similarity produced significant results in three of the four problem categories and marginally significant results ($p < .05$ for a one-tailed test) for the fourth category.

As in Experiment 1, a preference score for each student was calculated to determine if the students' preferences had been determined by complexity, inclusiveness, or similarity. The mean score for complexity was 0.35, and the standard deviation was 2.29 [$t(49) = 1.09$, $p > .05$]. The mean score for inclusiveness was 0.35, and the standard deviation was 1.67 [$t(50) = 1.48$, $p > .05$]. The results show that neither complexity nor inclusiveness had a significant influence on the students' preferences. In contrast, the mean score for similarity was 2.04, and the standard deviation was 1.84 [$t(49) = 7.96$, $p < .01$].

Report of strategies. After making their selections, the students were asked to choose from among six strategies the one that best described how they had made their selections. Twenty-three subjects said they chose the more similar solution regardless of its complexity, 12 subjects said they chose the less complex solution because it would be easier to understand, 9 subjects said they chose the more complex solution because it would more likely contain the information needed to solve the problem, and 5 subjects said they chose the less complex solution because simpler solutions are usually taught before complex solutions. One of the subjects said he chose the more complex solution because it would be more challenging, and one subject reported using an alternative strategy that differed from the listed strategies. This distribution of responses is very similar to the distribution found in Experiment 1.

The subjects also received the strategy questionnaire after attempting to use solutions to solve the problems, in order to determine whether their strategies would change. Essentially the same distributions of responses occurred after the subjects had the opportunity to use some of the solutions.

Using solutions. An analysis of variance was performed to compare the relative effectiveness of the less inclusive and isomorphic solutions. This analysis yielded results that corresponded to those obtained in the first experiment. Providing a less inclusive solution resulted in an improvement in performance from 5% correct on Trial 1 to 12% correct on Trial 2. Providing an isomorphic solution resulted in an improved performance from 6% correct on Trial 1 to 51% correct on Trial 2. Students, therefore, solved more than four times as many problems when using isomorphic solutions than when using less inclusive solutions [$t(51) = 7.04$, $p < .01$].

As in Experiment 1, the effect of problem similarity on the successful use of solutions was evaluated for each of the four categories. Table 11 shows that the solution to the more similar problem was more effective than the solution to the less similar problem only for the interest and work categories. The large improvement for the work category occurred because the students were successful in identifying the isomorphic work problem as the more similar problem. As indicated in Table 12, the students were able to use the isomorphic solutions to significantly improve their performance on each of the four test problems.

Discussion

The results of Experiment 2 replicated in large measure the results of Experiment 1. Clearly the more inclusive

Table 11
Effect of Problem Similarity on the Successful Use of Solutions in Experiment 2

Category	Problem Similarity	Percent Correct Trial 1	Percent Correct Trial 2	z score
Distance	Less	7	28	2.12*
	More	17	39	2.23*
Interest	Less	0	4	1.00
	More	9	35	2.12*
Mixture	Less	0	12	1.73
	More	4	24	1.87
Work	Less	0	14	1.73
	More	7	87	4.64*

*Significant at the $p < .05$ level.

Table 12
Effect of Solution Inclusiveness on the Successful Use of Solutions in Experiment 2

Category	Solution Inclusiveness	Percent Correct Trial 1	Percent Correct Trial 2	z score
Distance	Less	10	14	0.58
	Isomorphic	13	57	3.16*
Interest	Less	0	0	0
	Isomorphic	7	31	2.33*
Mixture	Less	3	10	1.44
	Isomorphic	0	26	2.45*
Work	Less	4	22	2.00*
	Isomorphic	3	86	4.90*

*Significant at the $p < .05$ level.

(isomorphic) solutions were more effective than the less inclusive solutions were, as had been predicted. Transfer to analogous problems was improved with the use of isomorphic solutions, which resulted in the solution of four times as many problems as did less inclusive solutions. The isomorphic solutions were significantly effective in all four problem categories.

As hypothesized, the subjects made their selections on the basis of perceived similarity, even though selection on that basis often resulted in the selection of the less inclusive solution. Although isomorphic solutions were shown to be far more effective in providing useful information for solving algebra word problems, students with limited problem-solving experience did not recognize the potential usefulness of the isomorphic problems. This finding corresponds to the research by Chi et al. (1982), Schoenfeld and Herrmann (1982), and Silver (1979), which indicates that novices are often insensitive to the mathematical structure of problems.

