Improving examples to improve transfer to novel problems

RICHARD CATRAMBONE
Georgia Institute of Technology, Atlanta, Georgia

People often memorize a set of steps for solving problems when they study worked-out examples in domains such as math and physics without learning what domain-relevant subgoals or subtasks these steps achieve. As a result, they have trouble solving novel problems that contain the same structural elements but require different, lower-level steps. In three experiments, subjects who studied example solutions that emphasized a needed subgoal were more likely to solve novel problems that required a new approach for achieving this subgoal than were subjects who did not learn this subgoal. This result suggests that research aimed at determining the factors that influence subgoal learning may be valuable in improving transfer from examples to novel problems.

A number of studies have indicated that learners rely heavily on worked-out examples when trying to solve novel problems (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985). Unfortunately, in a domain, novices have great difficulty separating the features of the examples that are necessary to the solution procedure from those that are incidental (Ross, 1987, 1989). In addition, learners have difficulty generalizing solutions from examples to structurally similar, but nonisomorphic, problems (Reed, Ackinclose, & Voss, 1990; Reed, Dempster, & Ettinger, 1985). Although certain training manipulations have succeeded in improving transfer from examples to novel problems to some degree (Lewis & Anderson, 1985; Zhu & Simon, 1987), in general, transfer has not been impressive (e.g., Gick & Holyoak, 1983; Reed et al., 1985; Ross, 1987, 1989).

Learners differ in what they extract from worked-out examples. Chi, Bassok, Lewis, Reimann, and Glaser (1989) found that good learners are more likely to try to understand *why* a particular step was taken in a solution. Good and poor learners, at least initially after studying a physics chapter without examples, seemed to have a similar level of declarative knowledge about mechanics. However, when studying an example, the good learners produced explanations that contained more "inferences about the conditions, the consequences, the goals, and the meaning of various mathematical actions described in the example" (p. 168). Thus, good learners seem to get more from examples, including a knowledge of goals,

even when starting at a knowledge level similar to that of poor learners.

The present experiments explore whether more learners can be turned into good learners by presenting them with examples that convey the subgoals relevant for solving problems in a domain. The term *subgoal* is used here to represent the task structure to be learned for solving problems in a particular domain (e.g., Catrambone & Holyoak, 1990; Dixon, 1987; Eylon & Reif, 1984). A subgoal groups a set of steps under a meaningful task or purpose (e.g., Anzai & Simon, 1979; Chi & VanLehn, 1991). I hypothesize that a person who has learned the subgoal will be in a better position to achieve it in a novel problem requiring a new or modified set of steps than will someone who has not learned the subgoal.

Learning subgoals is assumed to enhance performance because subgoals act as guides to the part or parts of the procedure demonstrated in examples that need to be changed for the current problem. Thus, subgoals narrow the space in which the solver has to search in order to determine what must be changed. For instance, Simon and Reed (1976) found that providing learners with a subgoal—in the form of a hint to achieve a particular state along the solution path in a problem involving missionaries and cannibals—aided their navigation through the problem space.

In the probability materials used in the first two experiments, one subgoal toward the overall goal of finding a particular probability is to find the probability of each of the individual events. If the steps for finding an individual event probability in a novel problem are not the same as those used in the example, then a person who has learned the subgoal to find the individual event probabilities will have a better chance of focusing on the steps of the procedure that must be changed—the steps involved in finding the individual event probabilities—than a person who has learned only a set of steps for finding the overall probability. For this second learner,

the steps for finding the individual probabilities are obscured because they are simply part of a longer set of steps for reaching the end goal. This learner would have fewer cues to direct him or her to the appropriate steps that need to be changed.

There may not be a theoretically "best" set of subgoals for solving problems in a domain. The particular subgoals that are taught might represent an instructor's judgment about how students should decompose problems into subproblems in order to solve novel problems most effectively. The judiciousness of the instructor's choice of subgoals can be measured by the success of the learners on novel problems.

A learner will be more likely to learn subgoals for the subparts of a general solution procedure if those subparts are emphasized in the examples' solutions. This claim hinges on the assumption that people will form a representation that is based on the most salient features of the example solution. The salience of a feature will vary, depending on the learner's expertise in the domain and how the solution is presented (Larkin, McDermott, Simon, & Simon, 1980; Ward & Sweller, 1990). In the present experiments, steps were labeled and visually separated as a technique for encouraging the formation of a subgoal.

An important first step in creating useful examples is to perform a task analysis to determine what elements need to be learned in order to solve problems in the domain of interest. *How* a domain is analyzed to produce those elements is by no means standardized. One promising approach is to create a set of production rules that solves problems or carries out procedures that one wants learners to be able to solve or learn (e.g., Anderson, Boyle, Farrell, & Reiser, 1987; Kieras & Bovair, 1986; Zhu & Simon, 1987). Examples can then be created in order to teach these productions to learners.

It does not seem necessary to make a commitment to a production rule formalism embodying a particular learning theory such as ACT* (Anderson, 1983) or Soar (Laird, Newell, & Rosenbloom, 1987) in order to derive the elements that need to be learned. However, a fundamental feature of most production rule systems—the goal structure—does provide a useful way to represent the knowledge needed to solve problems in a domain. Subgoals show the breakdown of a problem-solving procedure into subproblems (Anzai & Simon, 1979). Depending on the features of the examples and of the learner, the subgoals learned from examples could represent either a flexible and general approach to solving problems in a particular domain, or a rigid and superficial approach.

