Memory & Cognition
1990, 18 (6), 593-603

Learning subgoals and methods for solving probability problems

RICHARD CATRAMBONE
Georgia Institute of Technology, Atlanta, Georgia

and

KEITH J. HOLYOAK
University of California, Los Angeles, California

We hypothesize that typical example problems used in quantitative domains such as algebra and probability can be represented in terms of subgoals and methods that these problems teach learners. The "quality" of these subgoals and methods can vary, depending on the features of the examples. In addition, the likelihood of these subgoals' being recognized in novel problems and the likelihood of learners' being able to modify an old method for a new problem may be functions of the training examples learners study. In Experiment 1, subjects who studied examples predicted to teach certain subgoals were often able to recognize those subgoals in nonisomorphic transfer problems. Subjects who studied examples demonstrating two methods rather than one exhibited no advantages in transfer. Experiment 2 demonstrated that if the conditions for applying a method are highlighted in examples, learners are more likely to appropriately adapt that method in a novel problem, perhaps because they recognize that the conditions do not fully match those required for any of the old methods. Overall, the results indicate that the subgoal/method representational scheme may be useful in predicting transfer performance.

A consistent finding in the problem-solving literature is that learners are often unable to make appropriate use of prior information to solve new problems if the new problems differ from training examples in more than minor ways (e.g., Anderson, Farrell, & Sauers, 1984; Gick & Holyoak, 1980, 1983; Reed, Dempster, & Ettinger, 1985; Spencer & Weisberg, 1986). Learners often seem to acquire primarily superficial knowledge from training examples, consisting of little more than a series of memorized steps. For instance, some of the examples used in the present experiments involve finding the average number of times an event occurs per trial (such as the average number of errors per game made by the Detroit Tigers' infield) given the frequencies of the various *subevents* (e.g., the number of games in which the infield made 0 errors, the number of games in which they made exactly 1 error, and so on). Learners are good at memorizing, from examples, the steps of multiplying each subevent (0, 1, etc.) by its observed frequency, summing the results, and dividing by the number of trials in order to achieve the subgoal of finding the average frequency. However, they rarely notice that the steps of multiplying each subevent by its frequency and then summing them could be viewed as a way of finding the total frequency of the event. As a result, these learners are usually unable to find the average frequency of an event when a problem provides the total frequency of that event directly, rather than requiring that it be derived from the frequencies of the various subevents.

In the present paper, we will argue that the knowledge people gain from examples in domains that emphasize solving problems, such as probability, can be represented by subgoals and methods. Furthermore, manipulations of examples can affect which subgoals and methods a learner acquires, thus influencing his or her performance on novel problems. A subgoal is an unknown entity (numerical or conceptual) that needs to be found in order to achieve a higher level (sub)goal of a problem. A method is a series of steps for achieving a particular subgoal.

Identifying Subgoals and Methods in a Domain

The information that is considered to constitute subgoals and methods in the work presented here was identified in a task analysis of the examples and problems used in the experiments. There presumably do not exist platonic subgoals and methods for a domain. Rather, the problem-solving knowledge can be carved up in different ways. Our choice for representing the domain was based on our own understanding of the domain, plus the guidance of some experts. Solutions to the examples and test problems were examined, and the parts of the solution that seemed to be directed toward a single conceptual goal were identified as a method for that goal.

We assume that learners will form a minimal representation of an example solution (e.g., Singley & Anderson,

This research was supported by NSF Grant BNS 86-15316. Experiment 1 was reported at the Ninth Annual Meeting of the Cognitive Science Society, Seattle, July 1987. Please address all correspondence to Richard Catrambone, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332.

1989, p. 275). Thus, if an example can be processed in terms of a series of steps without the overhead of hierarchical organization, then, we assume, that is how it will be processed. The learner would acquire one *large* goal, to predict the probability of some event, and one large method for finding that goal, a series of mathematical steps. We assume that certain manipulations of examples can lead to a different organization that explicitly represents *smaller* subgoals, such as finding the total frequency of an event, and their associated methods, as well as the larger goal. Two of those manipulations are explored in the experiments presented here: learning multiple methods for finding a subgoal, and receiving elaborations of example methods.

In the present study, we will focus on differential knowledge conveyed by examples, rather than on individual differences in what people learn from examples. Although individual differences account for some performance differences (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989), the quality of training examples can also affect transfer performance. Thus, this research is aimed at determining whether learners' knowledge and features of training examples and target problems can be represented in terms of subgoals and methods, and whether this representation can be used to predict learners' performance. In addition, we will propose learning constraints that generate predictions regarding which subgoals and methods are taught by particular examples.

We hypothesize that it is more effective, with respect to transfer to a novel problem, for a learner to have his or her problem-solving knowledge represented in terms of relatively small subgoals and methods. A small subgoal is one that is achieved, through the use of an explicitly represented method, on the way to the achievement of some larger subgoal. We argue that problem-solving knowledge represented by a few large subgoals, which are achieved by large, linear methods. has less flexibility, because potentially useful subgoals are not represented explicitly. If one of these subgoals is needed to solve a new problem, a learner who has acquired a larger unitary method would be less likely to solve the problem than would a learner who has acquired smaller, decomposed methods and subgoals.

For example, with respect to the method discussed earlier for finding the average number of errors made by the Detroit Tigers' infield, it would be better if that method contained an explicit representation of the subgoal of finding the total number of errors made by the infield, rather than only a series of steps to achieve the higher level subgoal of finding the average number of errors. The additional subgoal of finding the total frequency would be useful for a novel problem that provided total frequency directly and did not provide the frequencies of the various subevents (i.e., games with 0 errors, 1 error, etc.). If the learner possessed this additional subgoal, he or she might be more likely to find the average frequency correctly than would a subject who only knew a single linear series of steps for finding an average frequency. We will argue that examples can strongly influence the type of subgoals and methods formed by a learner.