The results of Experiment 2, therefore, support the hypothesis that students would not show a significant preference for the isomorphic solutions, even though these solutions would be more beneficial. The results are consistent with previous studies that have demonstrated a gap between students' ability to spontaneously notice the similarity of isomorphic problems and their ability to use an isomorphic solution when told of its value (Gick & Holyoak, 1980; Reed, 1987). However, our results go further by showing the consequences of this gap: Students select analogous problems that are not as helpful as they could be.

EXPERIMENT 3

The students' inability to select good analogous problems may have been caused either by inexperience or by their lack of opportunity to study the solutions before making their selections. Our purpose in Experiment 3 was to investigate whether these factors would influence students' selections.

The subjects in Experiments 1 and 2 were tested in college algebra classes and therefore had similar preparation in mathematics. In contrast, the subjects in Experiment 3 were participants in the psychology subject pool and therefore had a more varied background in college mathematics courses. The second factor—familiarity with the analogous solutions—was varied by allowing the subjects to study the solutions to half of the problem sets before they made their selections. We were therefore able to determine whether either experience or seeing the solutions would increase the selection of the more inclusive solutions.

The problems consisted of three of the four sets from Experiment 1 and three of the four sets from Experiment 2. The work problems were eliminated from each of these sets, because most of the subjects selected the more inclusive work problem in both experiments. The remaining three sets resulted in a 41% solution rate for the more inclusive solutions in Experiment 1 and a 17% solution rate for the less inclusive solutions. The three sets from Experiment 2 resulted in a 38% solution rate for the more inclusive (isomorphic) solutions and an 8% solution rate for the less inclusive solutions.

Method

Subjects. The subjects were 85 undergraduates in the psychology subject pool at Florida Atlantic University. Eight subjects had not taken a college algebra course, 57 subjects had either taken or were currently enrolled in a college algebra course, and 20 subjects had taken or were currently enrolled in a calculus course. They received course credit for their participation.

Procedure. The instructions indicated that the purpose of the experiment was to determine how people select related problems to help them solve problems. The students were told that they would see the solutions to some of the problems before making their judgments. They were also informed that they would be spending 3 min on each page and that they should not move forward or backward in the booklet if they finished early.

The format of the questions was identical to the format used in Experiments 1 and 2. The three problems in a set appeared at the top of a page. The questions below the problems asked the subjects to select solutions, rate the complexity of the problems, and judge the similarity of the problems.

The three similar sets (from Experiment 1) and three isomorphic sets (from Experiment 2) appeared on alternate pages, starting with a similar set for approximately half the subjects and an isomorphic set for the remainder. The subjects were randomly assigned to one of two groups, distinguished by whether they received solutions to the similar sets or the isomorphic sets. The solutions consisted of the solution to the least inclusive and most inclusive problem for each of the similar sets and the solution to the least inclusive and isomorphic problem for each of the isomorphic sets. The solutions were the same solutions used during the problem-solving phase of Experiments 1 and 2 (see the Appendix). The subjects had 3 min to study the two solutions immediately before answering the questions about a problem set. If seeing the solutions is helpful, students should have been more likely to select the more inclusive solution when shown solutions for the similar sets and more likely to select the isomorphic solution when shown solutions for the isomorphic sets.

Results

We analyzed selections for Problem 2 that required choosing between a less inclusive solution and a more inclusive (or isomorphic) solution. We will first present the results showing how perceived complexity, inclusiveness, and perceived similarity influenced subjects' selections. We will then present the results showing how mathematical experience and the opportunity to see solutions influenced the selections.

Effect of complexity, inclusiveness, and similarity. Table 13 shows how perceived complexity, inclusiveness, and perceived similarity affected the selection of solutions. The results supported the previous findings that only perceived similarity consistently influenced how students made their choices.

For both the three similar problems and the three isomorphic problems, the subjects had a significant preference for the more complex solutions for one problem, the less complex solution for another problem, and no significant preference for the third problem. Inclusiveness also had an inconsistent influence across the six problems. The subjects showed a significant preference for the more in-

Table 13
Effects of Complexity, Inclusiveness, and Similarity on Selecting Solutions in Experiment 3

	Complexity			Inclusiveness			Similarity		
Category	Less	More	z score	Less	More	z score	Less	More	z score
				Similar					
Cost	35	48	1.32	35	49	1.42	22	61	4.17*
Distance	20	57	5.70*	21	64	4.56*	19	62	4.67*
Fulcrum	61	23	4.03*	61	23	4.03*	10	73	6.81*
				Isomorphic					
Distance	64	20	4.69*	64	21	4.56*	27	57	3.16*
Interest	23	61	4.03*	22	63	4.34*	16	67	5.49*
Mixture	47	38	0.87	68	17	5.42*	7	75	7.40*

*Significant at the $p < .01$ level.

clusive solution for two problems and the less inclusive solution for three problems, and no significant preference for one problem. In contrast, the subjects consistently preferred the more similar problem across all six problems.