Learning From Examples

A number of studies have shown that manipulations of examples have a powerful and systematic effect on performance on novel problems (e.g., Catrambone & Holyoak, 1990; LeFevre & Dixon, 1986; Pirolli & Anderson, 1985; Reder, Charney, & Morgan, 1986; Ross, 1984). Given the central role that examples play in prob-

lem solving, and given the assumption that people learn subgoals from the examples, it is important to investigate the conditions that influence subgoal learning.

One reason that many learners do not form the "right" subgoals (as determined by an experimenter or instructor) is because examples typically are not designed to convey them. This observation echoes an inadequacy in a mechanics example from a physics textbook noted by Chi et al. (1989, p. 149). In the example, a block is suspended from a ceiling by two pieces of rope joined at a knot and a third piece of rope extending from the knot to the block. The task is to find the magnitude of two of the forces, given the third force. The solution states that the knot where the three strings are joined should be considered the body. However, no explanation is given as to why this decision is made. The decision is made because, in order to find a force in terms of other forces, the forces must all act on a common point. In this problem, the only place where all three forces act is the knot. This critical subgoal of finding a common point where the forces are acting would be useful for many future problems. However, instead of conveying this subgoal, the example is more likely to convey a series of steps that may or may not be useful for other problems.

One question at this point: Why not directly state the subgoals to learners rather than embedding them in examples? There are two problems with this approach. First, learners exhibit a clear preference for learning from and referring to examples when faced with new problems (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985). Second, although there have been a small number of successes teaching solution procedures directly (Fong, Krantz, & Nisbett, 1986), most attempts have been unsuccessful (e.g., Reed & Bolstad, 1991).

Overview of Experiments 1 and 2

In the first two experiments, subjects studied examples that differed in whether they emphasized a subgoal that was predicted to be useful for solving novel test problems. It was predicted that a subject would be more successful at an unfamiliar part of a test problem if he or she had learned the relevant subgoal, compared with a subject who did not learn that subgoal.

The domain explored in Experiments 1 and 2 was probability. This domain was chosen because the training and testing materials can be relatively simple, and because prior work (Ross, 1987, 1989) has provided a useful manipulation.

Ross (1989; Experiment 1B) had subjects study example probability problems, such as ones involving permutations, and then solve several test problems. The mathematical roles of the entities (e.g., scientists, computers) in the examples and test problems were manipulated. For instance, Table 1 presents a permutation problem involving the determination of the probability that scientists will pick particular computers. The equation used for this example was $p = 1/[n(n-1)...(n-r+1)]$, where n is the number of choices available, and r is the number of choices being made. The test problems re-

quired some subjects to find, for instance, the probability that students will pick particular cars (humans picking inanimate objects), while other subjects had to find the probability that particular students would be assigned to particular cars (i.e., objects "picking" humans).

Corresponding mathematical roles are held by the humans and objects in the example in Table 1 and the test problem involving students picking cars. In both the example and the test problem, it is the number of humans, scientists and students, doing the choosing that provides the value for r. The number of objects, computers and cars, from which to choose provides the value for n. The second type of test problem, cars "picking" students, however, has reversed object correspondences: The humans provide the value for n (see Table 2B for another example of humans providing a value for n). Ross (1989) found that subjects were more successful at solving the first type of test problem than the second, presumably because their problem solving was guided to some degree by a feature correspondence approach. Specifically, if the number of objects provided the value of n in the examples, then subjects were likely to assign them this role in a test problem, even if it was the number of humans that should have provided this value in the test problem.

Besides working on permutation problems, Ross's subjects studied and solved combination problems. An example combination problem might ask for the probability that the seven hooks nearest the classroom door

Table 1
Permutation Training Example

The supply department at IBM has to make sure that scientists get computers. Today, they have 11 IBM computers and 8 IBM scientists requesting computers. The scientists randomly choose their computer, but do so in alphabetical order. What is the probability that the first 3 scientists, alphabetically, will get the lowest, second lowest, and third lowest serial numbers, respectively, on their computers?

Table 2
Test Problems

A. Permutation: People Picking Objects

As part of a new management policy, the Campbell Company is allowing the 20 company-owned vacation cottages to be used for vacations by their 14 plant managers. If the managers, in order of seniority, randomly choose a cottage from a list, what is the probability that the 4 managers with the most seniority get the most lavish, second most lavish, third most lavish, and fourth most lavish cottages, respectively?

B. Permutation: Objects Picking People

The secretaries at city hall are supposed to get new chairs this week. Today, city hall received 14 new chairs, and there are 11 secretaries requesting them. For inventory purposes, the property manager wants to assign the chairs in the order that they are unpacked. So, starting with the chair that is unpacked first, she randomly chooses a secretary to receive it, and continues until all the secretaries have chairs. What is the probability that the first 2 secretaries, alphabetically, will get the first and second chairs that are unpacked, respectively?

C. Combination: People Picking Objects

The Happy House Nursery School has had 17 hooks put up in the hall for the coats of their 14 students, with each student using 1 hook. The students each choose a hook at random as they come in one morning. What is the probability that the 7 tallest students get the 7 hooks closest to the classroom door? (It does not matter which of the particular 7 hooks closest to the door these students get, just as long as it is any 1 of the 7 closest.)

D. Combination: Objects Picking People

The Nashville Gnats Baseball team has a bus that has 30 seats. There are 25 players that are going on a road trip to play in a nearby town. To avoid arguments, the manager randomly chooses a player for each seat, starting with the seats in the front. What is the probability that the 6 pitchers get the 6 front seats? (It does not matter which of the particular 6 front seats the pitchers get, just as long as it is any 1 of the 6 in the front.)

