Journal of Experimental Psychology:
Learning, Memory, and Cognition
1994, Vol. 20, No. 2, 366–378

0278-7393/94/$3.00

Effects of Solving Related Proofs on Memory and Transfer in Geometry Problem Solving

Marsha C. Lovett and John R. Anderson

Three experiments investigate the relationship between memory and problem solving in the domain of geometry theorem proving. In Experiment 1, Ss memories for an original problem-solving episode were interfered with retroactively by solving a 2nd problem that had the same diagram, but no memory effects were observed that depended on the second problem's logical similarity to the original. Results suggest that the diagram is the basis for geometry problem-solving memories. Experiments 2 and 3 investigated problem-solving memories in use by examining Ss transfer to a 3rd (test) problem. As with the memory results, transfer was reduced when the 1st two problems had the same diagram relative to when they had 2 different diagrams. Transfer was reduced most in the condition with the greatest proportion of memory-interfering steps. Results suggest that the structure and quality of problem-solving memories affect problem-solving transfer.

As problem solvers gain experience in a particular domain, they not only develop their skills but also form problem-solving memories. These memories can play a role in subsequent problem solving. For example, while a person plays a game of chess, he or she is creating a memory trace that might include specific moves and board patterns from that game. Research suggests that the use of such memories can be helpful in subsequent play (de-Groot, 1965; Holding, 1985). We argue that the same is true of problem solving in geometry: Problem-solving memories can have a significant impact on transfer. First, we explore the structure and dynamics of geometry problem-solving memories by demonstrating that they can be retroactively affected by certain types of intervening problem-solving episodes (Experiment 1). Then, we investigate how these changed problem-solving memories can, in turn, affect subsequent problem-solving performance (Experiments 2 and 3).

The term *problem-solving memory* refers to an episodic memory trace that includes various features of a previous problem-solving experience (e.g., the problem statement, steps taken in the solution, operators used, and mistakes made). Several studies have examined subjects' problem-solving memories. For example, Egan and Greeno (1974) found that, after solving the Towers of Hanoi task in which a set of disks must be moved from one peg to another, subjects remembered problem-solving steps associated with top-level goals better than they remembered steps associated with lower level goals. This effect was maintained when subjects were taught a goal-based strategy but not when they were taught a perceptual strategy (Ruiz, 1987). Furthermore, solving the same Towers of Hanoi problem multiple times increased subjects' overall recognition memory for states and for state–operator pairs that occurred during problem solving (Ruiz, 1987). In the domain of mathematical word problems, Silver (1981) found that seventh-grade students exhibited differences in certain aspects of their memory for a given problem according to their performance level: Good solvers recalled more of the structural aspects of the problem they had solved than did poor solvers, but both groups recalled the same number of details. These results show that problem-solving memories can be influenced by the process of problem solving, but they do not reveal much about how or why such problem-solving memory differences occur.

In Experiment 1, we address these issues by examining a case in which the problems are multiple and distinct. More specifically, we focus on these questions: How are subjects' memories for an original geometry problem affected by solving a second, different geometry problem? What do these retroactive effects tell us about the structure of problem-solving memories?

Our theoretical framework for such retroactive influences comes not from the problem-solving literature but rather from research on memory for prose (Anderson & Bower, 1973; Bower, 1974; Bower, Black, & Turner, 1979; Crouse, 1971; Thorndyke & Hayes-Roth, 1979; Watkins & Kerkar, 1985). In general, this research has demonstrated that greater retroactive facilitation and interference occur in situations where the to-be-remembered structures are related: Facilitation is exhibited for those features shared between the structures, and interference is exhibited for those features not shared. For example, subjects who studied three stories of the same type (all short biographies) exhibited better memory for facts from the original biography that were shared by all three than did control subjects who read the original biography followed by two stories of different types. On the other hand, subjects who studied all three biographies exhibited poorer memory than did control subjects on facts unique to the original biography (Bower, 1974).

Marsha C. Lovett and John R. Anderson, Department of Psychology, Carnegie Mellon University.

Preparation of this article was supported in part by a National Science Foundation Graduate Research Fellowship to Marsha C. Lovett and by National Science Foundation Grant 91-08529 to John R. Anderson.

We thank Christopher Genovese, Akira Miyake, and our reviewers for their comments, which greatly improved the quality of this article.

Correspondence concerning this article should be addressed to Marsha C. Lovett, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890. Electronic mail may be sent to ml3n@andrew.cmu.edu.

Thorndyke and Hayes-Roth (1979) have proposed a memory schemata model that accounts for such facilitation and interference results in memory for prose. The fundamental tenet of this model is that memory traces for similar events (e.g., reading biographies of famous people) are grouped under a single schema. A schema consists of a set of interlinked components (e.g., OCCUPATION and RESIDENCE for a biography schema) that correspond to substructures common across the events. One thinks of the details of a particular memory trace (e.g., Thomas Jefferson was president of the United States; he lived in Monticello) as being attached to the appropriate components of the schema. According to this model, interference effects occur because different details from distinct events are attached to the same component of the same schema (e.g., Monticello and Mount Vernon would both be attached to RESIDENCE), complicating the retrieval of a desired piece of information (e.g., Where did Jefferson reside?) For interference to occur, it is first necessary that the schema itself be built up, a process that presumably takes place during early exposure to the similar events. Facilitation effects, on the other hand, occur when the same detail is repeatedly attached to a single component of the schema. These effects are primarily driven by an increasing likelihood of remembering the repeated detail during formation of the schema.

Thorndyke and Hayes-Roth (1979) sought empirical support for their model by asking subjects to study a variable number of training texts of the same type (e.g., biographies) and then to study and recall a new test passage, also of the same type. The test passage included sentences that were repeated identically throughout the training passages (shared information), sentences that had their details changed across the training passages (not-shared information), and sentences that were unrelated to any sentences during training (unrelated). Predictions of the schema model matched the memory results in several respects: (a) Recall of shared information increased with the number of training passages, (b) recall of not-shared information decreased with the number of training passages (after an initial rise), and (c) recall of unrelated information did not change with number of training passages.

These results and those by Bower (1974) are consistent with schema theory because memory effects were elicited only when subjects experienced multiple events of the same type (viz., read stories of the same type) and because memory facilitation and interference were specific to the shared and not-shared information, respectively. Nevertheless, the two studies measured these memory effects very differently: Bower measured the retroactive effects of interpolated stories on subjects' recall of the original story, whereas Thorndyke and Hayes-Roth (1979) measured the proactive effect of interpolated prose passages on subsequent memory performance.

In the three experiments reported here, we apply both variations of this paradigm to geometry theorem proving to investigate the structure and dynamics of geometry problem-solving memories. In Experiment 1, we measure the retroactive effects of solving a varying second geometry problem on an original problem-solving memory and thereby attempt to distinguish the structure of geometry problem-solving memories. In Experiments 2 and 3, we study problem-solving memories in action by measuring the impact of solving these same geometry problems on subsequent problem-solving performance.