Effect of mathematical experience and seeing solutions. We analyzed subjects' selections in a 3 (experience) × 2 (solutions) analysis of variance to determine whether either of these variables would influence the selection of the more inclusive solution. The selections for the three similar problems and the three isomorphic problems were separately analyzed.

The analysis for the similar sets revealed that neither experience [$F(2,79) < 1$] nor solutions [$F(1,79) < 1$] influenced subjects' preferences. The interaction was also nonsignificant [$F(2,79) < 1$, $MS_e = 0.70$] for all tests. The more inclusive solution was selected on 54% of the occasions for subjects who had not taken college algebra, 54% of the occasions for subjects who had taken college algebra, and 55% of the occasions for subjects who had taken calculus. The subjects who studied the similar solutions selected the more inclusive solution on 56% of their selections, compared with 51% for subjects who studied solutions for the isomorphic sets.

In contrast, seeing the solutions for the isomorphic sets significantly influenced the selection of the isomorphic problems [$F(1,79) = 4.41$, $MS_e = 0.54$, $p < .05$]. The subjects who studied the solutions to the isomorphic sets selected the isomorphic problem on 43% of the occasions, compared with 35% for the subjects who studied solutions to the similar sets. Neither experience [$F(2,79) = 2.84$] nor the experience × solutions interaction [$F(2,79) = 1.17$] was significant. The subjects who had taken a calculus course selected the isomorphic solutions on 42% of their selections, compared with 36% for students who had taken a college algebra course, and 50% for students who had not taken a college algebra course. The surprisingly high value of the latter group may have been caused by the small sample size, since there were only 8 subjects in this group.

The finding that mathematical experience did not have a significant influence on selections deviates from previous findings that expertise helps people identify isomorphic problems (Chi et al., 1982; Schoenfeld & Herrmann, 1982). However, the range of expertise was greater in the Chi et al. study, in which the novices were undergraduates and the experts were advanced students in a PhD program. In the Schoenfeld study, a within-subjects comparison was made before and after students took an intensive course on mathematical problem solving. Our results showed that showing students solutions significantly increased the selection of an isomorphic analogue, although the increase was not large.

EXPERIMENT 4

The failure to find an effect of mathematical experience on selecting solutions in Experiment 3 may have been caused by an insufficient range in experience. In Experiment 4, we included a group of undergraduates who were majoring in mathematics and planned to teach mathematics at a junior high or high school. They were all enrolled in an upper-division mathematics course, Basic Mathematic Concepts, and they had previously taken an average of six mathematics courses.

Method

Subjects. The subjects were 76 undergraduates at San Diego State University, including the 28 students who were majoring in mathematics. The remaining 48 students were currently taking either an introductory or a cognitive psychology course and received course credit for participating. This group included 29 students who had not taken any college algebra (or more advanced courses) and 19 students who had taken a college algebra course. Three of the 19 students had also taken a calculus course. All the subjects were tested in groups.

Procedure. The procedure was identical to the procedure in Experiment 3, except that the subjects did not receive solutions to any of the problems and the three similar and three isomorphic problems were blocked rather than alternated. Approximately half the subjects at each level of experience received the similar problems first and the remainder received the isomorphic problems first. This allowed us to evaluate whether presentation order would influence selections.

Results

Effect of complexity, inclusiveness, and similarity. Table 14 shows how perceived complexity, inclusiveness, and perceived similarity influenced the selection of solutions. Once again, only perceived similarity strongly influenced the selections.

Complexity and inclusiveness significantly influenced selections for only two of the six problems, and the bias

Table 14
Effects of Complexity, Inclusiveness, and Similarity on Selecting Solutions in Experiment 4

	Complexity			Inclusiveness			Similarity		
Category	Less	More	z score	Less	More	z score	Less	More	z score
				Similar					
Cost	39	36	0.23	37	39	0.34	15	60	5.08*
Distance	36	38	0.12	38	38	0.11	26	50	2.41*
Fulcrum	63	12	5.77*	60	15	5.08*	14	61	5.31*
				Isomorphic					
Distance	48	28	2.18*	47	29	1.95	21	55	3.78*
Interest	33	42	0.92	32	43	1.15	17	58	4.62*
Mixture	40	36	0.80	59	17	4.70*	13	63	5.62*

*Significant at the $p < .05$ level.

in these cases was toward the less complex and less inclusive problems. In contrast, the students showed a significant preference for the more similar problem for all six problems, replicating the results of Experiment 3.