Table 3
Solution Types Used for the Permutation Example in Table 1

Subgoal Solution

The equation needed for this problem is $1/[n*(n-1)*...*(n-r+1)]$. In this problem, $n = 11$ and $r = 3$. However, another way of approaching the problem is to think of it in the following way:

Probability of the first scientist (who comes first alphabetically) getting the computer with the lowest serial number = 1/11.

Probability of the second scientist getting the second lowest serial number = 1/10.

Probability of the third scientist getting the third lowest serial number = 1/9.

So, $1/11*1/10*1/9 = 1/990$ = overall probability.

Equation Solution

The equation needed for this problem is $1/n*(n-1)*...*(n-r+1)$. This equation allows one to determine the probability of the above outcome occurring. In this problem, $n = 11$ and $r = 3$. The 11 represents the number of computers that are available to be chosen, and the 3 represents the number of choices that are being focused on in this problem. The equation divides the number of ways the desired outcome could occur by the number of possible outcomes. So, inserting 11 and 3 into the equation, we find that $1/11*10*9 = 1/990$ = overall probability.

would be picked by the seven tallest students in a class (see Table 2C). The equation used to solve combination problems of this sort is $p = [h!(j-h)!]/j!$, where h is the number of entities (e.g., students) doing the choosing, and j is the number of entities in the pool from which things are chosen (e.g., hooks). Again, Ross demonstrated the object correspondence phenomenon.

Although Ross taught his subjects the procedures for solving both permutation and combination problems and examined transfer to problems in which the roles of humans and objects were switched, an examination of the two procedures shows that a more general procedure can be used to solve both problem types. Both permutation and combination problems can be analyzed by considering the individual event probabilities that contribute to an overall probability. This approach is demonstrated in the "subgoal" solution provided for the problem of the scientists and the computers (see Table 3). The combination problem, involving students and coathooks (Table 2C), can be analyzed in a similar way:

Probability that one of the seven tallest students will get a hook near the door = 7/17.
Probability that one of the remaining six tallest students will get a hook near the door = 6/16.
Probability that one of the remaining five tallest students will get a hook near the door = 5/15, etc.

So,

$$\frac{7}{17} * \frac{6}{16} * \frac{5}{15} * \frac{4}{14} * \frac{3}{13} * \frac{2}{12} * \frac{1}{11}$$

$$= \frac{7!}{17 * 16 * \dots * 11} = \text{overall probability.}$$

Combination problems have numerators that are no longer simply "1." Instead, they start at the number of acceptable choices and then are decremented just like the denominator.

The "subgoal" solution, presented in Table 3 for the scientists and computers permutation problem, is assumed to help learners form two goals. The first goal is to find the overall probability; this goal is assumed to be formed because it is explicitly stated in the example. The second is the subgoal to find each event probability—for example, the probability that the first scientist will get the computer with the lowest serial number, the probability that the second scientist will get the computer with the second lowest serial number, and so on. This subgoal is assumed to be formed because each individual event probability is explicitly labeled and spatially separate in the subgoal solution in Table 3. The method for finding an individual event probability will involve the steps of inserting a 1 in the numerator and placing the number of (remaining) objects in the denominator of each probability.

The "equation" solution, presented in Table 3 for the scientists and computers permutation problem, is assumed to help learners form only the goal to find the overall probability. The method for achieving this goal will consist of a set of steps for finding numbers from the problem statement and inserting them into the equation.

Performance Predictions

Subjects who study either the equation solution or the subgoal solution are predicted to perform well on the first test problem they are given (see Table 2A). This first problem is a permutation problem that is isomorphic to the training examples that the subjects studied, and humans and objects play the same roles that were in the examples (i.e., humans picking objects). The subjects can simply repeat the steps learned from the examples. Performance on this problem serves as a check that a subject has learned at least a set of steps.

Two of the three remaining test problems, one permutation problem and one combination problem (see Tables 2B and 2D), reverse the roles for humans and objects compared with the training examples; that is, objects are "picking" humans. What this means in terms of the solution is that the numbers that go into the denominator are based on the number of humans in the problem, not the number of objects. It is predicted that subjects will have difficulty with this aspect of the problems. However, those who study the subgoal solution (the subgoal group) are predicted to have learned the subgoal to find each individual event probability, and thus might be more likely to consider what the numerator and denominator mean in each event probability. As a result, these subjects have a better chance to consider modifying the denominator for this problem than subjects in the equation group.

The two combination test problems (see Tables 2C and 2D) are expected to cause difficulty on numerator performance since, unlike the examples' numerators, the numerators are no longer 1. Once again, the Subgoal group, by virtue of learning the subgoal to find each individual event probability, is predicted to outperform the equation group because the subgoal of finding each individual event probability might lead the Subgoal subjects to consider the numerator as a potential locus for change.

EXPERIMENT 1

In Experiment 1, I tested the hypothesis that transfer will be improved if subjects study training examples that emphasize a subgoal needed for novel problems. Besides examining problem-solving performance, another measure of subgoal learning was attempted by having subjects describe how to solve probability problems after studying the examples. If a subject's description includes a statement such as "find the probability of each event's occurrence," then this would be taken as additional support that he or she had learned that subgoal. The subgoal group should mention the subgoal of breaking the overall probability into a set of individual event probabilities more often than the equation group, since this is the major difference in the solution approaches presented to the two groups. The subjects who mention this subgoal should perform better than other subjects on the denominators for reversed-roles permutation and combination problems and on the numerators for the combination problems.