Transfer to Novel Problems

A novel problem could be novel for several reasons involving the transfer of subgoals and methods. A problem might involve a new subgoal; it might involve old subgoals but require that a modification of an old method be made so that one (or more) of the subgoals can be achieved; or it might require the use of a completely new method for an old subgoal to be achieved. To *modify* a method is to delete or change one or more of its steps in order to adjust to the givens in a problem.

An analysis of subgoals and methods provides a potentially useful way of representing the knowledge that learners acquire from examples, the knowledge that is available for modification, and the knowledge that is needed to solve a new problem. To the extent that manipulations designed to alter the subgoals and methods acquired by subjects are successful, this distinction will gain psychological validity. The success of the manipulations will be measured by the accuracy of the predictions of subjects' performance on test problems that are designed to contain variations on the subgoals and methods presumably learned from training examples.

It is worth it to highlight briefly how this approach differs from other work on transfer that also involves subgoals and methods. Singley and Anderson (1989) use a production rule formalism to predict transfer from training to test problems. The degree of overlap in the productions learned during training with the productions needed to solve a test problem provides a prediction of the amount of expected transfer. This approach has been successful, to varying degrees, in domains such as learning calculus, LISP, and text editing (see also Bovair, Kieras, & Polson, 1990). The empirical emphasis in Bovair et al.'s (1990) work has been on training manipulations introduced by intelligent tutors and degrees of practice. The current paper is focused more on learning through manipulations of the examples themselves. From a theoretical perspective, our goal is not so much to examine the detailed (i.e., production rule) overlap between an acquired procedure and a target procedure, but to examine whether conceptual units such as subgoals and methods provide a useful way to characterize problem-solving knowledge.

VanLehn (1982) has examined how children carve up the rules for subtraction as a function of the training problems they solve. He has demonstrated that a given sequence of training problems can lead to a representation of the subtraction procedure that has certain omissions. Using the representation presumably acquired by children after a certain lesson sequence, VanLehn could successfully account for most of the errors made by children as they attempted to solve novel subtraction problems that required knowledge beyond what they already possessed. In the present paper, we describe our attempts to manipulate the representations formed by adult learners from examples and to begin to develop a set of rules for predicting the conditions under which a subgoal and/or method will be learned.

Two Learning Assumptions

We make two assumptions concerning what a person learns from an example. The first assumption is that if a subgoal is labeled or clearly highlighted in an example's solution, then this subgoal will also be explicitly represented in the learner's knowledge base. Suppose a learner studies the example in Appendix A. After studying this example, the learner is predicted to have learned three subgoals: to find the average for some event, to find the expected probability of a subevent, and to find the expected frequency of a subevent.

The second assumption is that once a learner acquires a particular subgoal and method in the context of an example, that method can be used to achieve the subgoal in a new problem even if the context of the example and problem are different (i.e., even if superficial story line and the other subgoals and methods in the problem are different from those in the example). This assumption is based on the work of Singley and Anderson (1988, 1989) and of Foltz, Davies, Polson, and Kieras (1988), which suggests that the transfer of procedures or parts of procedures from one task to another can occur successfully, at least within the domain of learning text-editing skills. Also, it is likely that relatively experienced problem solvers, such as the students in our experiments, have learned that the specific story line is typically irrelevant to the procedure illustrated in an example (see also Schoenfeld, 1985).

Given the above learning constraints, a learner who studies the example in Appendix A would possess a method for finding the average frequency of EVENT (EVENT-AVG) that is *not* further broken down into a subgoal for finding the total frequency of EVENT. (This assumption is supported by some pilot work reported in Catrambone & Holyoak, 1987.) Rather, the learner learns to apply a series of mathematical operations without any meaning given to them other than the subgoal that they as a group enable the learner to achieve. The subgoal of finding the total frequency of the EVENT was not highlighted or labeled in the example solution, and therefore it is assumed not to be represented explicitly in the knowledge base.

On the basis of the analysis above, we would predict that a learner would have difficulty solving the novel problem in Appendix B, which asks for the expected frequency of some subevent but only provides the total frequency of EVENT and not the frequencies of the various subevents. A learner possessing the subgoals and method described above would be predicted to recognize the final subgoal of the problem (find $P[X = 2]$) as well as the intermediate subgoal of finding the average frequency of EVENT, but would be expected to have difficulty achieving this subgoal since the steps he or she has learned for achieving it could no longer be followed (i.e., because the observed frequencies of the various subevents are not given).

To solve the problem in Appendix B, the learner would need to understand that some of the steps in the method for finding EVENT-AVG involved finding the total EVENT frequency. Of course, even if the learner realized this, he or she would still have to recognize that this value (total EVENT frequency) is given directly in the problem statement. However, since the learner would at least possess the needed subgoal of finding total EVENT frequency, he or she might have a better chance at solving this problem than would someone who does not have this subgoal.

Learning and Transfer from Examples

Although learners can be given general procedures for solving problems in a domain, it is usually the case that learners prefer to study or refer to examples. This preference is manifested in the usefulness of allowing learners to have an example in front of them while working on a new problem (Pirolli & Anderson, 1985), in their explicit mentioning of examples when solving a problem (e.g., Chi et al., 1989; Lancaster & Kolodner, 1988; Ross, 1984), and in their following of examples rather than instructions when the two conflict (LeFevre & Dixon, 1986). Thus, the preference for using examples seems quite strong.

