Journal of Experimental Psychology:
Learning, Memory, and Cognition
1991, Vol. 17, No. 4, 753–766

0278-7393/91/$3.00

Use of Examples and Procedures in Problem Solving

Stephen K. Reed
San Diego State University

Cheryl Actor Bolstad
Florida Atlantic University

We studied how successfully students could use examples and procedures to construct equations for work problems. According to the proposed theory, the procedures indicate how to generate values that differ in structure from the example. The first experiment compared 3 groups of students who received a simple example, a set of procedures, or both. A mathematical model with 3 parameters (the probability of generating a correct value by matching the example, following a procedure, or using general knowledge) accounted for 94% of the variance for how the 3 instructional groups performed over 4 levels of transformation. A second experiment extended the predictions of the model to include either a complex example, a complex example and procedures, or a complex example and a simple example.

Two alternative approaches for instructing people about a task are to present either a detailed example or a set of procedures. Each method has its advantages and disadvantages. The advantage of an example is that it illustrates how the procedures are applied to a particular situation. For example, students in a college algebra class could be given a detailed solution to the following problem:

> Ann can type a manuscript in 10 hr, and Florence can type it in 5 hr. How long will it take them if they both work together?

The disadvantage of an example is that it may not be very helpful for solving problems that are slightly different. Students often have difficulty in solving variations of the examples, such as a problem in which one person worked more hours than the other (Reed, Dempster, & Ettinger, 1985).

The advantage of procedures or rules is that they can specify the component steps for solving a variety of problems. One rule might specify what to do when one person works longer than another, and another rule might specify what to do when rate rather than time is the unknown. The disadvantage of procedures is that they can be rather abstract and isolated, resulting in minimal understanding of the task as a whole. Thus learning to operate a device can be facilitated if a set of procedures is supplemented with additional material (functional, structural, or diagrammatic information) that enables students to better understand and integrate the procedures (Kieras & Bovair, 1984; Smith & Goodman, 1984; Viscuso & Spoehr, 1986).

Experiment 1 is based on Cheryl Actor Bolstad's master's thesis, under the supervision of Stephen K. Reed. This work was funded by Air Force Office of Scientific Research Grants AFOSR 88-0008 and AFOSR 89-0107 to Stephen K. Reed.

We are grateful to Howard Hock and Ed O'Brien for their comments on the thesis, to Rogers Hall, Gary Olson, John Sweller, and several anonymous reviewers for their comments on the manuscript, and to Javier Sanchez for his assistance with Experiment 2.

Correspondence concerning this article should be addressed to Stephen K. Reed, Department of Psychology, San Diego State University, San Diego, California 92182.

When students receive both procedures and examples for solving problems, they seem to emphasize examples. Pirolli and Anderson (1985) collected detailed verbal protocols of two college students and one 8-year-old child learning to program recursive functions. The protocols suggested that problem solving by analogy to worked examples was frequent during the initial programming attempts. A particularly dramatic demonstration of the power of examples occurred in a series of experiments conducted by LeFevre and Dixon (1986). Written instructions for solving inductive reasoning problems were accompanied by an example that conflicted with the instructions. Most of the subjects in the six experiments consistently used the example and disregarded the procedure described in the instructions.

Sweller and Cooper (1985) proposed that students need to be shown a wide range of worked examples to become proficient in solving problems. But they also anticipated a possible criticism of this conclusion—a worked example of every conceivable problem would require too many examples. An alternative approach would be to use some combination of examples and rules in which the rules would inform students what to do when the test problems differed from the examples.

We follow this approach in the two experiments reported in this article. The problem set for both experiments are work problems (see Appendix A for examples) that can be solved by the equation:

$$\text{Rate}_1 \times \text{Time}_1 + \text{Rate}_2 \times \text{Time}_2 = \text{Tasks}. \quad (1)$$

The rates refer to how long it takes each of two workers to complete a task, the times refer to how long each worker spends on the task, and the task refers to how many tasks they must complete. Even when two problems share the same equation, however, students are often unable to use the solution of one to solve the other because they cannot generate new values to fit the slots of the equation (Reed & Ettinger, 1987).