Although the emphasis so far has been placed on the differences in the solutions studied by the subgoal and equation groups, it is important to note that the solutions are similar in length and, presumably, clarity (see Table 3). One could construct a solution type that is arbitrarily unclear and demonstrate that subjects' performance on transfer problems is poor relative to subjects who study examples using a solution type that is arbitrarily clear. It is suggested that the equation solution contains potentially useful information for solving test problems, but that subjects who study that solution are likely to focus on how values are inserted into the equation rather than form a more general procedure.

Method

Subjects. The subjects were 66 students from introductory psychology classes at the Georgia Institute of Technology who participated for course credit. None of the subjects had taken a probability course prior to participating in the experiment.

Materials and Procedure. The subjects received a booklet containing training examples and test problems. All the subjects studied two isomorphic worked-out permutation example problems in which humans picked objects. Table 1 presents one of these examples.

The subjects were randomly assigned to the equation group ($n = 31$) or the subgoal group ($n = 35$). In the equation group's examples, an equation was used in order to solve the problem. The solution included an explanation of the meaning of the numbers being inserted into the equation. The subgoal group's examples divided the problem into finding each individual probability. Table 3 contains the solutions that were seen by the groups for one of the training examples.

The subjects were asked to study the examples carefully and were told that, after studying them, they would be asked to solve some problems. They were also told that they could not look at the examples when working on the problems. This restriction was intended to increase the likelihood that they would pay attention to the examples and how they were solved.

After studying the examples, the subjects were asked to describe how to solve problems in the domain. The instructions were: "Suppose you were going to teach someone how to solve probability problems of the type you have just studied. Please describe the procedure you would give someone to solve these problems. Please be as complete as possible."

After writing their descriptions, the subjects attempted to solve the four test problems in Table 2. The first problem was isomorphic to the examples (see Table 2A). The second problem (B) was a permutation problem like the examples, but humans provided the value of n (see Table 2B). The third and fourth problems (C and D) were combination problems. In the first combination problem, humans picked objects, which is the same notion shown in the examples (see Table 2C). In the second combination problem, objects picked humans (see Table 2D). Thus, the test problems represented a range of difficulty as a function of whether they involved permutations or combinations and whether humans were picking objects, or objects were picking humans.

The subjects worked at their own pace and were asked to show all their work. In general, they took about 30 min to complete the experiment. Each permutation problem was scored for whether a subject used the correct denominator. For instance, the solution to the second permutation problem is $1/11 * 1/10$. If a subject wrote $1/14 * 1/13$, confusing the roles of the chairs and secretaries, the denominator would be scored as incorrect. For combination problems, the numerator and denominator were both scored as correct or incorrect. Two raters independently scored the problems; their scores agreed 92% of the time. Any disagreements were resolved

by discussion. The frequencies with which the groups correctly found the denominators for the permutation problems and the denominators and numerators for the combination problems were analyzed by using the likelihood ratio chi-square test (G^2; Bishop, Fienberg, & Holland, 1975).

Results and Discussion

The overall performance differences between the groups can be summarized as follows: They did not reliably differ on denominator performance for the reversed-role problems (i.e., objects picking humans), but the subgoal group outperformed the equation group on the numerators for the combination problems.

As expected, both groups were quite successful in determining the denominator on the first permutation problem, which was isomorphic to the training examples and had humans and objects playing the same roles as in the examples [$G^2(1) = 0.5, p = .78$; see Table 4].

In the second permutation problem, objects picked humans—a reversal from the training examples. An error that the subjects frequently made on this problem was to use the number of chairs (14) as the starting point in the denominator rather than the number of secretaries (11). Although the groups did not significantly differ in finding the denominator [$G^2(1) = 1.11, p = .29$], there was an 11% advantage for the subgoal group (see Table 4).

The next problem was a combination problem in which humans picked objects. The subjects were fairly successful at finding the denominator and, as expected, did not differ significantly on this measure [$G^2(1) = .09, p = .76$; see Table 4]. However, they did have difficulty finding the correct numerator for this problem; many of the subjects simply used 1. As expected, the subgoal group outperformed the equation group on this measure [$G^2(1) = 6.39, p = .01$].

In the second combination problem, objects picked humans. As in the reversed-roles permutation problem, the subjects had difficulty finding the correct value for the denominator. The groups did not differ significantly on this measure [$G^2(1) = .06, p = .80$; see Table 4]. As in the first combination problem, the subjects had difficulty finding the correct numerator. As expected, the subgoal group was more successful than the equation group on this measure [$G^2(1) = 9.83, p = .002$].

Table 4
Performance (Percent Correct) on Experiment 1 Test Problems

	Group	
	Equation ($n = 31$)	Subgoal ($n = 35$)
Permutation Problem 1 (people pick objects)		
Denominator	94	89
Permutation Problem 2 (objects pick people)		
Denominator	23	34
Combination Problem 1 (people pick objects)		
Denominator	71	74
Numerator	13	40
Combination Problem 2 (objects pick people)		
Denominator	23	20
Numerator	10	43

The transfer results suggest that the subjects in both groups were equally misled by superficial role reversals of objects and humans in the denominator. However, the subgoal subjects were more likely to adapt their solution procedure to find the numerator correctly in combination problems.

Relationship between training examples and explanations produced by subjects on how to solve problems. The subjects' explanations of how to solve problems in the domain were scored for whether they mentioned the subgoal of dividing the overall probability into a series of individual probabilities. Six equation subjects and 5 subgoal subjects produced explanations that were too general or idiosyncratic to be scored. These explanations typically consisted of statements such as "I would read through the example and write it up on the board for the person." These subjects were excluded from the following analyses.