Prior research on transfer from training examples to target problems has often been focused on the effects of the relationships between superficial features of examples and problems (Ross, 1987, 1989a) or between the form of equations used in examples versus problems (Reed & Actor, 1989; Reed et al., 1985). This work has indicated that if a worked-out example and a target problem have similar objects (such as both problems' involving vehicles and car mechanics), the learner will try to give similar objects corresponding roles in the new problem and in the example, even when the problem structure implies different roles (Ross, 1987, 1989a). Similarly, learners will often try to repeat the same steps in a new problem by taking numbers from the problem statement and putting them into the equation without full regard for their "meaning" (Reed et al., 1985). These studies support the notion that learners rely heavily on examples and tend to learn superficially from them; however, such results do not provide guidance for how to design examples to promote more meaningful learning. The subgoal/method distinction provides a way of predicting the places in a novel problem where learners should succeed and have difficulty, and it may suggest ways of designing examples to increase meaningful learning—that is, to influence the number and type of subgoals and methods learned.

Overview of Experiments

In the present experiments, we investigated the usefulness of the subgoal/method distinction by examining the effects of different manipulations of training examples on whether learners acquire particular subgoals and methods, and whether they can modify old methods for new problems. The manipulations involved varying the superficial similarity of the examples, the methods and subgoals in the examples, and the presence of explanatory text in which we attempted to highlight the subgoals and methods for the learner. After they had studied examples, the subjects were given target problems to solve. Their performance on the problems was used to determine which subgoals and methods they seemed to possess, as well as their degree of success in modifying and constructing methods.

In these experiments, the participants were paid volunteers from undergraduate probability courses. In these courses, students learn counting rules (e.g., ordered and unordered sampling) and are then introduced to the notion of a random variable. The students next learn about certain basic discrete probability distributions such as the binomial, Poisson, and geometric. The binomial distribution is typically introduced first, followed by the Poisson distribution. The present experiments dealt with the Poisson distribution, and the students participated in them just before learning the Poisson in their classes.

The Poisson distribution is often used to approximate binomial probabilities for events that occur with some small probability p. The Poisson equation is

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!},$$

where λ is the expected value (the average) of the random variable X.

EXPERIMENT 1

Although prior work has shown that learners are heavily influenced by superficial feature overlap between examples and target problems (e.g., Ross, 1984, 1987, 1989a), there is little evidence that manipulations of superficial features among training examples affect transfer to nonisomorphic problems (e.g., Gick & Holyoak, 1983). However, there may be a systematic effect due to manipulations of structural features. In Experiment 1, we examined whether learners who study examples that demonstrate two methods for finding a subgoal would be more likely to recognize that subgoal in a new problem and to adapt an old method to achieve it, relative to learners who only study examples demonstrating one method. In addition, we tested the two learning constraints proposed earlier by examining whether the subjects appeared to have learned the predicted subgoals.

The sociologist example presented in Appendix A deals with an event occurring randomly in time. The Poisson distribution can also be used to model events occurring randomly in space. For instance, one could reasonably fit a Poisson distribution to the number of fossils found in each section of a partitioned quarry. This problem can be solved using the same steps as the sociologist example.

Clearly, a person could learn to solve problems of this type by memorizing the series of operations without forming all the "correct" subgoals. This could occur because learners usually focus on more superficial features of problems and on the operations used to achieve an end goal, for these aspects are easier to identify than the underlying subgoals and methods (see, e.g., Ross, 1989b; Schoenfeld & Herrmann, 1982). Students have considerable experience with the typical objects, such as decks of cards and blocks of wood, that populate the world of quantitative problems in domains such as probability and mechanics. It is therefore not surprising that these students would tend to focus their attention on features, and on mathematical operations involving those features, with which they are most familiar (Greeno, Riley, & Gelman, 1984; Hayes, Waterman, & Robinson, 1977).

There were three groups in Experiment 1. The superficial variety (SV) groups studied four superficially different, worked-out examples isomorphic to the sociologist and Detroit Tigers problems (the SV-T group) or the "cookie" problem (the SV-C group; see Appendix C, top). The procedural variety (PV) group studied two examples isomorphic to the sociologist problem and two examples isomorphic to the cookie problem. The first question of interest was whether subjects would learn the subgoals they were predicted to learn on the basis of the first learning constraint—that is, whether subjects would learn the subgoals of EVENT-AVG, expected probability, and expected frequency. The second question was how well subjects in the SV and PV groups could modify an old method to find λ (EVENT-AVG) in novel problems.

Four target problems were used in Experiment 1. The first was an isomorph to the sociologist example. This problem dealt with the number of errors made per game by the Detroit Tigers' infield over the course of a season. The problem should have been difficult for the SV-C subjects, since they were not predicted to have formed some of the necessary subgoals and methods during training (see below for the method for solving the cookie problem).

The second target problem was the cookie problem. Its final subgoal is to find EVENT-AVG. However, the method for achieving this subgoal is substantially different from the method used in the sociologist example. Instead of multiplying and summing a series of observed frequencies and dividing by trials, the average frequency is found by inserting the given value for $P(X = 0)$ (the desired probability of finding zero raisins in a cookie) into the Poisson equation and solving for λ. The SV-T group was expected to have great difficulty with this problem, because the method for finding λ in the cookie problem was not a straightforward modification of the method they learned during training. Nevertheless, the assumption that learned subgoals can be transferred would predict that these subjects would demonstrate that they were trying to find λ.

The third target problem is presented in Appendix C, (bottom). This "birthday" problem requires that the learner realize that λ can be calculated by dividing the number of people (total frequency of EVENT) by the number of days (total trials). Then λ can be used in the Poisson equation to find $P(X = 3)$.