Schema-Based Theories of Problem Solving

Our attempt to improve students' ability to transfer a solution was influenced by work in artificial intelligence on schema-based theories of problem solving. The Knowledge Representation Language (KRL) constructed by Bobrow and Winograd (1977; Winograd, 1975) provided the initial framework for such theories. The data structures of KRL are built of descriptions that are clustered into structures called *units*. Each unit is assigned to a category type and contains slots that are associated with the conceptual entities referred to by the units. Associated with each slot are a set of procedures that can be used to instantiate values for that slot.

The application of this approach to problem solving is illustrated by an artificial intelligence system called FERMI (Larkin, Reif, Carbonell & Gugliotta, 1988). FERMI is capable of calculating pressure drops in fluids, potential drops in electric circuits, and centers of mass for planar objects. Knowledge is stored at various levels of generality so a specific schema can inherit the content of a general schema (including slots and attached procedures), causing some transfer of knowledge across different domains.

Several investigators, including Larkin and her colleagues, have argued that such schema-based systems offer promise for improving instruction and providing insights into effective ways of organizing human knowledge. Greeno (1983) also discussed several examples of how schema-based learning might facilitate understanding in mathematical problem solving by either teaching new applications of an existing schema or new procedural attachments.

In our view a *schema* is a cluster of knowledge that provides a skeleton structure that can be instantiated or filled out with the detailed properties of a particular instance (Thorndyke, 1984). We propose that Equation 1 provides a skeleton structure for solving many work problems, but students will fail to solve these problems correctly if they cannot enter the appropriate values into the slots of the equation. An example shows them how to construct these values but is not helpful when the value has to be modified (such as when one worker works more hours than the other or when part of the task is already completed). In this case, it may be helpful to have a set of procedures or rules associated with each of the concepts in the equation that inform students how to construct a value for these different situations. According to this view, an equation provides an organizational framework for showing the formal relations among concepts, an example provides an integrated solution showing how values are instantiated for a particular problem, and procedures show how to modify these values for variations of the example.

We follow this general approach in the current study by giving students a detailed solution and a set of procedures that should help them apply the solution to similar problems. The first experiment compares three groups of students who receive either an example, procedures, or both the example and procedures. The data allow us to evaluate a simple mathematical model of how students attempt to solve problems in each of these three situations. We extend the model in the second experiment by including additional instructional conditions, such as providing two examples that span the test problems.

Experiment 1

The primary purpose of Experiment 1 was to evaluate a model of how students use an example and procedures. We compared three groups of students—one group received the set of procedures shown in Appendix B, a second group received a solution to the simple example in Appendix A, and a third group received both the example and procedures.

Appendix A also shows the set of test problems that differed from the simple example by either zero, one, two, or three transformations. The first test problem in Appendix A is equivalent to the example and therefore differs by zero transformations. The problems that differ by one transformation were transformed by changing either the rate, the time, or the task. A change in the rate involved expressing the rate of one worker relative to the other worker rather than as an independent number (see Problem 2). A time change occurred when one worker labored longer than the other (Problem 3). A change in task occurred when part of the task had been completed earlier (as in Problem 4). Problems that differ from the test problem by two transformations were created by changing either rate and time (Problem 5), rate and task (Problem 6), or time and task (Problem 7). And, of course, the test problem that differed by three transformations was created by changing the rate, the time, and the task (Problem 8).

The four transformation levels and three instructional methods enabled us to evaluate the predictions of a model for each of these 12 conditions. Because both the example and the procedures provide students with the basic equation for solving these problems, we assume that the probability of generating a correct equation is equal to the probability of correctly generating the values for the five quantities: $Rate_1$, $Time_1$, $Rate_2$, $Time_2$, and Tasks Completed. Students can generate these values by using either the information provided in the example, information provided in the procedures, or their general knowledge about these problems. According to our model, students attempt to generate the values by first using the example, then the procedures, and finally their general (prior) knowledge.