The subgoal group mentioned the notion of dividing the overall probability into a set of individual probabilities far more often than the equation group [83% vs. 8%; $G^2(1) = 35.3, p < .0001$]. This was expected, since the subgoal group studied example solutions that labeled and isolated individual probabilities, whereas the equation group did not.

Relationship between explanations and transfer performance. It was expected that the subjects who mentioned the subgoal of finding individual event probabilities in their explanations would be more likely to correctly find the denominator in the reversed-role problems and the numerator in the combination problems.

There was no difference in denominator performance for the first permutation problem for the subjects who mentioned the notion of dividing a probability into individual probabilities (the "IndProb" subjects) compared with the subjects who did not mention this notion in their explanations (the "OneProb" subjects) [$G^2(1) = .18, p = .67$; see Table 5]. This is not surprising, since in this problem humans choose objects, as in the examples.

As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct value for the denominator in the second permutation problem, which involved reversed roles for humans and objects [$G^2(1) = 3.84, p = .05$; see Table 5].

In the first combination problem, humans and objects played the same roles that were in the examples, and therefore it is not surprising to find that there was no difference between the groups in finding the correct value for the denominator in this problem [$G^2(1) = .15, p = .69$; see Table 5]. As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct value for the numerator [$G^2(1) = 6.26, p = .01$; see Table 5].

Although the anticipated difference between the groups in finding the correct value for the denominator in the second combination problem—a reversed-roles problem—was not found [$G^2(1) = 2.18, p = .14$], the IndProb group had a 15% advantage (see Table 5). As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct numerator value [$G^2(1) = 8.29, p = .004$; see Table 5].

These results follow the trend of those that were found when the instructional groups were compared on the transfer problems. This makes sense, since the subgoal subjects were by far the ones most likely to mention the notion of dividing an overall probability into individual probabilities.

It could be argued that learners who write "better" descriptions (e.g., mention the subgoal of finding individual event probabilities) are also the ones who are better at transfer. One defense against this argument is to note that the experimental manipulation of type of examples studied influenced transfer success.

Although the subgoal subjects did not reliably outperform the equation subjects on the denominators for reversed-role problems, they were clearly superior at finding the correct numerators for the combination problems. In addition, the IndProb group outperformed the OneProb group on finding the denominator for one of the reversed-role problems. Although most of the results support the predictions, the unreliability of the denominator effect for the reversed-role problems suggested that a second, more focused experiment would be appropriate.

EXPERIMENT 2

The procedure and materials for Experiment 2 were identical to those of Experiment 1, except for three features: (1) the subjects studied three rather than two training examples, (2) the subjects were not asked to write explanations of how to solve problems, and (3) for half of the subjects, the combination problems did *not* contain the last sentence shown for each combination problem in Table 2.

The number of examples presented to the subjects was increased to improve the likelihood that they would learn the procedures demonstrated in the examples and perhaps improve transfer to the test problems. They were

Table 5
Performance (Percent Correct) on Test Problems
as a Function of Subjects' Explanations in Experiment 1

	Group	
	OneProb* ($n = 27$)	IndProb† ($n = 28$)
Permutation Problem 1		
Denominator	89	93
Permutation Problem 2		
Denominator	14	37
Combination Problem 1		
Denominator	78	74
Numerator	14	44
Combination Problem 2		
Denominator	11	26
Numerator	11	44

*Does not mention breaking problem into individual probabilities.
†Mentions breaking problem into individual probabilities.

not asked to write explanations, since this task was unusual and may have influenced learning in unanticipated ways. Finally, the last sentence for each combination problem in Table 2 was removed for half of the subjects in order to examine whether they signaled to the subjects that these problems were different from the training examples and needed to be approached differently. Perhaps with this cue removed, the subgoal subjects would not outperform the equation subjects.

Method

Subjects. The subjects were 78 students from introductory psychology classes at the Georgia Institute of Technology who participated in the experiment for course credit. None of the subjects had taken a probability course prior to participating in the experiment.

Materials and Procedure. The subjects received a booklet containing training examples and test problems. They all studied three isomorphic worked-out permutation example problems in which humans picked objects (including the two used in Experiment 1).

The subjects were randomly assigned to the equation group ($n = 40$) or the subgoal group ($n = 38$). After studying the examples, they solved the four test problems in Table 2. The subjects worked at their own pace and were asked to show all their work. In general, they took about 25 min to complete the experiment. Two raters independently scored the problems, and their scores agreed 90% of the time. Any disagreements were resolved by discussion.

Results and Discussion

The results were similar to those from Experiment 1 and supported most of the predictions. The subgoal group strongly outperformed the equation group on the numerators for the combination problems and showed a trend toward superior performance on the denominator for one of the two reversed-role problems.

As expected, both of the groups did well in determining the denominator for the first permutation problem, and there was no significant difference in their performance [$G^2(1) = 0.63, p = .43$; see Table 6]. As in Experiment 1, both of the groups showed inferior performance in determining the denominator for the second permutation problem—a problem with reversed roles for humans and objects. As predicted, there was an advantage (20%) for the subgoal group, although this difference just missed significance [$G^2(1) = 3.47, p = .06$; see Table 6].

Performance on the combination problems was initially broken down as a function of training group and whether or not the problems contained the last sentence presented for each combination problem in Table 2. There was no main effect of sentence for either problem, nor was there an interaction between group and presence of the sentence; thus, the analyses and the results for the combination problems are collapsed across this factor in Table 6.