Effect of mathematical experience and presentation order. We analyzed the subjects' selections in a 3 (experience) × 2 (presentation order) analysis of variance to determine whether either of these variables would influence the selection of the more inclusive solution. The selections for the three similar problems and the three isomorphic problems were separately analyzed.

The analysis of the similar sets revealed that neither experience [$F(2,70) < 1$] nor presentation order [$F(1,70) < 1$] influenced the subjects' preferences. The interaction was also nonsignificant [$F(2,70) = 2.13$, $MS_e = 0.69$] for all tests. The more inclusive solution was selected on 38% of the occasions for the students who had not taken college algebra, 39% for the students who had taken college algebra, and 42% for the students who were mathematics majors. The subjects who received the similar problems first selected the more inclusive solution on 39% of their choices, compared with 40% for those students who received the isomorphic problems first.

In contrast, mathematical experience did have a significant effect on the selection of isomorphic problems [$F(2,70) = 4.37$, $MS_e = 0.59$, $p < .02$]. The isomorphic solution was selected on 37% of the occasions for the students who had not taken college algebra, 28% of the occasions for the students who had taken college algebra, and 50% of the occasions for the students majoring in mathematics. Neither presentation order [$F(1,70) = 2.15$] nor its interaction with experience [$F(2,70) < 1$] was significant. The subjects selected the isomorphic problems on 44% of their choices when the isomorphic problems occurred first and on 35% of their choices when the similar problems occurred first.

GENERAL DISCUSSION

Our objective was to identify variables that influence the selection of analogous solutions and to determine whether students would select effective solutions. In Experiment 1, students had to choose between two problems that belonged to the same category as the test problem. One problem was less inclusive than the test problem and the other problem was more inclusive than the test problem. In Experiment 2, students had to choose between a problem that was less inclusive than the test problem and a problem that was isomorphic to the test problem.

The same pattern of results occurred in both experiments: Students selected problems on the basis of perceived similarity. They did not show a significant preference for the more inclusive problems in Experiment 1 or the isomorphic problems in Experiment 2, although both sets of solutions were significantly more effective than were solutions to the less inclusive problems. The results therefore reveal a discrepancy between the variable that determines the selection of solutions (similarity) and the variable that determines the usefulness of solutions (inclusiveness). Furthermore, as has been shown in Experiments 3 and 4, neither mathematical experience nor showing students the solutions had much impact on the selection of more inclusive solutions, although both increased the selection of isomorphic solutions.

Holyoak and Koh (1987) have proposed that the retrieval of analogies is based on a summation of activation resulting from multiple shared features. Both structural features, which play a causal role in determining possible solutions, and salient surface features influence the selection in their model. A possible problem with this proposal is that a more inclusive problem shares more structural features with the target problem than a less inclusive problem does. According to the summed features view, students should therefore select the more inclusive problem (at least for the similar problem sets, in which there is a close correspondence between surface and structural features). Our results suggest that selecting problems on the basis of shared structural features is a better normative model than a descriptive model.

A practical issue related to this discrepancy is the question of how students can improve their ability to select appropriate analogies. Although the same pattern of results occurred in Experiment 1 and in Experiment 2, the answer may depend on whether the most effective analogy is a member of the same category, as in Experiment 1, or a member of a different category, as in Experiment 2. Selecting an effective analogy from the same category requires determining whether there is sufficient information in the analogous problem for solving the test problem. This requires comparing how the problems differ in the amounts of relevant information that they contain.

In contrast, students' inability to select isomorphic problems as analogous is caused by their inability to spontaneously map the features and relations in one problem onto the features and relations in the isomorphic problem. Such a mapping depends on the recognition of a common mathematical structure, which seems to require considerable experience (Chi et al., 1982; Schoenfeld & Herrmann, 1982).

The difficulty in noticing isomorphic problems may be enhanced when students have to retrieve analogous problems from long-term memory. Ross (1984) found that students are more likely to be reminded of an analogous problem when it has the same story content as the test problem has. This, of course, would reduce the probability of retrieving an isomorphic problem, which has a different story content. Gentner and Landers (1985) also found that students were reminded of stories that had the same content, although they could appropriately judge the soundness of the analogy when asked to rate pairs of stories that were simultaneously presented.