The model has three parameters. A student can generate a correct value by either correctly matching the structure of a corresponding value in the example (m), correctly applying a rule in the procedures (r), or correctly using general knowledge (g). Consider the predictions for the instructional group who receives the example and the procedures. When the test problem is equivalent to the example, a student can generate all five values by using the matching operation. The probability of generating a correct equation is therefore m^5—the probability that the student correctly applies the matching operation to each of the values in the example. When the test problem differs by one transformation the probability of a correct equation is m^4r. In this case the student can match four of the quantities but must use the procedures to generate the transformed value. Following the same logic, the probability of correctly generating an equation should be m^3r^2 for two transformations and m^2r^3 for three transformations. Assuming that it is easier to match values in an example than follow procedures ($m > r$), the model predicts a decline in performance as the number of transformations increase.

When students have only the example, they must rely on their general knowledge to generate the transformed quantities. The probability of constructing a correct equation should therefore be m^5 for zero transformations, m^4g for one transformation, m^3g^2 for two transformations, and m^2g^3 for three transformations. The generalization gradient should be steeper for the example group than for the example and procedures group if the rules increase the probability of correctly generating the transformed values ($r > g$).

When students have only the rules, there should not be a generalization gradient. In this case, the probability of constructing a correct equation should simply be r^5, the probability of correctly applying a rule to generate each of the five values. The following experiment was designed to collect the data required to evaluate the model.

Method

Subjects. The subjects were 65 students in two college algebra classes at Florida Atlantic University and were tested during class. The students ($n = 47$) in one class were ready to begin working on word problems in the course. The students in the other class ($n = 18$) had just begun working on word problems but hadn't received any of the problems used in the experiment. The students in each class were randomly assigned to the three instructional conditions, resulting in 22 students in the example group, 21 students in the procedures group, and 22 students in the example and procedures group. The experimenter informed them that they would receive copies of the instructional material and the correct answers when they completed the task.

Procedure. Students were informed that the purpose of the experiment was to compare several different instructional methods for teaching students how to construct equations for algebra word problems. All students were initially given 3 min to attempt to construct a correct equation for the example problem. They then spent 5 min studying the instructional material, which consisted of a detailed solution of the example for the example group, the set of procedures in Appendix B for the procedures group, and both the example and procedures for the example and procedures group.

The eight test problems occurred on a single page in the order shown in Appendix A for approximately half of the subjects in each group and in the reverse order for the remainder. Students had 16 min to construct the eight equations and could work on the problems in any order. They were allowed to refer back to the instructional material as they worked on the problems.

Results

An equation was scored as correct if it had the same structure as Equation 1, with the correct values entered for the two rates, the two times, and the task. A value was scored as correct if it was mathematically equivalent to the correct answer; for instance, both $1 - 1/3$ and $2/3$ were scored as correct for the tasks variable in Problem 4. Equations that were mathematically equivalent to the correct answer were also scored as correct, such as adding 1/3 to the left side of the equation rather than subtracting 1/3 from the right side of the equation. However, mathematically correct transformations of Equation 1 rarely occurred. We did require that students group values such as $h + 2$ when one worker labored more hours than the other to show that the rate was multiplied by the entire quantity, rather than simply by h. We were flexible in allowing any letter to represent the unknown variable, as long as the same letter was used for both of the time variables in the equation.

Figure 1 shows how well the three groups performed at each of the four transformation levels. The results confirm the expected generalization gradients for the two groups that received the example. Also, as expected, the gradient was not as steep for the group that received the example and procedures. We first report an analysis of variance (ANOVA) of these results to determine which differences are significant. We then evaluate how well these results fit the predictions of the proposed model (see Table 1 for data on individual problems).[1]

Tests of significance. The data in Figure 1 were analyzed in a two-factor ANOVA in which instructional method was a between-subjects factor and transformations was a within-subjects factor. Significant effects were found for instructional method, $F(2, 62) = 6.01$, $MS_e = 0.273$, $p < .01$, transformations, $F(3, 186) = 60.68$, $MS_e = 0.057$, $p < .001$, and their interaction, $F(6, 186) = 9.91$, $MS_e = 0.057$, $p < .001$.