In the first combination problem, humans picked objects, and, as anticipated, the subjects were fairly, and equally, successful at finding the denominator [$G^2(1) = .009, p = .92$; see Table 6]. As expected, the subgoal

Table 6
Performance (Percent Correct) on Experiment 2 Test Problems

	Group	
	Equation ($n = 40$)	Subgoal ($n = 38$)
Permutation Problem 1		
Denominator	90	95
Permutation Problem 2		
Denominator	22	42
Combination Problem 1		
Denominator	85	84
Numerator	0	29
Combination Problem 2		
Denominator	30	32
Numerator	0	32

subjects outperformed the equation subjects at finding the correct numerator [$G^2(1) = 17.7, p = .0001$; see Table 6].

The predicted superior performance by the subgoal subjects did not occur for finding the correct value for the denominator in the second combination problem, a reversed-roles problem [$G^2(1) = 0.02, p = .88$; see Table 6]. As expected, the subgoal subjects outperformed the equation subjects at finding the correct numerator [$G^2(1) = 19.6, p = .0001$; see Table 6].

It was hypothesized that the subgoal subjects learned the subgoal to find individual event probabilities, but the equation subjects did not. This subgoal was predicted to aid performance in finding the denominator for reversed-role problems and the numerators for combination problems. Both experiments clearly demonstrated the numerator effect. Experiment 1 demonstrated a trend toward the denominator effect for the reversed-role problems, primarily when the subjects were partitioned into those who mentioned, or failed to mention, the subgoal of finding individual event probabilities in their explanations. This trend was stronger for the permutation problem than for the combination problem. Experiment 2 also showed this effect more strongly for the permutation problem. It is not clear if there are certain features of these two problems that differentially affected denominator performance. The subjects' explanations of how to solve problems were consistent with the claim that the subgoal subjects were more likely to learn the subgoal to find individual event probabilities than the equation subjects.

The fact that the denominator effect was less reliable than the numerator effect may be due to the role of superficial features. The training examples presumably led the subjects to expect objects to provide the value for the denominator, and perhaps this expectation tended to override any benefits due to learning subgoals, especially since a value for objects was always provided in each problem. The tendency to put a 1 as the numerator may have been more easily overridden by subgoal learning because its connection with either humans or objects is less clear.

EXPERIMENT 3

Experiment 3 was an attempt to generalize the findings from the first two experiments to another domain: algebra word problems. This domain was chosen because prior work had demonstrated poor transfer from training to transfer problems despite attempts to improve examples (e.g., Reed et al., 1985). Consider the algebra example in Table 7A, in which one has to determine how long it would take someone to do a job given that certain information about their work rate and time and another person's work rate and time are provided. This problem involves using an equation for determining work that requires representing each worker's work rate and time: $(rate_1 \times time_1) + (rate_2 \times time_2) = 1$.

Learners are good at memorizing how to solve problems that are isomorphic to the one in Table 7A. In this problem, both of the workers' rates are represented as constants. The time spent working by Worker 1 is represented as a variable, and Worker 2's time is represented as a function of that variable. However, learners may not encode the example solution by determining a representation for each rate and time and then inserting these representations into the equation. Instead, they have a more superficial understanding of the solution procedure, which involves matching the form used in the example, finding similar values in the problem statement, and inserting them into the equation. As a result, if a new problem requires a different representation of the rates and times, these learners may be unable to solve the problem. That is, the learners may not have learned that certain subgoals exist—the subgoals of representing each worker's rate and time—and that these subgoals might be achieved differently (i.e., different ways of representing rate and time depending on the givens in the problem) from the way they were achieved in the example.

For instance, the problem in Table 7B requires that Worker 1's rate be represented as a variable. In addition, instead of having the workers' times be represented as a variable and a function of that variable, the times are represented as a constant and a function of that constant. Nevertheless, the new representations can be inserted into the same equation that was used for the example in Table 7A. Similarly, the problem in Table 7C requires that Worker 1's rate be represented as a variable, and Worker 2's rate be a function of that variable. Their times are both represented as constants. These representations are different from those used in the examples.

It is hypothesized that if the representations for rates and times are highlighted separately from the equation in the example solutions, then learners will be more likely to learn that rate and time are individual representations that must be determined for each worker. In addition, it is hypothesized that if subjects learn the subgoals of representing workers' rates and times, then they will be more likely to correctly solve a novel problem requiring novel representations for rates and times than would subjects who do not learn those subgoals.

Table 7
Example and Test Work Problems in Experiment 3

A. Mary can rebuild a carburetor in 3 hours and Mike can rebuild one in 4 hours. How long would it take Mary to rebuild a carburetor if she and Mike work together, but Mike works for $\frac{1}{2}$ hour more than Mary?

Solution
    $1/3$ = Mary's rate
    t = time Mary spent rebuilding carburetor
    $1/4$ = Mike's rate
    $t + 1/2$ = time Mike spent rebuilding carburetor
    $(1/3*t) + [1/4*(t + 1/2)] = 1$
    $7/12*t = 1 - 1/8$
    $t = 7/8*12/7 = 3/2$ hours = time Mary spent rebuilding carburetor

B. Mr. Jones can refinish a dresser in 5 hours. After working for 2 hours, he is joined by his wife. Together they finish the job in 1 hour. How much of the job could his wife do in 1 hour when working alone?