Experiment 1

To determine the basis for geometry problem-solving memories, we created a problem-solving situation in which retroactive memory effects could occur. Subjects solved an original geometry proof problem followed by a second problem and then were asked to recall the original. According to the theoretical framework described earlier, retroactive effects should only occur when the two problems use the same memory structure—when their memories are associated with the same schema. Recall that, in the case of memory for prose, story type (e.g., biographies) provided the structure by which subjects' memories were facilitated and impaired (Bower, 1974; Thorndyke & Hayes-Roth, 1979). Geometry problem solving, in contrast, has two factors that could potentially provide the framework on which problem-solving memories are based: the diagram and the logic of the problem's solution. Although it is possible that both factors could be represented in problem-solving memories, we were interested in distinguishing which one might have the greater influence. Therefore, we formulated two competing hypotheses for the structure of geometry problem-solving memories. The diagram hypothesis states that the problem diagram forms the basis of geometry problem-solving memories and, hence, predicts that memory facilitation (for shared features) and interference (for not-shared features) should occur when the diagrams of the two geometry problems are similar. The logic hypothesis, on the other hand, states that the problem's logic is the structure of these memories and thus predicts that memory facilitation (for shared features) and interference (for not-shared features) will occur when the logical structures of the two proofs are similar.

Although it may seem counterintuitive to consider the diagram as structurally important to geometry problem-solving memories, there are several arguments that support this idea. For instance, in many geometry textbooks, the diagram is used to record the steps in a proof, linking important features of the problem's solution to the diagram. (All of our subjects made such marks on their diagrams to represent proof steps.) There is also considerable information contained in a geometry diagram that may help direct the course of problem solving. Geometry experts, for example, use a variety of proof schemata that rely heavily on certain perceptual configurations in the diagram (Koedinger & Anderson, 1990). Similarly, diagrammatic representations that ease subjects' memory and attention load and that place related pieces of information close together facilitate problem solving in physics (Larkin & Simon, 1987). By establishing diagram and logic as two reasonable candidates for the structure of geometry problem-solving memories, we created a situation in which retroactive memory effects could be a powerful discriminator.

Method

Subjects. Subjects were 48 Carnegie Mellon University undergraduates, participating either for credit or for payment. Two additional

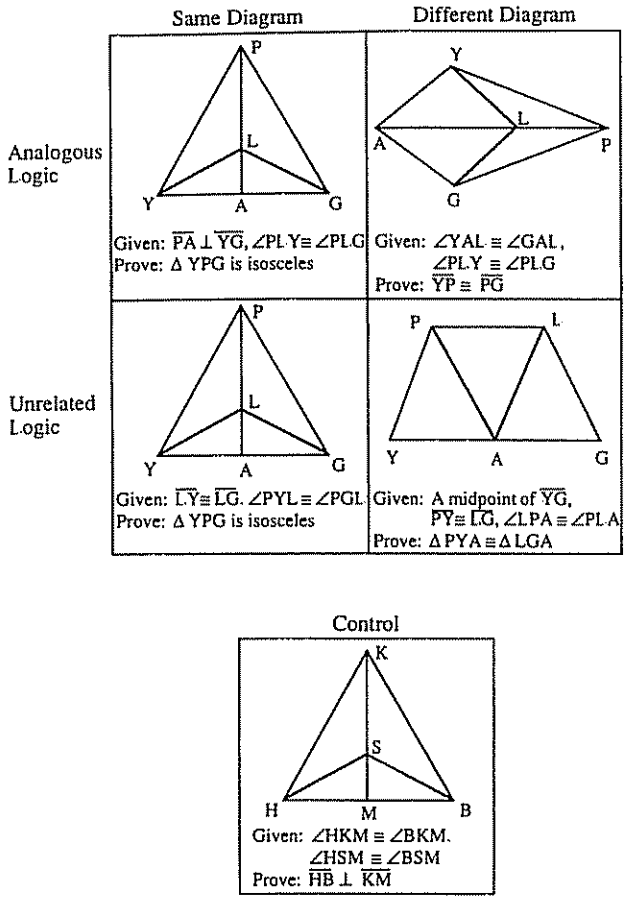


Figure 1 Geometry problems solved by subjects in each condition. The control problem was the only problem solved by subjects in the control condition and the original problem solved by all other subjects. Each problem in the 2 × 2 matrix corresponds to the second problem solved by subjects in that condition

subjects were removed from the study because they were unable to solve the problems.

Design We used a 2 × 2 design of Diagrammatic Similarity (same or different) × Logical Similarity (analogous or unrelated), with each factor representing the second problem's relationship to the original. A second problem's proof was considered analogous to the original problem's as long as it used the same logical structure (e.g., the givens lead to proving one pair of triangles congruent, that congruence leads to proving another pair of triangles congruent, and that congruence proves the goal). Sample solutions for the original problem and for the same diagram, analogous logic problem are presented in Appendix A to demonstrate this correspondence. A second problem was considered to have unrelated logic when it shared none of these structural components with the original problem

Sixteen subjects were randomly assigned to each level of the logic factor and participated in two sessions, one for each level of the diagram factor. Thus, the logic factor was between subjects, and the diagram factor was within subjects, with ordering of the two diagram levels counterbalanced. (In Experiments 2 and 3, the design was changed slightly, making diagram and logic both between-subjects factors.) An additional 16 subjects participated in a control condition that provided a baseline measure of memory for the original problem when no second problem was solved.

Materials Figure 1 shows the original problem (labeled *control*) and a 2 × 2 matrix of second problems (according to their relationship to the original). A second set of problems with the same relationships (see Appendix B) was required by the inclusion of the second session. Matching of problem set and session number was counterbalanced.

Procedure The experiment was conducted in two sessions. Each session began with a geometry review in which subjects read a summary of the theorems and definitions to be used in the experiment. This page was in subjects' view during all problem solving. Next, subjects solved the geometry problem or problems corresponding to their condition and taken from one of the two available problem sets. Finally, subjects were given a series of memory tests regarding the original problem. In the second session, subjects followed the same procedure except that the problems were taken from the other set and at the opposite diagrammatic level. Time between the two sessions averaged 2 days.

For each problem, subjects were instructed to write a two-column proof. The left column was for statements, and the right column was for justifications of each statement. After the original problem's proof was successfully completed and the solution sheet removed, subjects (in the experimental conditions) solved a second problem. The experimenter was present during all problem solving to answer questions and to provide hints and corrections. Such information was given in the most uniform manner possible: Only the next correct step in the optimal solution path was given. Hints were given when subjects did not write down a step or make any marks on the diagram for more than 30 s. On average, subjects received less than one hint per session.

To equate the time between beginning the original problem and beginning the memory tests, subjects worked on an unrelated puzzle after finishing the geometry problems and until a total of 35 min had elapsed since beginning their first proof. (Pilot testing revealed that 35 min was the maximum time necessary to complete the two proofs.)

The memory test given to all subjects included the recall of four types of information about the original problem: its theorems, diagram with labels, given statements, and steps. First, subjects were asked to write down the theorems and definitions they remembered using in their original proof. Second, on a separate sheet, subjects were asked to draw the original problem's diagram (including the letters labeling it). Subjects were asked to add letters next to the points on the diagram for labels they could not remember. Next, subjects were asked to write down the original problem's "givens" (the statements given to be true) and the problem's goal. Those who misremembered these statements were corrected so that, at this point, all subjects had their own copy of the original problem before them. Finally, subjects were asked to say out loud all the steps they could remember from their solution of the original problem. Two points were emphasized to the subjects: (a) to avoid re-solving the problem by saying only what they remembered and (b) to state their steps according to the particular solution they had used when working on the original problem. These protocols were recorded on audiotape

At the end of each session, we asked subjects to rate the problems' similarity on a scale from *very dissimilar* (1) to *very similar* (7) to verify that subjects were sensitive to both the diagrammatic and logical similarity relationships. Subjects rated their problems as more similar in the same diagram than in the different diagram conditions, $F(1, 30) = 25.06$, $MS_e = 1.18$, $p < .01$. Similarly, subjects in the analogous logic conditions rated their problems as more similar than did subjects in the unrelated logic conditions, $F(1, 30) = 25.01$, $MS_e = 1.08$, $p < .01$. The lack of a Diagram × Logic interaction suggested that these factors contributed to subjects' similarity ratings in an additive fashion, $F(1, 30) = 1.20$, $MS_e = 1.18$, $p > .10$. Therefore, subjects seemed to be highly sensitive to both similarity manipulations.