The fourth target problem was the "football" problem (see Appendix B), which is isomorphic to the birthday problem. The football problem was expected to be easier than the birthday problem, since the latter could lead subjects to think that there was something special about January 1 instead of realizing that the problem could apply to any day of the year. Thus, subjects could be confused about what events and trials were in the birthday problem. The football problem was more straightforward and should have produced performance more closely tied to subjects' ability at modifying old methods, rather than their ability to identify events and trials.

Method

Subjects. Fifty students from a probability class were recruited and were paid for their participation.

Materials and Procedure. The subjects in Experiments 1 and 2 were run in groups of anywhere from 1 to 10 people. The subjects were first given a booklet to study. The cover sheet contained a description of the relationship between the binomial and Poisson distributions and provided the Poisson equation. The next four pages contained the four worked-out Poisson distribution examples. The subjects were told to study the examples carefully, since after studying them they would be asked to solve four problems. They were also told that they could refer back to the cover page but not to the examples while working on the problems. This restriction was intended to increase the likelihood that the subjects would pay attention to the examples and how they were solved. The SV-T group (n = 13) studied four training examples isomorphic to the sociologist example. The SV-C group (n = 15) studied four training examples isomorphic to the cookie problem. The PV group (n = 22) studied two sociologist isomorphs and two cookie isomorphs. After studying the examples for about 15-20 min, the subjects then received four target problems to solve in the following order: Detroit Tigers, cookie, birthday, and football.

The subjects worked at their own pace for the entire experiment. In general, the subjects took about 30-45 min to complete the experiment. They were asked to show all their work but could use a calculator for the basic arithmetic.

Results and Discussion

The subjects' answers to the target problems were first scored as correct or incorrect. The solution and error frequencies were analyzed using the likelihood ratio chi-square test (G^2; Bishop, Fienberg, & Holland, 1975).

The subjects' written solutions were analyzed to evaluate whether subjects were trying to find subgoals such as λ and $P(X = x)$. It should be noted that, given that the subjects knew that all problems dealt with the Poisson distribution, it would not be too surprising if their protocols indicated that they were trying to find λ or some $P(X = x)$ in the test problems even if the protocols otherwise indicated that the subjects really did not know what they were doing. Nonetheless, the subjects' attempts to find these unknowns serve as a rough measure of whether or not they possessed a particular subgoal.

Tiger and cookie problems. Whereas all of the SV-T and PV subjects solved the Detroit Tigers problem correctly, only 13% of the SV-C subjects did. This difference is, of course, significant [$G^2(2) = 45.5, p < .0001$]. Similarly, whereas most of the SV-C and PV subjects solved the cookie problem correctly (87% and 86%, respectively), a much lower percentage (31%) of the SV-T subjects did [$G^2(2) = 13.9, p < .001$]. The 13% solution rate of SV-C subjects on the Detroit Tigers problem is not significantly less than the 31% success rate of the SV-T subjects on the cookie problem [$G^2(1) = 1.27, p < .3$].

Of the 9 SV-T subjects who failed to solve the cookie problem, 7 failed to realize that the final subgoal was to solve for λ, all 9 failed to realize the $P(X = 0) = .01$ was provided in the problem, and 3 of them tried to make up a frequency table as an aid to solve the problem. In contradiction to Learning Constraint 1, this indicates that most SV-T subjects did not recognize solving for λ as a subgoal, but rather simply attempted to apply the operations from the examples. This observation is further supported by the fact that most of these subjects went on to calculate expected frequencies rather than stopping after having calculated λ. It appears that these SV-T subjects did not make up the observed frequency table to calculate λ per se, but rather created the table so that they could apply the stereotyped operations and thereby achieve the subgoal of finding an expected frequency.

Of the 13 SV-C subjects who were scored as not solving the Detroit Tigers problem, all of them tried to calculate λ. However, 12 of these subjects tried to calculate λ by taking an observed frequency for some X (usually $X = 0$) and plugging that into the Poisson equation to solve for λ. This is the "wrong" method to use, because it ignores available frequency data that could give a better estimate of λ. Moreover, of these 12 subjects, 11 stopped after solving for λ and did not go on to calculate the expected frequencies (a new subgoal) for the various values of X. The remaining subject who got the problem wrong calculated λ using the frequency table approach, but did not go on to generate predicted frequencies for the various values of X.

The above results suggest that SV-C subjects had learned the subgoal of finding EVENT-AVG. However, the fact that 13 of the 15 SV-C subjects did not use λ to find expected frequencies suggests that they were mostly unable to recognize the new subgoal of finding expected frequencies. Rather, they seemed to consider λ as being the only possible final subgoal.

The fact that all the SV-C subjects tried to find λ in the Tigers problem supports the assumption that to label a subgoal in an example helps a learner learn that subgoal and recognize it in a new context. However, the fact that 7 of the 13 SV-T subjects did not try to find λ in the cookie problem seems to contradict the assumption. It may be, however, that the novel way in which the cookie problem was phrased ("How many raisins will a cookie contain on the average in order to achieve this result?") could have made it difficult for SV-T subjects to recognize that they were being asked to find λ. That is, the problem asked for a prediction of λ in order to achieve some desired outcome (only 1 cookie out of 100 having zero raisins) rather than a calculation of λ based on some known outcome (observed frequencies). The understanding that the relationship between λ and an outcome can go in either direction may have been too large a leap for most subjects.

Finally, with respect to the Tigers and cookie problems, it is worth noting that both the SV–T and SV–C subjects had difficulty in trying to come up with the correct method for finding λ for the novel problem. For both groups, the old method could not be adapted to the required method: rather, an essentially new method was needed. It seemed to be beyond the skills of most of the subjects to create a new method.