Solution [not seen by subjects]
    $[1/5*(2+1)] + (w*1) = 1$
    $3/5 + w = 1$
    $w = 2/5$ = wife's rate
    so, in 1 hour, wife could do 2/5 of job

C. Barbara and Connie can finish a job in 6 hours when they work together. Barbara works twice as fast as Connie. How much of the job could Connie do in 1 hour when working alone?

Solution [not seen by subjects]
    $(2c*6) + (c*6) = 1$
    $12c + 6c = 1$
    $18c = 1$
    $c = 1/18$ = Connie's rate
    so, in 1 hour, Connie could do 1/18 of job

Subjects who learn the subgoal to represent each worker's rate should be more successful at representing the rate as a variable in the problem in Table 7B and rate as a variable and rate as a function of a variable in the problem in Table 7C. Subjects who learn the subgoal to represent each worker's time should be more successful at representing the time as a constant and a function of a constant in the second problem (B) and as a constant in the third problem (C).

Method

Subjects. Sixty-two students from introductory psychology classes at the Georgia Institute of Technology participated for course credit.

Materials and Procedure. The subjects studied three isomorphic example word problems dealing with work, including the example in Table 7A. The "rate and time label" (RTL) group ($n = 21$) studied examples in which the representations for rates and times were presented separately from the equation (see Lines 1–4 in the solution to the example in Table 7A). The "time label" (TL) group ($n = 20$) studied examples that presented the representations for each worker's time (i.e., Lines 2 and 4 from the example in Table 7A). The "rate label" (RL) group ($n = 21$) studied examples that presented the representations for each worker's rate (i.e., Lines 1 and 3 from the example in Table 7A).

After studying the examples, the subjects received three problems to solve. One was isomorphic to the training examples, and the other two involved new and old ways of representing rate and/or time (see Tables 7B and 7C).

Results and Discussion

All of the groups performed well at representing rate and time in Problem 1, which was isomorphic to the

Table 8
Performance (Percent Correct) on Test Problems in Experiment 3

	Group		
	RTL ($n = 21$)	TL ($n = 20$)	RL ($n = 21$)
Problem 1			
Rate Worker 1 (constant)	100	100	100
Rate Worker 2 (constant)	100	100	95
Time Worker 1 (variable)	100	100	100
Time Worker 2 (function of a variable)	95	95	100
Problem 2			
Rate Worker 1 (constant)	100	95	100
Rate Worker 2 (variable)	86	55	86
Time Worker 1 (function of a constant)	81	90	57
Time Worker 2 (constant)	90	90	57
Problem 3			
Rate Worker 1 (variable)	90	60	90
Rate Worker 2 (function of a variable)	86	50	86
Time Worker 1 (constant)	90	100	62
Time Worker 2 (constant)	90	100	62

Note—RTL, rate and time label; TL, time label; RL, rate label.

training examples (see Table 8). This is not surprising, since the subjects could match the representations from the training examples and simply plug in the new values.

In Problem 2 (Table 8), almost all of the subjects correctly represented the rate as a constant for Worker 1. Again, this is reasonable since this was the representation used in the training examples. The RTL and RL groups were significantly more successful than the TL group at representing the rate for Worker 2 as a variable [$G^2(2) = 6.63, p = .04$]. The RTL and TL groups were more successful than the RL group at representing the rate for Worker 2 as a constant [$G^2(2) = 8.78, p = .01$] and Worker 1 as a function of that constant [$G^2(2) = 6.50, p = .04$].

In Problem 3 (Table 8), the RTL and RL groups were significantly more successful than the TL group at representing the rate for Worker 1 as a variable [$G^2(2) = 7.59, p = .02$] and Worker 2 as a function of that variable [$G^2(2) = 8.63, p = .01$]. The RTL and TL groups were more successful than the RL group at representing the rates for Workers 1 and 2 as constants [$G^2(2) = 13.66, p = .001$, in both cases]. Across the problems, the most common errors were that the subjects either left out the representation for time or rate in the equation, or wrote that not enough information was given in the problem.

The results from this experiment are consistent with the hypothesis that subjects who learned the subgoals of representing workers' rates and times would represent them more successfully on novel problems. However, alternate explanations exist. One is that the highlighting manipulation essentially provided labels for the variables, and thus made them more meaningful. This "meaningfulness" helped the subjects to properly use the variables in the equation for the novel transfer problems. The results from Experiment 3, taken alone, are not sufficient to discriminate between this explanation and the subgoal

explanation. However, the pattern of results across the three experiments is consistent with a subgoal explanation. It is not clear how the "subgoal" solution in the first two experiments made the variables more meaningful.

One could consider the subgoal explanation of the results as being an attempt to make the notion of meaningfulness more precise. Subgoals can be viewed as a way of making a solution procedure more meaningful by providing guideposts that the solver must reach en route to achieving the overall solution to the problem. In the algebra experiment, the guideposts were to explicitly represent the rate and time for each worker. Even if these guideposts are somewhat arbitrary, at least from the point of view of the solver, they still provide organization and guidance that may make the solver less likely to stray from the correct solution path (cf. Mawer & Sweller, 1982).

It is worth noting that performance on the test problems was generally quite good. This suggests that even for relatively sophisticated subjects (most of the students in the experiment had at least one term of college calculus), examples that more effectively convey subgoals can improve transfer performance.