The percentage of correct equations for each of the instructional groups across the four transformations was 15% for the procedures group, 34% for the example group, and 42% for the example and procedures group. A Newman-Keuls test indicated that both the example and the example and procedures group differed significantly from the procedures group, but these two groups did not differ significantly from each other.

The effect of instruction was also analyzed at each transformation level because of the Instruction × Transformation interaction. The ANOVA revealed a significant effect at zero transformation, $F(2, 62) = 17.75$, $MS_e = 0.158$, $p < .001$. The effect of instruction was not significant at one transformation, $F(2, 62) = 2.61$, $MS_e = 0.121$, or at higher levels of transformation.

This analysis was supplemented with a planned comparison of the example and procedures and the procedures group for each transformation level. These two groups differed significantly for 0, $t(62) = 5.18$, $p < .001$, and one transformation, $t(62) = 2.21$, $p < .05$.

Predictions of the model. The purpose of the model was to fit the 12 data points in Figure 1 by estimating values for the three parameters. The model has the basic form:

$$\text{probability correct} = m^x r^y g^z,$$

in which x is the number of values generated by matching the example, y is the number of values generated by using the rules, and z is the number of values generated by using general knowledge. The parameters m, r, and g were estimated by using multiple linear regression after using logs to create a linear equation:

$$\log(\text{probability correct}) = x \times \log m + y \times \log r + z \times \log g. \qquad (2)$$

Applying Equation 2 to the 12 data points in Figure 1 resulted in parameter estimates of .96 for m (the probability

[1] Our analysis of the data in terms of number of transformations required that we combine the data from Problems 2, 3, and 4 to form the one-transformation problems and from Problems 5, 6, and 7 to form the two-transformation problems. Results for each of these six problems are reported in Table 1.

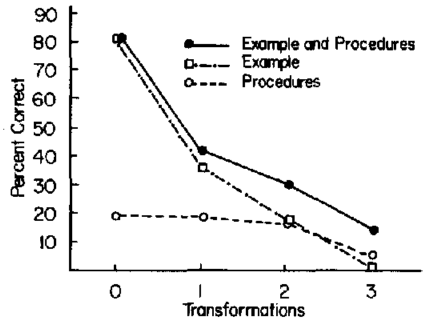

Figure 1. Percentage correct equations for the 3 instructional groups over the 4 levels of transformation in Experiment 1.

of correctly matching the example), .65 for r (the probability of correctly applying a rule), and .45 for g (the probability of correctly applying general knowledge). Table 2 shows the observed and predicted values. The model accounts for 94% of the variance.

Evaluation of the parameters. The parameter estimates attempt to predict the probability of correctly generating the entire equation from the probability of correctly generating each of its values. We tested the estimates by separately scoring how many of the values were correct in each of the generated equations.[2] For the example and example-and-procedures groups, 28 of the 40 values (8 equations × 5 values) could be generated by matching the example. Subjects in the example groups correctly generated 77% of these quantities, and subjects in the example and procedures group correctly generated 76% of the quantities. Subjects in the example group were correct on 37% of the 12 transformed quantities, compared with 50% for subjects in the example and procedures group. And finally, subjects in the procedures group, who had to rely on the procedures to generate all of the quantities, were correct on 46% of the values.

Table 1
Percentage Correct on Individual One- and Two-Transformation Problems

Group	Problem 2	3	4	5	6	7
Experiment 1						
Example	55	18	36	9	36	9
Procedures	19	14	24	14	19	19
Example and procedures	55	32	45	27	41	23
Experiment 2						
Procedures	14	8	14	3	7	0
Simple example	62	21	21	21	31	21
Complex example	28	14	45	14	48	48
Simple example and procedures	59	28	62	31	48	34
Complex example and procedures	34	45	62	41	62	69
Simple example and complex	83	48	79	55	66	66

Note. Problems 2, 3, and 4 are one-transformation problems and Problems 5, 6, and 7 are two-transformation problems for all groups except the complex example and the complex example and procedures groups. The classification is reversed for these two groups.