Birthday and football problems. There were no significant differences in solution rates among the three groups for the birthday or football problems (which were not isomorphic to any of the training examples). Overall, 48% of the subjects solved the birthday problem [SVC–T = 46%, SV–C = 47%, PV = 50%: $G^2(2) = .06$, $p = .97$] and 46% solved the football problem [SV–T = 31%, SV–C = 53%, PV = 50%: $G^2(2) = 1.72, p = .42$]. Two SV–T subjects who solved the birthday problem failed to solve the football problem, producing a nonsignificant difference [46% vs. 31%: $G^2(1) = .65, p < .5$] in success rates on those problems for that group. Both of these subjects set up the football problem correctly but calculated λ as 539/1,200 rather than 539/300. That is, they were confused about whether the fact that the tackles were made over four games mattered, and therefore multiplied the number of pieces of the field by the number of games. This finding is reminiscent of Reed et al.'s (1985) subjects, who needlessly (and incorrectly) multiplied amounts together in algebra word-mixture problems rather than realizing that these amounts were ready to be used directly in the equation.

The percentages of subjects who tried to find λ in the birthday problem (including subjects who solved the problem correctly) in order to find $P(X = 3)$ were 92%, 60%, and 86%, for the SV–T, SV–C, and PV groups, respectively. The difference among the groups is marginally significant [$G^2(2) = 5.27, p < .08$]. This pattern of results suggests that subjects who studied examples isomorphic to the Detroit Tigers problem (SV–T and PV subjects) may have been somewhat more likely to recognize that EVENT–AVG has to be found in order to find an expected probability. If the SV–T and PV groups are combined and compared to the SV–C group, the performance difference is significant [$G^2(1) = 4.97$, $p < .05$].

Performance on the football problem was similar to the performance on the birthday problem. The percentages of subjects who tried to find λ were 100%, 67%, and 95%, for the SV–T, SV–C, and PV groups, respectively. This difference is significant [$G^2(2) = 9.46$, $p < .01$], supporting the claim that the SV–C subjects lacked the necessary subgoal relationship of finding an expected probability by first finding EVENT–AVG. Again, if the combined SV–T and PV groups are compared with the SV–C group, the performance difference is significant [$G^2(1) = 8.52$, $p < .01$]. Overall, the performance of the SV–C subjects on the Detroit Tigers problem and the performance of the SV–T and PV subjects on the birthday and football problems support the assumption that subgoals, once learned, can be recognized in new problems that have a different context and contain only a subset of the subgoals from the studied examples. However, some of the results, such as the fact that SV–C subjects tried to find EVENT–AVG less often than did other subjects on the football and birthday problems, lend less support to this assumption (although even the SV–C subjects recognized this subgoal over 60% of the time in these new contexts).

Besides examining whether subjects tried to find EVENT–AVG in the birthday and football problems, it is instructive to examine the errors they made in trying to achieve this subgoal. For the birthday problem, 3 of the 7 unsuccessful SV–T subjects tried to find EVENT–AVG by setting it equal to 365/500 rather than 500/365. This suggests some confusion between events and trials. An additional subject found EVENT–AVG correctly (and labeled it as such) but for some reason did not go on to solve the problem. Thus, using a lenient criterion that allows the above responses, we could say that 4 of the 7 unsuccessful SV–T subjects were more or less able to adapt an old method to find EVENT–AVG. Applying this analysis to the PV group, only 1 of the 11 unsuccessful PV subjects adapted the old method in this way. Interestingly, 3 PV subjects tried to find a value for $P(X = 0)$ and used this (incorrectly calculated) value to find λ through the Poisson equation. That is, they tried to use the approach from the cookie problem to find EVENT–AVG. Although all 3 of these subjects used this value of λ to calculate $P(X = 3)$, it seems that they suffered from some negative transfer from the cookie-type training isomorphs. For the football problem, 7 of the 9 unsuccessful SV–T subjects set λ = 539/1,200 or 300/539, but only 3 of the 11 unsuccessful PV subjects did so. Four of the 11 PV subjects used a cookie-type approach to find λ, again suggesting negative transfer.

The results above suggest that to label a subgoal in training examples helps learners learn the subgoal and recognize it in new problems. This result provides initial support for the value of researchers' and teachers' conceptualizing examples and test problems in terms of the subgoals they presumably teach. A second result from Experiment 1 is that someone who learns two methods for achieving a subgoal does not seem to possess any advantage in modifying one of those methods in a novel problem, relative to someone who learns only one modifiable method. If anything, there is an increased possibility of negative transfer due to a less appropriate old method's intruding on the new problem. Perhaps direct instruction regarding the conditions under which each method is applied might help learners realize that sometimes neither of the old methods should be applied without adaptation to a novel problem; that is, learners would be better able to recognize that an old approach is not correct (see, e.g., Morris & Rouse, 1985). Negative transfer would therefore be reduced, since learners would be more likely to try to modify, rather than import, an old method to solve a novel problem requiring a modification in the old method. This issue was examined in Experiment 2.