Results and Discussion

Recall that the diagram hypothesis predicted that memory interference and facilitation would occur when subjects solved problems with the same diagram. In contrast, the logic hypothesis predicted these effects would occur when subjects solved problems with analogous logic. By using our memory data, we were able to distinguish between the two hypotheses according to this difference. Across many of the memory tests, we found significant differences in subjects' memories for the original problem based on the diagrammatic similarity of their second problem. No main effects of logical similarity were revealed, however. These findings, summarized in Table 1, are consistent with the diagram hypothesis but not the logic hypothesis. The following sections provide details on specific results.

Memory for steps Subjects recalled a smaller percentage of steps from the original proof when their two problems had the same diagram as opposed to different diagrams (see Table 2). In contrast, there was no effect of the logical similarity between subjects' two problems. Furthermore, subjects' memories were worse in the same diagram conditions than in the control conditions, $t(40) = 2.06$, $p < .05$. These results support the diagram hypothesis and suggest that retroactive interference may be the mechanism at work.

The hypotheses under consideration also made specific predictions with respect to interference versus facilitation; both predicted that interference would apply to memory for items that were not shared between the two problems and that facilitation would apply to memory for items that were shared.

Table 1
Summary of Analyses on Memory Data in Experiment 1

Memory test	Diagram	Logic	Interaction
Steps			
$F(1, 30)$	5.84*	< 1	< 1
MS_e	348	567	348
Not-shared			
$F(1, 30)$	8.79**	1.20	1.11
MS_e	559	569	559
Shared			
$t(15)$	< 1[a]	—	—
Given statements			
$F(1, 30)$	7.66**	3.50	13.1**
MS_e	862	1350	862
Not-shared			
$t(15)$	4.39[b]**	—	—
Shared			
$t(15)$	< 1[a]	—	—
Diagram labels			
$F(1, 30)$	3.95*	< 1	< 1
MS_e	1060	1530	1060
Theorems			
$F(1, 30)$	1.31	< 1	2.03
MS_e	284	396	284
Not-shared			
$t(15)$	1.64[b]	—	—
Shared			
$t(15)$	< 1[a]	—	—

Note. Each row displays the results of either a 2 × 2 analysis of variance (Diagram × Logic) or a paired t test on subjects' percentage recall data. Dashes indicate data insufficient for analysis
[a]Analogous logic cells only [b]Unrelated logic cells only
*$p < .05$ **$p < .01$

Table 2
Percentage Recall on Various Memory Tests in Experiment 1

Test	Diagram: Same	Diagram: Different
Steps		
Analogous logic	67	78
Unrelated logic	65	80
Control	80	—
Givens		
Analogous logic	91	84
Unrelated logic	47	94
Control	91	—
Diagram labels		
Analogous logic	34	45
Unrelated logic	32	54
Control	71	—
Theorems		
Analogous logic	66	65
Unrelated logic	61	71
Control	73	—

Note. Dashes indicate that diagrammatic and logical similarity do not apply to the control condition.

Therefore, we categorized each step from each subject's original proof as shared if it had an isomorphic step corresponding to it in the subject's second proof, and not-shared if it had no such corresponding step in the second proof. (Isomorphic steps are defined as those with the same statement and justification, but they are not required to occur in the same position in their respective proofs.) Table 3 presents the mean percentages of not-shared steps and shared steps recalled by subjects in each condition. Subjects exhibited greater interference in their memory of not-shared steps when their two problems had the same diagram compared with different diagrams, exactly as the diagram hypothesis predicted.

It is important to note that subjects in the same diagram, unrelated logic condition were at a greater disadvantage than subjects in the same diagram, analogous logic condition. Not only were they subject to the same interference effects on not-shared steps (e.g., both conditions exhibited similarly low recall percentages), but subjects in the same diagram, unrelated logic condition had an absolute number of not-shared steps that was twice as large as that of subjects in the same diagram, analogous logic condition. Specifically, out of the grand mean of 11.2 total steps taken to solve the original problem, subjects in the same diagram, unrelated logic condition had on average 10.4 not-shared steps, whereas subjects in the same diagram, analogous logic condition had only 4.7 not-shared steps. According to the diagram hypothesis, this greater number of not-shared steps should entail a greater absolute quantity of memory interference (e.g., more interfered-with steps) even though the percentage measures of interference were approximately the same across the two same-diagram conditions.

Analyses of memory for shared steps were not as meaningful because of the small numbers of shared steps in the unrelated logic cells. A t test comparing the two analogous logic cells did not show facilitation according to diagrammatic similarity.

Memory for givens As shown in Table 2, percentage recall for givens was quite good for all subjects except those in the

Table 3
Percentage Recall for Not-Shared and Shared Steps and Theorems in Experiment 1

Test	Diagram: Same	Diagram: Different
Not-shared steps		
Analogous logic	56	80
Unrelated logic	65	80
Shared steps		
Analogous logic	78	77
Unrelated logic	—	—
Not-shared theorems		
Analogous logic	—	—
Unrelated logic	60	72
Shared theorems		
Analogous logic	70	75
Unrelated logic	—	—

Note. Dashes indicate that small absolute numbers made percentages uninterpretable.

same diagram, unrelated logic condition. (Note that the original, to-be-remembered problem always had two given statements, making an individual subject's recall for givens necessarily 0%, 50%, or 100%.) Here, overall recall measures are not very meaningful because of the nature of the givens in the analogous versus unrelated logic problems—shared givens only occurred in the analogous logic conditions, and not-shared givens only occurred in the unrelated logic conditions. (This situation was unavoidable because making a second problem's logical structure analogous to the original problem's required that the two problems' givens be shared.) Because both hypotheses under consideration make predictions on the basis of whether items are shared, separate analyses for shared and not-shared givens were appropriate. The particular breakdown of shared and not-shared givens permitted one comparison only within each level of the logic factor. In the case of shared givens, there was no facilitation in subjects' memories according to diagrammatic similarity, $t(15) < 1$. In the case of not-shared givens, however, subjects did exhibit interference in the same diagram cell, as predicted by the diagram hypothesis, $t(15) = 4.39, p < .01$.

Memory for the diagram. The diagram itself was remembered perfectly by all but 2 of the 48 subjects. One subject missed drawing one line segment from the original diagram, and another drew the diagram from the second problem instead of the original. Such striking memory for the diagram indirectly supports the diagram hypothesis because it indicates that subjects could easily store and retrieve diagrammatic configurations.

Subjects' memory specific to the diagram's labels, however, exhibited severe diagram-based interference effects as is shown in Table 2. Subjects solving a second problem, which always had different labels from the original, recalled a much lower percentage of the original problem's labels than did the control subjects, $F(1, 78) = 8.56, p < .01$. Within the four experimental conditions, subjects' memory for labels exhibited greater interference in the same diagram conditions, again providing support for the diagram hypothesis. No such difference was demonstrated with respect to logical similarity. These results are quite intuitive because it seems natural that subjects' memory for the labels would be tied to the diagram.