Comparing these results with the parameter estimates reveals (a) an excellent qualitative fit that corresponds to the expected pattern of relations but (b) a poor quantitative fit in which the parameter estimates are too high. The pattern of relations is consistent with the assumption of the model that the success of using the example or rules is independent of whether the example and rules occur separately or are combined. The probability of correctly generating a quantity from an example was almost identical for subjects in the example and example-and-procedures groups, and the probability of correctly applying a rule was very similar for subjects in the procedures and example-and-procedures groups. Also consistent with the parameter estimates is the finding that students did much better in generating quantities when they could match the example and did better on the transformations when they had the procedures.

However, the parameter estimates of $m = .96$, $r = .65$, and $g = .45$ indicate that subjects should have done better in generating the individual quantities. Because all the individual quantities are correct in the correct equation, this discrepancy is caused by the incorrect equations being more incorrect (having fewer correct values) than is predicted by the estimated parameters. Part of this discrepancy may be caused by the absence of any answer on 62 of the 520 items. We suspect that had subjects generated equations for these items, the equations would be partially correct, increasing the number of correct quantities in the incorrect equations.

Transformations. The claim that some subjects are performing consistently well across problems can be evaluated by analyzing whether a subject's success on problems with one transformation indicates that they will do well on problems with two or three transformations. This analysis was guided by the concept of knowledge spaces. The knowledge state of a subject can be represented by the particular subset of problems that the subject is capable of solving (Falmagne, Koppen, Villano, Doignon, & Johannesen, 1990).

An interesting special case involves problems that consist of components that can be tested either individually or in logical combinations. For instance, if there are three basic components, c_1, c_2, c_3, then the logical combinations are c_1c_2, c_1c_3, c_2c_3, and $c_1c_2c_3$. One assumption is that a hierarchical relationship exists in which success on any combination requires success on all of the individual components (Marshall, 1981). If this assumption is correct, it will be unnecessary to separately test individual components if a person is successful on the combination. However, the hierarchical assumption does not imply that mastery of the individual components implies mastery of the combination. Failure on the combi-

[2] The analysis of individual quantities raises the issue of whether it is possible to correctly generate all five quantities but incorrectly generate the equation. This rarely occurred. For example, using the large data base from Experiment 2, we found only six instances in 1,392 equations. Two instances occurred when subjects forgot the addition sign. Three instances occurred when subjects incorrectly paired the rate and time quantities in the two products. The last instance also involved a grouping error.

Table 2
Observed and Predicted Values for Three Groups in Experiment 1

Transformations	Example			Procedures			Example and procedures		
	Observed	Predicted	Model	Observed	Predicted	Model	Observed	Predicted	Model
0	82	81	m^5	19	12	r^5	82	81	m^5
1	36	37	m^4g	19	12	r^5	42	55	m^4r
2	18	17	m^3g^2	17	12	r^5	30	37	m^3r^2
3	0	7	m^2g^3	5	12	r^5	14	25	m^2r^3

Note. The predictions are based on parameter estimates of $m = .96$ (the probability of correctly matching the example), $r = .65$ (the probability of correctly applying a rule), and $g = .45$ (the probability of correctly applying general knowledge).

nation can result from either the necessity to perform the component skills in a certain order or simply some confusion factor (perhaps resulting from cognitive overload) when several components are in the same item.