EXPERIMENT 2

The difficulties that subjects in Experiment 1 had in recognizing when and how to modify an old method suggest that explicit mention of why a particular method was used in each example might help learners recognize when and how to adapt an old method for use in a novel problem. There is evidence from other research that when learners are given help in focusing on the goal structure of examples, they are more likely to apply the information presented in examples to a new problem correctly (Brown, Kane, & Echols, 1986; McKendree, 1986). Schoenfeld (1979) suggested that students can learn to do mathematical proofs from examples more effectively if heuristics for determining when to apply a particular approach are presented along with the examples. In addition, researchers have argued that learners often need explicit guidance concerning the underlying concepts of examples in order to be able to apply the examples to new problems (Gick & Holyoak, 1987; Mayer & Greeno, 1972; Stein et al., 1982). Lewis and Anderson (1985) showed that learners were more likely to learn correspondences between the givens in a geometry problem and appropriate solution methods for new problems if the learners explicitly indicated during training which method they thought was appropriate for each example and what information they would need in order to solve it.

Although it has been shown that additional explanatory material can help learning, there is evidence that not just any additional material aids learning. For example, Kieras and Bovair (1984) found that only information that helped learners infer how to operate a device enabled them to learn more quickly than a group that learned to operate the device by trial and error. Similarly, Reder, Charney, and Morgan (1986) found beneficial effects for explanatory material that focused on the syntax of computer commands but not for material that focused on elaborations of computer concepts.

Experiment 2 was designed to examine whether explicitly identifying the subgoals and methods in training examples and explaining why a particular method was chosen would aid subjects in recognizing subgoals in the target problems and reduce negative transfer resulting from their trying to apply an old method inappropriately rather than modify an old method.

Method

Subjects. Seventeen students from a probability class were recruited and were paid for their participation.

Materials and Procedure. There were two groups: the highlight group (n = 10) and the no-highlight group (n = 7). All subjects received a booklet containing a cover sheet, four training problems, and six target problems. No-highlight subjects received the same training examples as did the PV group in Experiment 1: that is, they studied two problems isomorphic to the Detroit Tigers problem and two problems isomorphic to the cookie problem. Highlight subjects received the same training examples; however, the solutions were annotated with comments indicating the subgoal and method that were being used in each part of the solution procedure. The comments also indicated why the subgoals and methods were appropriate for that example. Two of the annotated problems are presented in Appendix D, with the goal and method comments in the solutions highlighted. Care was taken to ensure that the annotated comments did not include other helpful information for understanding the problems, such as reminding the student that λ represented an average.

The six transfer problems involved the Detroit Tigers, cookie, birthday, and football, as well as "London," and the "Grateful Dead." The London problem was similar to the Detroit Tigers problem, except that EVENT-AVG was given directly in the problem and observed frequencies for different values of X were not specified (so EVENT-AVG could not be calculated). The Grateful Dead problem was like the birthday problem, except that EVENT-AVG was given directly in the problem. The football problem was modified so that 76 tackles were made over the course of one game. This was done to eliminate the confusion demonstrated by some subjects in Experiment 1 concerning whether or not the fact that the tackles were made over four games was relevant to the solution. The Detroit Tigers and cookie problems were thus isomorphic to examples studied by all subjects, whereas the birthday, football, London, and Grateful Dead problems were not.

Results and Discussion

Both groups did well on the problems that were isomorphic to the training examples. Nine out of 10 highlight subjects and all 7 no-highlight subjects solved the Detroit Tigers problem correctly. Nine out of 10 highlight subjects and 6 out of 7 no-highlight subjects solved the cookie problem correctly.

The number of subjects who solved the birthday problem did not differ significantly between the two groups [highlight = 4 out of 10 (40%), no-highlight = 2 out of 7 (29%); $G^2(1) = .238, p = .63$). However, their performance did differ on the football problem [highlight = 9 out of 10 (90%), no-highlight = 3 out of 7 (43%); $G^2(1) = 4.5, p < .04$]. Thus, although the annotations provided in the training examples did not seem to help the highlight subjects when they solved the birthday problem, these subjects did benefit on the football problem. This difference in performance on the football problem is due to 5 highlight subjects' finding EVENT-AVG correctly on the football problem but not finding it correctly in the birthday problem. No-highlight subjects had difficulty finding EVENT-AVG in both problems, although 1 subject who was unsuccessful on the birthday problem was successful on the football problem. These results suggest that the football problem might provide an easier mapping onto events and trials than the birthday problem does. Thus, when subjects are given training that encourages them to recognize the conditions for applying a learned method, and when this is combined with a novel problem that provides an easier mapping task, then the old method is more likely to be successfully modified.

All but 1 subject in each group solved the London problem correctly and all but 2 subjects in each group solved the Grateful Dead problem correctly. This suggests that, as predicted by Learning Constraint 2, most subjects were fairly good at realizing the need to achieve an old subgoal (finding EVENT-AVG) in a new context. Furthermore, most subjects were good at recognizing that EVENT-AVG was provided directly in these two problems. This recognition was probably aided by the fact that EVENT-AVG was labeled with the term "average" in the London and Grateful Dead problem statements.

In sum, subgoal and method annotations seem to provide some benefit in modifying an old method, at least when the target problem provides only a moderately difficult mapping task. Annotations do not seem necessary for enhancing the noticing of old subgoals, at least for the problems tested here. As in Experiment 1, manipulations of subgoals and methods were correlated with performance differences, again supporting the usefulness of the subgoal/method representational scheme.

GENERAL DISCUSSION

Two assumptions concerning the conditions under which learners acquire and transfer subgoals and methods have been proposed in this paper. A subgoal was assumed to be learned if it was labeled in an example solution, and it was assumed that this subgoal could be noticed in a new context. Across the two experiments, the subjects were generally successful in recognizing when a presumably learned subgoal needed to be achieved in a novel problem.