We suspect that teaching students to select isomorphic problems as a basis for analogical reasoning will be a challenging task, requiring the teaching of considerable domain-specific knowledge about the formal structure of problems. In contrast, teaching students to select more inclusive problems as a basis for analogical reasoning would seem to require less domain-specific knowledge, because students can more readily use surface information as a basis for determining inclusiveness. Glaser (1984) and Polson and Jeffries (1985) have recently raised the issue of how much domain-specific knowledge is required to teach general heuristics. For the heuristic of selecting an analogous problem, we expect that the answer will depend on whether the best analogy comes from the same category (Experiment 1) or a different category (Experiment 2). In either case, our results demonstrate that there is a need for students to make better selections. There is also a need for instruction on the use of analogous solutions, because students solved only one third of the problems in Experiment 1 and one half of the problems in Experiment 2 when given the better solutions.

REFERENCES

Barsalou, L. S. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 629-654.

Carroll, J. M., & Mack, R. L. (1985). Metaphor, computing systems, and active learning. *International Journal of Man-Machine Studies*, **22**, 39-57.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, NJ: Erlbaum.

Gentner, D. (1980). *The structure of analogical models in science* (BBN Rep. No. 4451). Cambridge, MA: Bolt Beranek and Newman Inc.

Gentner, D., & Landers, R. (1985, November). *Analogical reminding: A good match is hard to find*. Paper presented at the International Conference on Systems, Man & Cybernetics, Tucson, AZ.

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

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, **39**, 93-104.

Holyoak, K. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19, pp. 59-87). Orlando, FL: Academic Press.

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

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.

Polson, P. G., & Jeffries, R. (1985). Instruction in general problem-solving skills. In J. W. Segal & S. F. Chipman (Eds.), *Thinking and learning skills* (Vol. 1, pp. 417-455). Hillsdale, NJ: Erlbaum.

Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 124-139.

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

Reed, S. K., & Ettinger, M. (1987). Usefulness of tables for solving word problems. *Cognition & Instruction*, **4**, 43-59.

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

Schoenfeld, A. H. (1979). Explicit heuristic training as a variable in problem-solving performances. *Journal for Research in Mathematics Education*, **10**, 173-187.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **5**, 484-494.

Silver, E. A. (1979). Student perceptions of relatedness among mathematical problem solvers. *Journal for Research in Mathematics Education*, **10**, 195-210.

APPENDIX

Less Inclusive Solution

A pilot flew 1,575 miles in 7 hours. What was his rate of travel?

The problem is a distance-rate-time problem in which

$$\text{distance} = \text{rate} \times \text{time}$$

We begin by constructing a table to represent the rate, time, and distance. We want to find the rate of travel. Let r represent the number we want to find and enter it into the table below. It took 7 hours to travel the 1,575 miles. The table below shows these values.

Rate (mph)	Time (hours)	Distance (miles)
r	7	1,575

The following equation allows us to solve for r:

$$1{,}575 = r \times 7$$

More Inclusive Solution

A pilot flew his plane from Milton to Brownsville in 5 hours with a 25 mph tailwind. The return trip, against the same wind, took 1 hour longer. What was the rate of travel without any wind?

The problem is a distance-rate-time problem in which

$$\text{distance} = \text{rate} \times \text{time}$$

We begin by constructing a table to represent the rate, time, and distance for each leg of the trip. Let r represent the rate of travel without any wind. The rate of travel with the wind of the initial trip was $r+25$. The rate against the wind on the return trip was $r-25$. It took 5 hours for the initial trip and 6 hours for the return trip. We can now represent the distance between the two cities by multiplying the rate and time for each leg of the trip. The table below shows these values.

	Rate (mph)	Time (hours)	Distance (miles)
Initial Trip	$r+25$	5	$(r+25)\times 5$
Return Trip	$r-25$	6	$(r-25)\times 6$

Because the distance of the initial trip is the same as the distance of the return trip, we set the two distances equal to each other. The following equation allows us to solve for r:

$$(r+25)\times 5 = (r-25)\times 6.$$

Isomorphic Solution

The Williams gave their son a 5-year loan at an adjustable rate. If the interest rate increases by 2% they would receive the same amount of interest over the first 4 years as they would receive over the entire 5 years at the current rate. What is the current rate?

The percentage of the loan that is owed in interest is equal to the interest rate multiplied by the length of the loan. Let r equal the current interest rate. Because r represents the percentage of the loan owed in one year, the amount of interest owed after 5 years is $5\times r$. The first line of the table shows this information.

Interest Rate	Length of Loan (years)	Interest (% of Loan)
r	5	$5\times r$
$r+.02$	4	$4\times(r+.02)$

The amount of interest owed after 4 years at the higher interest rate is $4\times(r+.02)$, as shown in the bottom line. Setting these two amounts equal to each other yields the equation:

$$5\times r = 4\times(r+.02).$$

(Manuscript received August 10, 1987;
revision accepted for publication June 6, 1989.)