Marshall's analysis provides a framework to test how individual subjects perform across problems that require one, two, and three transformations. If there is cognitive overload, then subjects who correctly solve two single-transformation problems may not solve a problem that contains both transformations. This hypothesis can be evaluated in the top half of Table 3, which shows how success in solving two one-transformation (1t) problems (zero, one, or two solutions) relates to solving the corresponding two-transformation (2t) problem. The data in the bottom row of Experiment 1 do not support an overload hypothesis: 33 of the 36 correct solutions (92%) of the two 1t problems were accompanied by a correct solution of the 2t problem.[3]

Another comparison related to cognitive overload is the number of subjects who correctly solve two 1t problems but fail to solve the 2t problem (3 cases in Table 3) versus the number of subjects who solve the 2t problem but fail to solve both 1t problems (10 cases in Table 3). The hierarchical model discussed previously suggests that there should be more subjects in the first group, but our model correctly predicts that there should be more subjects in the second group. The prediction is based on the greater number of pattern-matching operations that are required to correctly solve two 1t problems (eight) than are required to solve one 2t problem (three). Both cases require that students are correct on the two transformations. The fact that they occur in the same problem for the 2t case does not matter because our model assumes that the probability of correctly generating a transformed value is independent of how many transformations appear in the problem.

A greater challenge of this assumption is the three-transformation problem (Problem 8). This problem appears very difficult, and it seems reasonable that some students might be able to solve all three 1t problems but fail on the 3t problem. Unfortunately, the cognitive overload hypothesis is difficult to test for this condition because so few students solved all three 1t problems and the 3t problem. Three students solved the three 1t problems but failed the 3t problem, 1 student solved the 3t problem without solving all three 1t problems, and 3 students solved all four problems. This pattern is quite different from the pattern of the 2t problems and suggests that there is cognitive overload with three transformations. This hypothesis is also consistent with the finding that our model overpredicts performance on the 3t problem (see Table 2).

Table 3
Effect of Solving One-Transformation (1t) Problems on Solving the Corresponding Two-Transformation (2t) Problem

Experiment	1t solutions	2t solution: No	2t solution: Yes
1	0	100	2
	1	49	8
	2	3	33
2	0	156	1
	1	66	28
	2	15	82

Note. These results are total cases based on transformations of rate and time (Problems 2, 3, and 5), rate and task (Problems 2, 4, and 6), and time and task (Problems 3, 4, and 7).

Discussion

We believe the proposed model provided a useful framework for analyzing the results. First, it provided a reasonably good fit of how subjects in the three instructional groups performed over four levels of transformation. Second, detailed analyses of the data were generally supportive of various assumptions regarding the context-free nature of the proposed operations across instructional conditions and transformations. Success in both pattern matching and rule application was not influenced by whether examples and rules occurred together or alone. In addition, there was no evidence in the two-transformation problems for a cognitive overload hypothesis that increasing the number of transformations increases the difficulty of each individual transformation. However, there may be some overload in the three-transformation problem.

A disappointing aspect of our results was the ineffectiveness of the rules, and one goal of Experiment 2 was to provide more effective rules. There are several ways to modify the

[3] A model does not need to assume cognitive overload to generate a generalization gradient. A person who could solve two of the three 1t problems should be able to solve the 0t problem but would fail on two of the three 2t problems and on the 3t problem.

rules that may increase their effectiveness. First, the relevant rules could be elaborated to provide more information. For example, students sometimes fail to place parentheses around the expression $h + 1$ to represent that one worker labored for 1 hr more than the other. Parentheses are required to indicate that $h + 1$, rather than simply h, is multiplied by the rate. Providing such information in the rules should increase their effectiveness. Second, rules that are not needed to solve a particular set of test problems could be eliminated. It would be desirable to have an extensive set of rules, but the gradual introduction of the rules might be a more effective instructional technique.

A second goal of Experiment 2 was to apply the model to a greater range of instructional conditions. This would enable us to test the generality of the model and determine whether there are better instructional techniques for generating transformed values.

Experiment 2

We designed Experiment 2 to modify the rules and expand the instructional conditions used in Experiment 1. Appendix C shows the modified set of rules, which we hoped would be more effective. Representing the relative rate of work is explained by using numbers (1/10 and 1/5) rather than variables (r and $4r$ in the previous set) because rate is never the unknown variable in the test problems. Students were also instructed to place parentheses around a quantity such as $h + 3$ when representing how much more time one worker spends on the task than another worker. And finally, the rules explicitly stated how to calculate the number of tasks when part of the task had been completed. The part completed should be subtracted from the total number of tasks.