It was hypothesized that if a method for achieving a subgoal consisted of a linear series of steps rather than a hierarchy of other subgoals, it would be hard to adapt this method. This assumption was largely supported by the finding that subjects had difficulty adapting an old method in novel problems, even though the old method implicitly contained the needed subgoal. Because this subgoal was not explicit, the learner was less able to modify old steps to create the necessary method to achieve that subgoal. It would be interesting to teach learners, in the context of different examples, all the necessary subgoals to solve a novel problem but not all the methods needed to achieve those subgoals in the novel problem. This situation would allow an examination of method modification without imposing the additional burden of recognizing a new subgoal.

In some sense, improved performance via smaller subgoals and methods seems contradictory to the notion of knowledge compilation (see, e.g., Anderson, 1983). However, these two notions are different. The size of the initial subgoals and methods is assumed to be strongly influenced by the examples the learner studies, whereas compilation is largely a function of practice. There is nothing to prevent a learner, with experience, from combining methods that are frequently executed in order into a large, compiled method. However, representations of the smaller, precompiled methods, may continue to coexist with the new compiled method, especially if the original methods continue to be used in a variety of combinations across different types of problems.

The use of annotations in examples to highlight subgoals and methods seems to increase the likelihood that a learner will modify an old method rather than apply it without adaptation, perhaps because the learner is more likely to encode the conditions for applying the old method. However, further work should be done to examine which features of the annotation are most beneficial for reducing negative transfer (i.e., applying an old method without adaptation) and for aiding the adapting of an old method. For example, it may be that the justification for using a particular method is the most important feature for reducing negative transfer (Lewis & Anderson, 1985), and it may be that calling the learner's attention to the methods themselves (a form of elaboration) is what helps the learner modify the method in a novel problem (Reder et al., 1986).

The subjects' difficulty on the birthday problem relative to the football problem in Experiment 2 suggests that learners may need practice in mapping the parts of problems to the appropriate concepts or variables, perhaps without the additional requirement of solving the problem (cf. Sweller, Mawer, & Ward, 1983). This suggestion could also be explored in a domain such as mechanics, where students may need practice in setting up problems, such as drawing free-body diagrams, apart from actually solving them. Since mapping (perhaps "translating" is a better term) is only part of what is required to solve a problem, it may get less attention than it deserves unless the student is fortunate enough to recognize that mapping is a separate subtask. Students typically are not so fortunate (see also Larkin, 1983). Teachers rarely seem to isolate underlying subgoals or highlight the conditions for why a particular method is used when they present examples. In addition, examples done in class can involve a consistent mapping, further preventing the student from even recognizing that mapping is a separate component.

It may be possible to develop a theoretically motivated methodology for constructing examples for textbooks in quantitative domains. This methodology would involve first identifying the subgoals and methods that students need to learn (Bovair et al., 1990; Catrambone, 1990) and then designing examples that should lead to the acquisition of these subgoals and methods and the ability to modify the methods. This characterization in terms of subgoals and methods could help teachers construct examples that convey the concepts they wish students to learn, as well as help them construct more diagnostic test problems. Ultimately, careful theory-driven development of training examples may make it easier for learners to see beyond the superficial features of problems to the subgoals and methods that need to be acquired.

REFERENCES

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

ANDERSON, J. R., FARRELL, R., & SAUERS, R. (1984). Learning to program in LISP. *Cognitive Science*, 8, 87-129.

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

Bovair, S., Kieras, D. E., & Polson, P. (1990). The acquisition and performance of text editing skills: A production system analysis. *Human-Computer Interaction*, **5**, 1-48.

Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, **1**, 103-121.

Catrambone, R. (1990). Specific versus general procedures in instructions. *Human-Computer Interaction*, **5**, 49-93.

Catrambone, R., & Holyoak, K. J. (1987). Transfer in problem solving as a function of the procedural variety of training problems. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society* (pp. 36-49). Hillsdale, NJ: Erlbaum.

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.

Foltz, P. W., Davies, S. E., Polson, P. G., & Kieras, D. E. (1988). Transfer between menu systems. In *Proceedings CHI '88 Human Factors in Computing Systems* (pp. 107-122). New York: Association for Computing Machinery.

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

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

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning: Contemporary research and applications* (pp. 9-46). Orlando, FL: Academic Press.

Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology*, **16**, 94-143.

Hayes, J. R., Waterman, D. A., & Robinson, C. S. (1977). Identifying the relevant aspects of a problem text. *Cognitive Science*, **1**, 297-313.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, **8**, 255-273.

Lancaster, J. S., & Kolodner, J. L. (1988). *Problem solving and learning in a natural task domain* (Tech. Rep. No. GIT-ICS-89/02). Atlanta: Georgia Institute of Technology.

Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 75-98). Hillsdale, NJ: Erlbaum.

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.

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

McKendree, J. (1986). *Impact of feedback content during complex skill acquisition*. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh.

Morris, N. M., & Rouse, W. B. (1985). The effects of type of knowledge upon human problem solving in a process control task. *IEEE Transactions on Systems, Man, & Cybernetics*, **15**, 689-707.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skill. *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., & Actor, C. (1989). *The use of examples and procedures in problem solving*. Unpublished manuscript.

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. (1989a). 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.

Ross, B. (1989b). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438-469). Cambridge, England: Cambridge University Press.

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

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic Press.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 484-494.

Singley, M. K., & Anderson, J. R. (1988). A keystroke analysis of learning and transfer in text editing. *Human-Computer Interaction*, **3**, 223-274.

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

Spencer, R. M., & Weisberg, R. W. (1986). Context-dependent effects on analogical transfer. *Memory & Cognition*, **14**, 442-449.

Stein, B. S., Bransford, J. D., Franks, J. J., Owings, R. A., Vye, N. J., & McGraw, W. (1982). Differences in the precision of self-generated elaborations. *Journal of Experimental Psychology: General*, **111**, 399-405.