Memory for theorems. There were no significant differences among the four experimental conditions with respect to subjects' percentages of theorems recalled. Separate analyses of memory for not-shared and shared theorems tested the specific interference and facilitation predictions of the diagram and logic hypotheses. Because of small absolute numbers of shared and not-shared theorems in certain conditions, these percentage data were interpretable in only half of the cells. Subjects' recall of not-shared theorems (measured in the unrelated logic conditions only) revealed greater interference when the original and second problems had the same diagram (see Table 3), but this effect did not reach statistical significance, $t(15) = 1.64$; $p < .12$. Subjects' recall of shared theorems (measured in the analogous logic conditions) revealed no reliable effect of diagrammatic similarity, $t(15) < 1$.

Summary of results. In general, the results of Experiment 1 supported the hypothesis that a geometry problem's diagram is the underlying structure for problem-solving memories. After solving two proofs with the same diagram, subjects exhibited significant interference in their memory for the original problem's steps, given statements, and diagram labels. Indeed, in the case of memory for steps, interference was specific to the not-shared steps, exactly as predicted by the diagram hypothesis. In contrast, none of the memory measures indicated a logic-based effect.

By applying the memory schemata model proposed by Thorndyke and Hayes-Roth (1979) to the domain of geometry problem solving, we can better understand these results. In the case of geometry, a memory schema corresponds to a problem diagram (see Figure 2). Its components are specific relations between pairs of objects in the diagram (e.g., $\angle HKM = \angle BKM$). The details of a particular geometry problem-solving memory that are attached to the appropriate components are indications of whether that relation was established in the course of the proof (e.g., Yes or No). The key implication of this model is that geometry problem-solving memories that rely on the same diagram will be associated with the same memory schema. Furthermore, in such a schema, shared and not-shared aspects of those memories will be represented differently: A shared step, for example, is represented by a component with Yes attached twice and a not-shared step by a component with a Yes and a No. It is possible to consider more informative details (e.g., each Yes connected to a node indicating the justification used, etc.) that could account for memory for givens and theorems as well. For the purposes of this example, however, we will record only Yes and No detail information and show how even this simple notation can account for subjects' memory for steps.

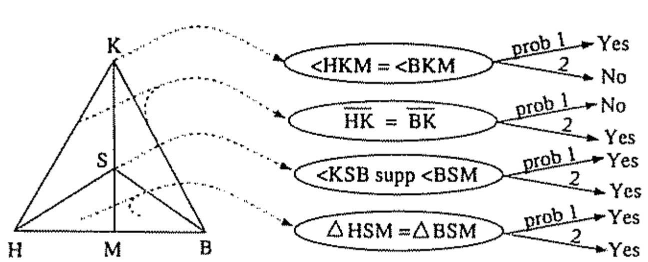


Figure 2. Diagram-based memory schema for geometry problem solving. Components of the schema are possible relations in the diagram, and details of the problem-solving memory are Yes or No, indicating whether that relation was established in the proof.

In the study by Thorndyke and Hayes-Roth (1979), interference is explained by the fact that different details from distinct events are attached to the same component of the schema, making it hard to retrieve one such detail for a particular event. In the geometry schema of Figure 2, the attached details are simply Yes or No instead of specific facts. Hence, a not-shared step (e.g., $\angle HKM = \angle BKM$) is represented in such a way that the No attached because of Problem 2 interferes with the Yes attached because of Problem 1, making more difficult the recall of whether that step was specifically used in Problem 1. If the two problems had different diagrams, however, this difficulty would not arise because the two memories would be associated with two different schemata.

For shared steps, the situation is similar except that Yes would be twice attached to a single component of the same schema when the two problems had the same diagram (e.g., $\angle KSB$ is supplementary to $\angle BSM$), and Yes would be attached to two completely separate schemata when the two problems had different diagrams. Facilitation is expected in the former case, relative to the latter. Thorndyke and Hayes-Roth (1979) suggested that this facilitation mainly occurs during the process of building the schema. This is because, as the schema is established, the accessibility of its components increases, directly affecting the probability of recalling a shared detail. In the current experiment, facilitation from establishing a new schema is unlikely because these subjects had substantial previous geometry experience (e.g., they were all able to solve the problems after merely reading a page-long review sheet). Thus, the lack of observed facilitation in memory for shared steps might be a result of the subjects' ability to draw on an already established schema in which facilitation effects no longer play a large role.

Experiment 2

What might these memory results entail for problem-solving transfer? Our preliminary answer is that subjects should show more problem-solving transfer in those conditions that resulted in better memory for problem-solving episodes and less transfer in the conditions with worse memory. This is by no means a logical necessity; in the literature on problem-solving transfer and analogical problem solving, quality of memory is rarely considered (e.g., Faries & Reiser, 1988; Gentner, 1989; Gick & Holyoak, 1983; Singley & Anderson, 1989). One exception is the finding that subjects' remindings during problem solving are sensitive to many of the same interfering and facilitating effects found in the memory literature (Ross, 1982). Another is the set of results from Thorndyke and Hayes-Roth (1979), mentioned above, that the memory structures subjects develop during training have an impact on their subsequent memory performance. If subjects make use of their original problem-solving memory when they solve a subsequent problem (e.g., by analogical problem solving), the quality of that memory should have an effect on transfer.

In Experiment 2, we tested the hypothesis that problem-solving transfer depends on the quality of problem-solving memories by substituting a test problem for the memory tests of Experiment 1. The test problem was designed to be identical to the original (except for labels on its diagram) to create a situation most conducive to use of the original problem's memory. We predicted that subjects' performance on the test problem would reflect our memory results from the previous experiment: Subjects solving two problems with the same diagram, a situation that caused memory interference for changed aspects of the original proof, should take longer to solve the test problem than subjects solving two problems with different diagrams. In particular, we predicted that the same-diagram, unrelated logic condition, which was most disadvantaged by this memory interference, would have the longest solution times. In contrast, we predicted that control subjects, who solved only the second problem followed by the test problem, would not exhibit these effects because their problem memories would not have been differentially affected by the diagrammatic and logical similarity of an interpolated problem. Instead, each of these subjects would have only a single problem-solving memory to transfer to the test problem, equal in its quality but different in its content across conditions. Thus, we predicted that the relevance of this problem-solving memory would account for control subjects' transfer: Subjects whose second problem was logically analogous to the test should solve the test problem faster than subjects whose second problem was logically unrelated to the test.

Method

Subjects. Subjects were 48 Carnegie Mellon University undergraduates, participating either for credit or for payment. Complete data from three other subjects were not available due to a technical error, so these subjects were not included in our analyses.

Design. The design resembled that of Experiment 1 in that the second proof varied in its relationship to the original according to diagrammatic similarity (same or different) and logical similarity (analogous or unrelated). In Experiment 2, however, both types of similarity were between-subjects factors. Subjects were randomly assigned to one of the four conditions (with the constraint that there were 12 subjects per cell), and this assignment was maintained across each of two experimental sessions. Subjects solved a three-problem sequence (original problem, second problem, test problem) in one session and a two-problem sequence (second problem, test problem) in another session. The two-problem sequence served as a control to compare subjects' performance on the test problem when they had not solved the original problem. Again, two problem sets were required for the two-session design. Assignment of sequence type (three-problem or two-problem sequence) to the first and second sessions was counterbalanced, as was the matching of problem set to session number.