In addition, we added three new instructional conditions to the three conditions evaluated in Experiment 1. Except for the change in rules, the initial conditions remained the same. Each new condition included an example (the complex example in Appendix A) that was equivalent to the most complex test problem (Problem 8). The three new conditions were (a) the complex example, (b) the complex example and procedures, and (c) the complex and simple examples.

We included a complex example in the new conditions to test the generality of the model. Transformations from the simple example in Experiment 1 created more complex problems. Thus the steep generalization gradients could be caused by both increased complexity and increased dissimilarity from the example. When a complex example is used, the transformations produce simpler problems. This should produce a generalization gradient that is less steep than the gradients obtained for the simple example and allow us to apply the model to a different pattern of results.

The instructional condition that included both a simple and a complex example should be interesting for practical and theoretical reasons. The practical reason is that the ineffectiveness of the rules in Experiment 1 requires the search for other approaches. Examples that are equivalent to Problems 1 and 8 in Appendix A span the set of eight test problems because the information needed to solve each test problem is contained in the two examples. According to the proposed model, students should apply the pattern-matching operation to the quantities in the two examples. All five quantities can be obtained from the simple example for Problem 1. Problems 2–4 can be solved by matching the simple example on four quantities and the complex example on one quantity. Problems 5–7 can be solved by matching the complex example on four quantities and the simple example on one quantity. And finally, Problem 8 can be solved by matching the complex example on all five quantities.

A fourth issue was forced upon us during the analysis of the results. We discovered that we had unintentionally modified Problem 8 to read that John works 1 hr longer, rather than Paul works 1 hr longer. In the other test problems (as in the complex example), students were asked to find the time of the person who worked fewer hours when the hours differed. This issue is somewhat analogous to the role of object correspondences studied by Ross (1987), who found that reversing object correspondences caused a significant decrement in substituting the correct values into a formula. We will discuss the impact of this change when presenting the results.

Method

Subjects. The 174 subjects were enrolled in introductory psychology courses at San Diego State University and received course credit for their participation. They were tested in small groups and assigned randomly to the six instructional conditions except for the constraint that there would be an equal number of subjects (29) in each condition. The instructional conditions consisted of the three conditions used in Experiment 1 (simple example, procedures, simple example and procedures) and three new conditions (complex example, complex example and procedures, simple example, and complex example).

Procedure. The procedure was nearly identical to the procedure followed in Experiment 1. Students were initially given 5 min to attempt to construct a correct equation for the simple and complex examples. They then spent 5 min studying the instructional material, which consisted of either a single example, procedures, an example and procedures, or two examples. Finally, they were given 20 min to construct equations for the eight test problems that appeared on a single page in the two orders described for Experiment 1.

Results

Pretest. There was only one correct equation—the equation for the simple example—on the pretest.

Effect of instruction. The data were analyzed in a two-factor ANOVA in which instructional method was a between-subjects factor and transformations was a within-subjects factor. For subjects in the simple example, procedures, simple example and procedures, and simple example and complex example groups, transformations were measured from the simple example as described for Experiment 1. For subjects in the complex example and the complex example and procedures groups, transformations were measured from the complex example. Problem 8 was zero transformations, Problems 5–7 were one transformation, Problems 2–4 were two transformations, and Problem 1 was three transformations from the complex example.

Significant effects were found for instructional method, $F(5, 168) = 10.96$, $MS_e = 0.386$, $p < .001$, transformations, $F(3, 504) = 44.91$, $MS_e = 0.072$, $p < .001$, and their interaction, $F(15, 504) = 7.65$, $MS_e = 0.072$, $p < .001$. The percentage of correct equations for each of the instructional groups across the four transformations was 7% for the procedures, 32% for the complex example, 38% for the simple example, 45% for the complex example and procedures, 47% for the simple example and procedures, and 65% for the simple and complex examples. According to a Neuman-Keuls analysis, the procedures group performed substantially worse than all other groups at the $p < .01$ confidence level. The group that received two examples performed substantially better at the $p < .01$ level than the groups that received either the procedures, simple example, or complex example and substantially better at the $p < .05$ level for the groups that received the procedures and an example. None of the other paired comparisons was significant.