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, **112**, 639-661.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *Journal of Mathematical Behavior*, **3**, 3-71.

APPENDIX A
Example Poisson Distribution Problem

Eighty-four sociologists were asked how many journal articles they had published in the last 3 years. It was observed that 23 sociologists published 0 articles, 29 published 1 article, 17 published 2 articles, and 15 published 3 articles. Fit a Poisson distribution to x; that is, give the expected frequencies for the different values of x based on the Poisson model.

Solution:

$$E(X) = \frac{[0(23)+1(29)+2(17)+3(15)]}{84} = \frac{108}{84} = 1.3$$

$$= \lambda$$

$$= \text{average number of articles published per sociologist}$$

$$P(X=x) = \frac{[(e^{-\lambda})(\lambda^x)]}{x!} = \frac{[(2.718^{-1.3})(1.3^x)]}{x!} = \frac{(.27)(1.3^x)}{x!}$$

(so, for example, when $x = 3$, $P(X = 3) = \frac{(.27)(1.3^3)}{3!} = .10$)

Fitted Poisson Distribution:

x	Expected Frequency
0	$.27 \cdot 84 = 23$
1	$.35 \cdot 84 = 29$
2	$.23 \cdot 84 = 19$
3	$.10 \cdot 84 = 8$

APPENDIX B
Another Example Poisson Distribution Problem

A football field was cut up into 300 equally sized pieces. A total of 539 tackles were made on the field over the course of four games. Use the Poisson distribution to determine the probability that a randomly chosen piece of the field had exactly 1 tackle made on it.

APPENDIX C
Cookie and Birthday Problems

Suppose you were making a batch of raisin cookies and you did not want more than one cookie out of 100 to be without a raisin. How many raisins will a cookie contain on the average in order to achieve this result? Use the Poisson distribution to find your answer.

Suppose you took a random sample of 500 people and found out their birthdays. A "success" is recorded each time a person's birthday turns out to be January 1. Assume there are 365 days in a year, each equally likely to be a randomly chosen person's birthday. Fit a Poisson distribution to x (the number of people born on January 1) and find the predicted likelihood that exactly 3 people from the sample are born on January 1.

APPENDIX D
Examples of Annotated Problems for Highlight Subjects (Experiment 2)

Example 1

In the 432 years from 1500 to 1931, war broke out somewhere in the world a number of times. The table below gives the distribution of the number of years on which x wars broke out, for $x = 0, 1, 2, 3, 4$ (or more). Fit a Poisson distribution to x and give the expected frequencies for the different values of x based on the Poisson model.

Number of Wars Beginning in a Given Year	Observed Frequency
0	223
1	142
2	48
3	15
4 or more	4
Total	432

Solution:

The goal of this problem is to create an expected frequency table. In order to do that, we must calculate the expected probabilities for the different values of x using the Poisson equation. We can then multiply these expected probabilities by the total number of "chances" in order to get the expected frequencies.

Since we know the values of x we are interested in, the only value we need to find before we can use the Poisson equation is λ.

One way to determine λ (the next problem shows a different way) is to divide the total number of events by the total number of chances.

$$E(X) = [0(223)+1(142)+2(48)+3(15)+4(4)]/432$$
$$= 299/432 = .692$$
$$= \lambda$$
$= $ average number of wars that broke out each year

Now that we have determined λ, we can calculate the expected probabilities for various values of x. The Poisson equation becomes:

$$P(X = x) = [e^{-.692}.692^x]/x! = [(.5006)(.692)^x]/x!$$

(so, for example, when $x = 3$, $P[X = 3]$

$= [(.5006)(.692)^3]/3! = .0276$).

We now multiply the expected probabilities by the total number of chances in order to get the expected frequencies for each x.

Fitted Poisson Distribution:

x	Expected Frequency
0	.5006·432=216
1	.3464·432=150
2	.1199·432=52
3	.0276·432=12
4 or more	.0048·432=2

Example 2

An archaeologist found a horizontal quarry surface that had been previously divided up into 30 squares about 1 m on a side. He dug in each square in order to find specimens of the extinct mammal *Ditolestes motissimus*. Out of all the squares, 16 of them contained no specimens. Use the Poisson distribution to determine how many specimens were found on the average in each square.

Solution:

Our goal is to find λ.

Note that, in the prior problem, finding λ was not the final goal, but in this problem it is. However, unlike the prior problem we cannot calculate the exact value of λ from the observed values of x, since we are not given all of the observed values. A different way to find λ is to use the Poisson equation, $P(X = x) = [e^{-\lambda}\lambda^x]/x!$, and solve for λ rather than $P(X = x)$. In order to solve for λ though, we need a value for at least one $P(X = x)$. We are given enough information to calculate $P(X = 0)$: $P(X = 0) = 16/30 = .53 =$ probability of a square having no specimens, so:
Now we can use the Poisson equation to estimate λ:

$$.53 = [e^{-\lambda}\lambda^0]/0! = P(X = 0)$$

now, since $0! = 1$ and $\lambda^0 = 1$, then

$$.53 = e^{-\lambda}$$

Solve for λ by taking natural log of both sides:

$$\ln[.53] = \ln[e^{\lambda}]$$

$$-.635 = \lambda$$

$.635 = \lambda =$ average number of specimens found in each square

We could have used this approach to finding λ in the prior problem by just calculating $P(X = 0)$ in that problem and plugging it into the Poisson equation to get λ. However, in the prior problem we would have been ignoring the rest of the information (the observed frequencies for all the other values of x), which helped us to calculate λ more precisely.

(Manuscript received November 1, 1989;
revision accepted for publication March 22, 1990.)