Note that several aspects of the two-problem sequences are derived from the corresponding three-problem sequences. For example, we use the label *second problem* to denote what is actually the first problem in the two-problem sequence. Also, we use the diagrammatic and logical similarity factors to describe the four possible two-problem sequences, even though these labels were defined in terms of similarity to the original problem, which is absent in the two-problem sequences. Because the test problem is identical to the original in its diagram and

Table 4
Test Problem Solution Times (in Minutes) for Three-Problem and Two-Problem Sequences in Experiment 2

	Diagram	
Sequence	Same	Different
Three-problem		
Analogous logic	5.8	5.7
Unrelated logic	7.4	5.1
Two-problem		
Analogous logic	8 5	8 2
Unrelated logic	11.6	14 3

logic, these similarity relationships still apply in the two-problem sequences

Materials The original and second problems used during the two sessions came from the two problem sets used in Experiment 1. For each problem set, the test problem was identical to the original except for a change of labels

Procedure Subjects studied a geometry review sheet (as in Experiment 1) and were asked to solve a sequence of geometry problems during each of two sessions Only three changes were made to the procedure First, problems were presented at 20-min intervals to equalize any effects caused by the different amounts of time needed for the various problems Subjects who finished the current problem in less than the time allotted were asked to hand in their solution and work on an unrelated puzzle for the time remaining Second, to provide subjects with enough help to solve the problems correctly but not to introduce unnecessary bias through experimenter involvement, subjects were asked whether they had any questions at approximately 5-min intervals As in Experiment 1, only the next step in the optimal solution path was given The total number of hints was less than one per subject per problem Third, 2 subjects in each of the four conditions provided talk-aloud protocols throughout all problem solving to allow for a more detailed investigation of how and when subjects made use of their problem-solving memories Their solution times on the first geometry problem averaged 9 2 min, $SD = 2.9$, compared with the rest of the subjects, who averaged 13 2 min, $SD = 5.0$. On the test problem in the three-problem sequence, however, these differences seemed to disappear: There was no reliable difference between the protocol subjects and the others, nor was the pattern of performance across conditions different between the protocol subjects and the others, all $ps > .10$

As in Experiment 1, we tested subjects' sensitivity to the various problem relationships by asking for similarity ratings after subjects had solved two problems (in the two- and three-problem sequences) and after they had solved all three problems (in the three-problem sequence only). Subjects rated their first two problems as more similar in the same diagram conditions than in the different diagram conditions, $F(1, 44) = 5.81$, $MS_e = 1.10$, $p < .05$. Similarly, subjects in the analogous logic conditions rated their first two problems as more similar than did subjects in the unrelated logic conditions, $F(1, 44) = 9.61$, $MS_e = 1.10$, $p < .01$ Furthermore, the lack of a Diagram × Logic interaction replicated our previous finding that these factors contributed to subjects' ratings in an additive fashion, $F(1, 44) < 1$ In addition, subjects rated the similarity of all three problems in the three-problem sequence significantly higher than they had rated their first two problems, suggesting that subjects were also sensitive to the similarity added by the test problem's isomorphic relationship to the original, $t(47) = 6.82$, $p < .001$.

Finally, at the end of the three-problem sequence, we asked subjects, "Did you notice any particular relationships among the problems you solved?" and "Did you ever think back to a previous problem solution?" to probe their awareness of the identity relationship between the original and test proofs Their responses were used to explore how subjects made use of their memories

Data analysis Except where noted, we used first-problem solution times as the covariate in analyses of test-problem solution times to account for individual differences in geometry problem solving In all analyses of covariance, the covariate was significant, $p < .01$.

Results and Discussion

In light of the first experiment's results, we expected that subjects' performance on the test problem from the three-problem sequence would reflect less transfer in the same diagram conditions than in the different diagram conditions. In particular, we expected the least transfer in the same diagram, unrelated logic condition because in the previous experiment, a greater absolute number of not-shared steps failed to be recalled in this condition compared with the same diagram, analogous logic condition, making it the most memory-disadvantaged For the two-problem sequence, we predicted greater transfer in the analogous logic conditions than in the unrelated logic conditions, but we predicted no differences according to diagrammatic similarity. Therefore, our predictions for performance on a particular test problem were based not only on the preceding second problem but also on the sequence type in which the test problem occurred. In this way, our predictions were sensitive to subjects' complete solution histories.

Solution times As shown in Table 4, solution times for the test problem in the three-problem sequence were longer in the same diagram conditions than in the different diagram conditions, $F(1, 43) = 4.26$, $MS_e = 4.26$, $p < .05$ Test solution times did not differ with respect to the logical similarity between the original and second problems, $F(1, 43) < 1$. The finding that diagrammatic similarity but not logical similarity led to longer test problem solution times matched our prediction that the memory-impaired conditions from Experiment 1 would exhibit less transfer. There was also a marginally significant interaction of Diagram × Logic, $F(1, 43) = 3.585$, $p < .07$ This finding is consistent with our prediction that the same diagram, unrelated logic condition, which was the most disadvantaged according to our memory measures, would have the longest solution times $F(1, 43) = 7.62$, $MS_e = 4.26$, $p < .01$. The other three conditions did not differ significantly from each other, $F(2, 32) < 1$.

These results are reminiscent of the transfer results compiled and modeled by Osgood (1949). In these experiments, Osgood measured transfer in the context of a three-problem sequence similar to the one above and found that performance on the third (test) problem, identical to the original, varied with the similarity relationships between the original and second problems For Osgood, each of these problems consisted of learning a stimulus–response pair, and the relationship between the original and second pairs was manipulated by independently varying the similarity of the two stimuli and the similarity of the two responses Subjects exhibited substantial transfer when the original and second pairs had similar stimuli and similar responses, and they exhibited extreme interference when the two pairs had similar stimuli but opposite responses. These results correspond to some of our findings above, with

diagrams viewed as the stimuli and proofs viewed as the responses: The first situation (similar stimuli and similar responses) is like the same diagram, analogous logic condition above in which we also found substantial transfer, and the second situation (similar stimuli and opposite responses) corresponds to our same diagram, unrelated logic condition where subjects demonstrated the least transfer. An interesting difference between our experiment and the experiments studied by Osgood is the greater complexity of geometry problem-solving memories compared with stimulus–response pair memories. Also, our performance measures always indicated positive transfer; the degree of positive transfer varied similarly with Osgood's positive and negative transfer effects. Osgood's model of these transfer effects is purely descriptive, making only general predictions and lacking an underlying explanation. In contrast, our predictions rely on specific aspects of geometry problem-solving memories (e.g., shared and not-shared steps) and on memory models, like that of Thorndyke and Hayes-Roth (1979), to explain how problem-solving memories can account for transfer

Solution times for the test problem of the two-problem sequence (see Table 4) revealed a main effect of logic, $F(1, 43) = 12.43$, $MS_e = 16.96$, $p < .001$, but no diagram effect and no Diagram × Logic interaction, $Fs < 1$. These results confirm our prediction with respect to solution times on the test problem in the two-problem sequence: Subjects benefited from their problem-solving memory to the degree that the previous problem was logically relevant, but memory interference did not come into play

Another important aspect of these solution time data was that subjects were faster at solving the test problem in the three-problem sequence than they were at solving the same test problem in the two-problem sequence regardless of their condition, $t(47) = 6.421$, $p < .001$. Furthermore, the latency pattern for the three-problem sequence, but not the two-problem sequence, matched our memory results. Therefore, these results strongly suggest that the effects we found for test problem solution times in the three-problem sequence were due to transfer from the original problem.