GENERAL DISCUSSION

The aim of the present study was to examine whether examples that teach a subgoal structure for solving problems in a domain could be created and whether learning these subgoals would help subjects solve problems that required novel methods for them. The numerator performance results and some of the denominator performance results in the probability experiments, and the rate and time results in the algebra experiment, suggest that subgoals can be conveyed to learners through examples, and that learning these subgoals helps people achieve them in novel problems. This is quite encouraging in comparison with the usual finding in the problem-solving literature, which has shown poor transfer from training materials to test problems that require more than a simple repetition of a set of memorized steps (e.g., Reed et al., 1985; Ross, 1987, 1989). These findings suggest that examples that emphasize a useful subgoal structure can help turn learners into the "good" learners observed by Chi et al. (1989), who tended to find meaning, such as goals, for the mathematical steps in the examples and who made use of this information when solving novel problems.

Mayer and Greeno (1972) experimentally manipulated the meaningfulness of instruction in solving binomial probability problems by varying whether the instruction focused on mechanical operations or on concepts that were presumed to be part of subjects' prior knowledge. They found that the "mechanical" group was more successful at solving familiar problems, whereas the "concept" group was better at answering "understanding" questions about the domain, such as whether the number of successes could be greater than the number trials. The present study extends Mayer and

Greeno's findings by showing that learning subgoals promotes transfer to novel problems while also helping learners exhibit some level of understanding, as shown by the explanations produced by the subgoal subjects in Experiment 1.

CONCLUSIONS

The present results are consistent with the claim that learning subgoals will help learners determine which parts of a solution procedure need to be modified in order to solve novel problems. The results also provide a starting point for determining how to "emphasize" a subgoal in an example. The approach used here was to create example solutions that isolated components of the procedure that could be construed as subgoals. The visual separation and the labeling, and perhaps their interaction, may have all played a role in subgoal learning.

Smith and Goodman (1984) examined subgoal learning by comparing a group of subjects who followed a set of steps for assembling an electric circuit with a group who received a structurally oriented "explanatory schema" with the steps. This schema consisted of statements that provided a rationale for carrying out sets of steps. Each rationale was essentially a statement of a goal that the steps were achieving (e.g., "The next thing that you will have to do is to assemble the on–off switch"). When assembling a new circuit, the subjects who had previously received the explanatory schema were more accurate at building the substructures corresponding to the goals, even though the required steps were not identical to the ones followed during training.

The present results, and those from Smith and Goodman (1984), suggest that research aimed at determining factors that affect subgoal learning, such as the use of labels and visual separation of steps in examples, would have clear pedagogical benefits. In addition, it is important, both in terms of theory development and the production of effective training materials, to explore what constitutes a good subgoal structure for a given domain. Perhaps a theory-motivated technique can eventually be developed for determining effective subgoal structures for any given domain.

REFERENCES

ANDERSON, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

ANDERSON, J. R., BOYLE, C. F., FARRELL, R., & REISER, B. J. (1987). Cognitive principles in the design of computer tutors. In P. Morris (Ed.), *Modeling cognition.* New York: Wiley.

ANZAI, Y., & SIMON, H. A. (1979). The theory of learning by doing. *Psychological Review,* 86, 124-140.

BISHOP, Y. M. M., FIENBERG, S. E., & HOLLAND, P. W. (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.

CATRAMBONE, R., & HOLYOAK, K. J. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition,* 18, 593-603.

CHI, M. T. H., BASSOK, M., LEWIS, R., REIMANN, P., & GLASER, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science,* 13, 145-182.

CHI, M. T. H., & VANLEHN, K. (1991). The content of physics self-explanations. *Journal of the Learning Sciences,* 1, 69-106.

DIXON, P. (1987). The processing of organizational and component step information in written directions. *Journal of Memory & Language,* 26, 24-35.

EYLON, B., & REIF, F. (1984). Effects of knowledge organization on task performance. *Cognition & Instruction,* 1, 5-44.

FONG, G. T., KRANTZ, D. H., & NISBETT, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology,* 18, 253-292.

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

KIERAS, D. E., & BOVAIR, S. (1986). The acquisition of procedures from text: A production system analysis of transfer of training. *Journal of Memory & Language,* 25, 507-524.

LAIRD, J. E., NEWELL, A., & ROSENBLOOM, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence,* 33, 1-64.

LARKIN, J., McDERMOTT, J., SIMON, D. P., & SIMON, H. A. (1980). Expert and novice performance in solving physics problems. *Science,* 208, 1335-1342.

LeFEVRE, J., & DIXON, P. (1986). Do written instructions need examples? *Cognition & Instruction,* 3, 1-30.

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

MAWER, R. F., & SWELLER, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 8, 252-259.

MAYER, R. E., & GREENO, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology,* 63, 165-173.

PIROLLI, P. L., & ANDERSON, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology,* 39, 240-272.

REDER, L. M., CHARNEY, D. H., & MORGAN, K. I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition,* 14, 64-78.

REED, S. K., ACKINCLOSE, C. C., & VOSS, A. A. (1990). Selecting analogous problems: Similarity versus inclusiveness. *Memory & Cognition,* 18, 83-98.

REED, S. K., & BOLSTAD, C. A. (1991). Use of examples and procedures in problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 17, 753-766.

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.

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

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

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

SIMON, H. A., & REED, S. K. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology,* 8, 86-97.

SMITH, E. E., & GOODMAN, L. (1984). Understanding written instructions: The role of an explanatory schema. *Cognition & Instruction,* 1, 359-396.

WARD, M., & SWELLER, J. (1990). Structuring effective worked examples. *Cognition & Instruction,* 7, 1-39.

ZHU, X., & SIMON, H. A. (1987). Learning mathematics from examples and by doing. *Cognition & Instruction,* 4, 137-166.