Hint requests. The number of subjects (out of 12 per cell) who received at least one hint on the test problem in the three-problem sequence was quite small, as is shown in Table 5. In fact, these numbers were too small to be reasonably compared with each other. Nevertheless, according to McNemar's test for dependent proportions (Agresti, 1990), a comparison between the hint data in the two- and three-problem sequences (in Table 5) revealed that fewer subjects received hints on the test problem in the three-problem sequence (i.e., when they had previously solved an identical problem) than in the two-problem sequence, $\chi^2(1, N = 48) = 5.4$, $p < .05$. This difference confirms the idea that subjects were transferring from the original problem.

Hint requests for the test problem in the two-problem sequence suggested that subjects' need for hints decreased as the logical similarity of the previous problem increased, $\chi^2(1, N = 48) = 7.06$, $p < .01$. This result was consistent with our prediction that subjects' test performance in the two-problem sequence would depend on the relevance of the previous problem's solution and, hence, show a logic-based effect.

Table 5
Number of Subjects (out of 12 per Cell) Needing Hints on the Test Problem in Three-Problem and Two-Problem Sequences in Experiment 2

	Diagram	
Sequence	Same	Different
Three-problem		
Analogous logic	2	3
Unrelated logic	2	0
Two-problem		
Analogous logic	2	0
Unrelated logic	5	9

Protocols. The talk-aloud protocols, taken from two subjects in each condition, provided more than an hour and a half of subjects' verbalizations, but only on four occasions did subjects make reference to a previously solved problem. All four of these occasions occurred during solution of the test problem in the three-problem sequence. Two were specific references in that subjects verbalized explicit analogical problem solving and took a correct step as a result. One subject each in the same diagram, analogous logic condition and the different diagram, analogous logic condition made these references. The other two excerpts involved a general reference to a previous problem, but subjects failed to make use of the previous solution. These references came from one subject each in the same diagram, analogous logic condition and the different diagram, unrelated logic condition.

It is interesting to note that subjects exhibited explicit analogical problem solving in only two out of the four remindings and that reference to a previous solution was rare. Why were subjects not making more references to their previous proofs? It certainly was not because they failed to verbalize what they were thinking. In general, the protocol subjects verbalized many elements in the current problem that must have been passing through short-term memory (Ericsson & Simon, 1984). External memory for the current problem may have facilitated these verbalizations, whereas referring to a problem no longer in view may have been considerably more difficult. This difference may have made explicit analogical problem solving unlikely in this experiment, and it suggests that subjects might have been using their previous problem-solving memories in a more implicit way.

Lack of awareness and implicit memory. Awareness data also support the possibility of some form of implicit transfer of problem-solving memories. Recall that at the end of the three-problem session, we asked subjects about their awareness of the identity relationship between the original and test problems. Subjects were categorized as "totally aware" if they spontaneously verbalized awareness of the identity or if they responded to the probes in a way that demonstrated explicit awareness, $n = 23$. Subjects were categorized as "not totally unaware" if they did not fulfill either of the above criteria, $n = 25$. We divided subjects into these two groups to distinguish between subjects who would definitely have been able to consider explicit use of their original problem-solving memory and subjects who would not necessarily have been considering doing so. These two categories of awareness were then used as the two levels of an additional factor in the analysis of subjects' solution times for the test problem in the three-problem sequence. In this analysis, we used subjects' solution times for the test problem in the two-problem sequence as the covariate in order to focus on the effect of having solved the original problem and to factor out the effects of solving different second problems. This analysis revealed no significant main effect of awareness, nor any significant interactions involving awareness, all $ps > .10$. This suggests that, although we found a large effect of the original problem in reducing solution times for the test problem, this benefit did not depend on whether subjects were aware of the identity relationship between these problems. Moreover, the lack of any interactions involving awareness indicates that the pattern of solution times across conditions was not affected by awareness. This issue is further explored in Experiment 3.

Summary of results. In Experiment 2, we found that subjects' test problem performance was consistent with our predictions. Subjects took significantly longer to solve the test problem in the three-problem sequence when their original and second problems had the same diagram, paralleling the memory results of Experiment 1. Furthermore, subjects exhibited the worst performance in the same diagram, unrelated logic condition—just as our memory interference explanation predicted. On the test problem in the two-problem sequence, solution times and hint requests were reduced in the analogous logic conditions. This result confirms the prediction that, given equal quality of problem-solving memory across conditions in the two-problem sequence, transfer will be facilitated to the degree that the particular problem-solving memory in each condition is logically relevant. Because the diagrammatic similarity manipulation, which interfered with subjects' memories in Experiment 1, reduced problem-solving transfer in the three-problem sequence but not in the two-problem sequence, it appears that subjects' problem-solving performance in the three-problem sequence was influenced by the quality of their memories for the original problem-solving episode.

We also found that many subjects lacked awareness of the identity relationship between the original and test problems and that, regardless of such awareness, the quality of subjects' original problem-solving memory had an effect on transfer. Our protocol subjects rarely demonstrated explicit analogical problem solving, further suggesting that subjects' memory transfer may be somehow implicit.

Experiment 3

One of the main goals of Experiment 3 was to replicate our previous results, using a different test problem. First, the test problem was changed from being isomorphic to the original problem to being only analogous. This provided a more typical testing situation and reduced the strong bias to make use of previous problem-solving memories as was the case in Experiment 2. The test problem was also changed so that its diagram was different from the original problem's diagram and from all of the possible second problems' diagrams. This change was made so that subjects could not choose a particular problem-solving memory to help solve the test problem by cuing off the test problem's diagram. Such a cuing strategy could have explained the results of Experiment 2 in that (a) it would have allowed subjects in both different diagram conditions to choose the correct analogue (viz., the original problem) merely by selecting the unique previous problem whose diagram matched the test problem's and (b) it would have led subjects in the same diagram conditions to choose an incorrect analogue (viz., the second problem) half of the time because both of their previous problems' diagrams matched the test problem's. This choice would have been especially detrimental in the same diagram, unrelated logic condition because, here, the second problem's unrelated logic made it an inappropriate analogue.

With a different diagram for the test problem in Experiment 3, diagram cuing could no longer account for the predicted pattern of results. Reduced transfer effects in the same diagram conditions caused by differential interference in problem-solving memories, however, still could account for such a pattern. Therefore, as in Experiment 2, we predicted that subjects' performance on the test problem in the three-problem sequence would exhibit less transfer in the same diagram conditions than in the different diagram conditions. Also, as in Experiment 2, we predicted that transfer performance in the two-problem sequences would vary with logical similarity.

The other main goal of this experiment was to discern the explicit–implicit nature of subjects' memory usage. An attempt was made to experimentally manipulate subjects' explicit awareness of the relevance of the original problem's solution to solving the test problem. Half of the subjects solving the test problem in the three-problem sequence received a hint telling them to "think back" to the original problem to solve the current problem; the other half did not receive such a hint. If subjects were at least attempting to make explicit use of previous problem-solving memories, then an increased awareness of the potential usefulness of the original problem's solution should increase transfer.

Method

Subjects. Subjects were 100 undergraduates at the University of California at Berkeley who were paid for their participation.

Design. This experiment included 10 conditions—8 experimental and 2 control. The 8 experimental conditions used the three-problem sequence and followed a 2 × 2 × 2 design, with logical and diagrammatic similarity (from Experiment 2) as two factors and presence of the hint described above as the third factor. As in Experiment 2, the diagrammatic and logical similarity factors were manipulations of a variable second problem's relationship to the original problem. (Again, the test problem had a different diagram from both preceding problems, no matter what the diagram factor implied about the relationship between the original and second problems' diagrams.) The control conditions were two-problem sequences, created by deleting the original problem from the three-problem sequences. The hint–no-hint factor did not apply to the two-problem sequence, so only the second problem (i.e., the sole problem preceding the test) needed to be varied. In contrast with Experiment 2, we did not maintain the distinction between same and different diagram conditions in the two-problem sequence because such labels would be misleading (i.e., the second problem's diagram was always different from that of the test problem). Instead, we distinguished the two-problem sequences according to the logical similarity between their two proofs—analogous and unrelated.

Materials. The original and second geometry problems for this experiment were only slightly changed from the previous experiments. The test problem (see Appendix B) was specially designed for this experiment, with its diagram different from any of the other problems' diagrams and its logical structure analogous (but not isomorphic) to the original problem's logic. Therefore, the test problem had the same logical structure as the original, but its steps were not isomorphic to the original problem's.

Procedure. The procedure for this experiment did not deviate significantly from that of Experiment 2. The differences were that subjects participated in one session only, subjects were not asked for similarity ratings or asked about the relationship between the original and third problems (because the hint–no-hint factor manipulated this awareness), and no protocol subjects were included.

Data analysis. As in Experiment 2, analyses of test-problem solution times used first-problem solution times as the covariate. In all cases, the covariate was significant, $p < .01$.

Results and Discussion

Solution times. Subjects' solution times for the test problem in the three-problem sequence were used as the main dependent variable. These data revealed a significant main effect of logical similarity and a main effect of diagrammatic similarity, $F(1, 71) = 4.89, MS_e = 9.51, p < .03$, and $F(1, 71) = 3.59, MS_e = 9.51, p < .06$, respectively. Subjects took longer to solve the test problem when their two previous problems had unrelated logic. Subjects also took longer to solve the test problem when their two previous problems had the same diagram, replicating Experiment 2 and matching the memory data. The hint–no-hint factor, however, did not reach statistical significance nor did any interactions involving it. Accordingly, we collapsed over this factor in our presentation of the data and in all subsequent analyses. Table 6 contains these solution time data for the three-problem sequence conditions and for the two control conditions.

As in Experiment 2, the planned comparison between the same diagram, unrelated logic condition and the other three conditions revealed that subjects in the same diagram, unrelated logic condition had longer solution times than the others, $F(1, 77) = 8.06, MS_e = 9.5, p < .01$. The other three conditions were not significantly different from one another, $F(2, 56) < 1$. This replicated the result of Experiment 2 and confirmed the prediction that problem-solving transfer was most impaired in the same diagram, unrelated logic condition, the most disadvantaged condition with respect to memory interference.

Subjects' test-problem solution times in the control conditions (two-problem sequences) also confirmed our predictions. Here, subjects in the analogous logic condition had shorter test-problem solution times than did subjects in the unrelated logic condition, $F(1, 17) = 15.19, MS_e = 14.54, p < .001$. This result replicated the Experiment 2 finding that, without the possibility for retroactive interference effects on the original problem memory, subjects exhibited transfer on the basis of the logical similarity of their single previous problem-solving memory.

Hint requests. To capture individual subject's transfer in terms of hint requests, we tallied the subjects whose number of hint requests increased, decreased, or stayed the same between the original problem and the test problem in the three-problem sequence. The number of subjects whose hint requests increased was 2, decreased 10, and stayed the same 8 in the same diagram, analogous logic condition; increased 2, decreased 4, and stayed the same 14 in the same diagram, unrelated logic condition; increased 0, decreased 15, and stayed the same 5 in the different diagram, analogous logic condition; and increased 2, decreased 10, and stayed the same 8 in the different diagram, unrelated logic condition. These three degrees of hint requests corresponded to situations of negative transfer, positive transfer, and no transfer, respectively. Subjects in the same diagram, unrelated logic condition were distributed in these categories quite differently from subjects in the other three conditions, $\chi^2(2, N = 80) = 8.94$, $p < .05$. Specifically, fewer subjects in the same diagram, unrelated logic condition exhibited positive transfer in terms of hint requests, again demonstrating that this was the most disadvantaged cell.

Summary of results. In this experiment, we replicated three important results of Experiment 2. First, transfer to the test problem in the three-problem sequence was reduced when the original and second problems had the same diagram. Second, this reduction in transfer was most pronounced in the same diagram, unrelated logic condition. These two findings, that the conditions that caused greater memory interference also led to reduced transfer, again strengthened the argument that quality of problem-solving memories can impact problem-solving transfer. Third, in the two-problem sequence, transfer varied with the logical similarity of the problems, supporting the notion that memory effects across the two sequence types were qualitatively different. In addition, the lack of a hint effect in this experiment provided converging evidence that the memory effects we observed in problem-solving transfer were not necessarily based on explicit memory use.

The above results generalized those of the previous experiment because they were observed with a new test problem that was analogous (not isomorphic) to the original and that had a diagram different from all the previous problems. The fact that these results were maintained when the test problem had a different diagram from all previous problems weakened the diagram cuing explanation described earlier and yet continued to support our quality of problem-solving memories explanation.

One new effect, however, was revealed in this experiment—solution times in the three-problem sequence were shorter in the analogous logic conditions than in the unrelated logic

Table 6
Test Problem Solution Times (in Minutes) for Three-Problem and Two-Problem Sequences in Experiment 3

Sequence	Diagram: Same	Diagram: Different
Three-problem		
Analogous	8.5	7.7
Unrelated	10.6	8.7
Two-problem		
Analogous	—	11.2
Unrelated	—	16.3

Note. Dashes indicate certain conditions not present in the design.

conditions. This effect was not predicted, but it might be explained by a combination of the diagram-based effect on third problem solution times and by the finding that subjects in the same-diagram, unrelated logic condition took longer to solve the third problem than did subjects in the other three conditions.

General Discussion

This series of experiments has provided results that begin to answer our preliminary questions about the relationship between problem-solving memories and transfer. The first is that the diagram, not the logical structure of the proof, appears to serve as the basis of memory for proof. We believe that this result is not specific only to geometry. In other domains, the diagram seems central to problem solving even when that problem solving could also be done symbolically. Memory for symbolic problem solving may be tied to semantic referents of those symbols in many domains. This may reflect the frequent finding of better memory for visual information than for symbolic information.

The second result is that the quality of memory for a specific problem-solving episode seemed to impact later success in solving analogous problems. We believe that this result provides another piece of evidence for the point of view that problem-solving transfer is strongly influenced by memory for other specific problem-solving episodes.

The third major result is somewhat peculiar in light of the second. This is the lack of evidence that awareness of the past problem's relationship to the current problem facilitates problem-solving transfer. In all cases, the original problem's solution was designed to be helpful in solving the test problem. And yet our protocol data, subjects' responses to the awareness probes, and subjects' assignment to a condition in which they were told about the relationship all seemed to indicate that the original problem memory influenced subjects' test solutions without their necessarily being aware of it.

With such striking performance effects and yet almost no data supporting explicit memory use, it seems that any mechanism to account for our observations must rely to some degree on implicit memory. Brooks (1978) has argued that implicit memory phenomena could be tied to unconscious use of past instances in an artificial grammar task (Reber, 1967, 1976). Similarly, Marescaux, Dejean, and Karnas (1990) have argued that the same is true for tasks in which subjects must manipulate a variety of input variables to maintain a target output value (Berry & Broadbent, 1988). Perhaps in geometry, subjects are also able to implicitly make use of previously solved problems, through biases to use certain steps or through the repetition of particular patterns of steps. Investigation of these possibilities, currently under way, may reveal new characteristics of problem-solving transfer and may eventually help us to better understand the underlying processes therein.

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory* (pp. 488–494). Hillsdale, NJ: Erlbaum.

Berry, D. C., & Broadbent, D. E. (1988). Interactive task and the implicit–explicit distinction. *British Journal of Psychology, 79,* 251–272.

Bower, G. H. (1974). Selective facilitation and interference in retention of prose. *Journal of Educational Psychology, 66,* 1–8.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177–220.

Brooks, L. (1978). Non-analytic concept formation and memory for instances. In E. Rosh & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Crouse, J. H. (1971). Retroactive interference in reading prose materials. *Journal of Educational Psychology, 62,* 39–44.

de Groot, A. D. (1965). *Thought and choice in chess*. New York: Basic Books.

Egan, D. E., & Greeno, J. G. (1974). Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning, and problem solving. In L. W. Gregg (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Faries, J. M., & Reiser, B. J. (1988). Access and use of previous solutions in a problem solving situation. In *Proceedings of the Tenth Annual Conference of the Cognitive Science Society* (pp. 433–439). Hillsdale, NJ: Erlbaum.

Gentner, D. (1989). Mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). Cambridge, England: Cambridge University Press.

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

Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science, 14,* 511–550.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–99.

Marescaux, P., Dejean, K., & Karnas, G. (1990). *Acquisition of specific or general knowledge at the control of a dynamic simulated system: An evaluation through a static situations questionnaire and a transfer control task* (Report 2PR2GK). Brussels, Belgium: Knowledge Acquisition and Use in Dynamic Task Environments (KAUDYTE) Project.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56,* 132–143.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855–863.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 88–94.

Ross, B. H. (1982). *Remindings and their effects on learning a cognitive skill* (Tech. Rep. No. CIS-19). Palo Alto, CA: Xerox Research Center.

Ruiz, D. (1987). *Learning and problem-solving: What is learned while solving the Towers of Hanoi?* Unpublished doctoral dissertation, Stanford University.

Silver, E. A. (1981). Recall of mathematical information: Solving related problems. *Journal for Research in Mathematics Education, 12,* 54–64.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambrdge, MA: Harvard University Press.

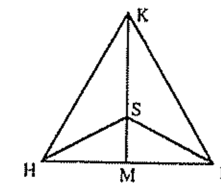
Thorndyke, P. W., & Hayes-Roth, B. (1979). The use of schemata in the acquisition and transfer of knowledge. *Cognitive Psychology, 11,* 82–106.

Watkins, M. J., & Kerkar, S. P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language, 24,* 666–678.

Appendix A

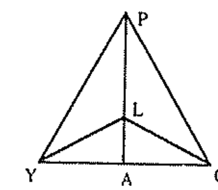
Sample Solutions

Original problem — Given: $\angle HKM \equiv \angle BKM$, $\angle HSM \equiv \angle BSM$
Prove: $\overline{HB} \perp \overline{KM}$



1. $\angle HKM \equiv \angle BKM$	1. Given
2. $\angle HSM \equiv \angle BSM$	2. Given
3. $\angle HSK + \angle HSM = 180°$	3. Supplementary angles
4. $\angle BSK + \angle BSM = 180°$	4. Supplementary angles
5. $\angle HSK \equiv \angle BSK$	5. Supplement theorem
6. $\overline{KS} \equiv \overline{KS}$	6. Reflexive property
7. $\Delta HSK \equiv \Delta BSK$	7. Angle-Side-Angle postulate
8. $\overline{HS} \equiv \overline{SB}$	8. Corresponding parts of congruent triangles are congruent (CPCTC)
9. $\overline{SM} \equiv \overline{SM}$	9. Reflexive property
10. $\Delta HSM \equiv \Delta BSM$	10. Side-Angle-Side postulate
11. $\angle HMS \equiv \angle BMS$	11. CPCTC
12. $\angle HMS + \angle BMS = 180°$	12. Supplementary angles
13. $\angle HMS$ right angle	13. Right angles theorem
14. $\overline{HB} \perp \overline{KM}$	14. Definition perpendicular lines

Same-analogous problem — Given: $\overline{PA} \perp \overline{YG}$, $\angle PLY \equiv \angle PLG$
Prove: ΔYPG is isosceles

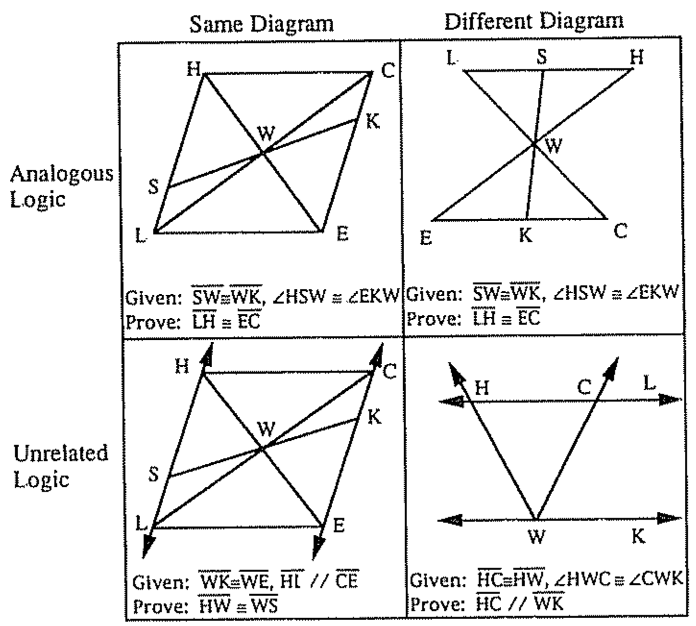


1. $\overline{PA} \perp \overline{YG}$	1. Given
2. $\angle PLY \equiv \angle PLG$	2. Given
3. $\angle PAY \equiv \angle PAG$	3. Right angles theorem
4. $\angle PLY + \angle YLA = 180°$	4. Supplementary angles
5. $\angle PLG + \angle GLA = 180°$	5. Supplementary angles
6. $\angle YLA \equiv \angle GLA$	6. Supplement theorem
7. $\overline{LA} \equiv \overline{LA}$	7. Reflexive property
8. $\Delta LYA \equiv \Delta LGA$	8. Angle-Side-Angle postulate
9. $\overline{LY} \equiv \overline{LG}$	9. CPCTC
10. $\overline{PL} \equiv \overline{PL}$	10. Reflexive property
11. $\Delta PLY \equiv \Delta PLG$	11. Side-Angle-Side postulate
12. $\overline{PY} \equiv \overline{PG}$	12. CPCTC
13. ΔYPG is isosceles	13. Definition isosceles triangle

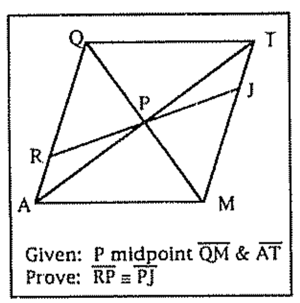
(Appendix B follows on next page.)

Appendix B

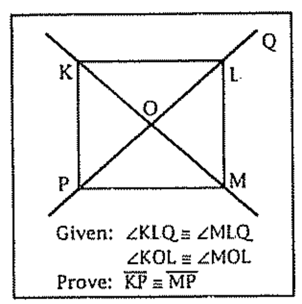
Additional Problems From Experiments 1, 2, and 3



Second problems according to condition (Experiments 1 and 2)



Original problem
(Experiments 1 and 2)



Test problem
(Experiment 3)

Received September 19, 1991
Revision received April 1, 1993
Accepted April 16, 1993 ■