As mentioned previously, we accidently changed one word in Problem 8. Problem 8 then stated that John, rather than Paul, worked 1 hr longer. To compensate for this change, we did a second analysis of the equations for Problem 8 and scored them as correct if they were correct except for interchanging the variables h and $h + 1$. This was the way in which the variables were assigned in the complex example and the other test problems. This change had virtually no effect on the three instructional groups that did not receive a complex example, but increased the overall score by approximately 5% for those groups that received the complex example. The rescored means were 7% correct for the procedures, 38% correct for the simple example, 38% correct for the complex example, 48% correct for the simple example and procedures, 49% correct for the complex example and procedures, and 70% correct for the two examples. The F tests and Neuman-Keuls analysis of the rescored data produced the same pattern of results, at the same confidence levels, as reported above.

Predictions of the model. Because Problem 8 should be zero transformations from the complex example, we used the more lenient criteria of allowing reversed assignments of h and $h + 1$ when judging the correctness of an equation. Figure 2 shows the generalization gradients for four of the instructional groups, in which transformations are measured from the simple example. Figure 3 shows the transformations for the other two groups, in which transformations are measured from the complex example.

The most extensive version of our model contains six parameters to account for the 24 data points in Figures 2 and 3. The parameters result from crossing the three operations described previously with two levels of complexity. We again used multiple linear regression to estimate the parameters, which accounted for 94% of the variance as measured by the square of the correlation coefficient. The six parameters involve generating a correct value by either of the following: matching the simple example ($m = .97$), matching the complex example ($m = .91$), applying a rule to a simple transformation ($r = .66$), applying a rule to a complex transformation ($r = .56$), applying general knowledge to a simple transformation ($g = .64$), or applying general knowledge to a complex transformation ($g = .52$).

The application of the model to the six instructional groups is shown in Table 4. Consistent with Figures 2 and 3, transformations are measured from the simple example (Figure 2) except for the complex example and the complex example

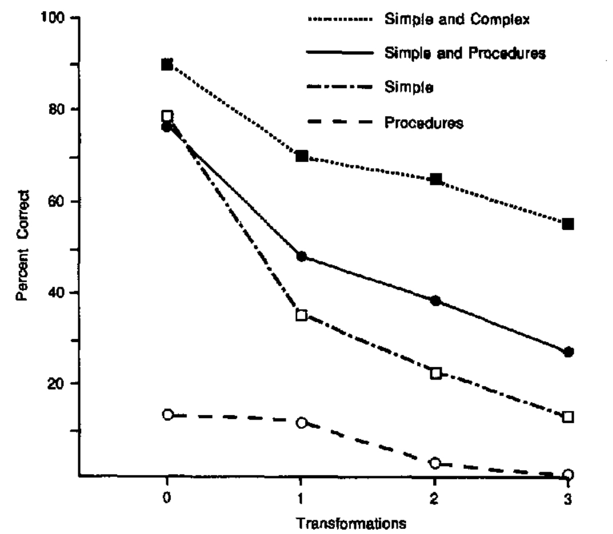

Figure 2. Percentage correct equations for 4 of the instructional groups over the 4 levels of transformation in Experiment 2. (Simple = simple example; complex = complex example.)

and procedures conditions (Figure 3). The model predicts that the probability of generating a correct equation for the procedures group ranges from .13 for the simple test problem to .08 for the complex test problem, in which three of the five quantities are complex.

Predictions for the two single-example conditions follow the same pattern as in Experiment 1 by replacing pattern matching with general knowledge for each transformation. The parameters $m = .97$ and $g = .52$ were used for the simple example and $m = .91$ and $g = .64$ were used for the complex example. Notice that transformations from the simple example produce more complex problems, whereas transformations from the complex example produce simpler problems.

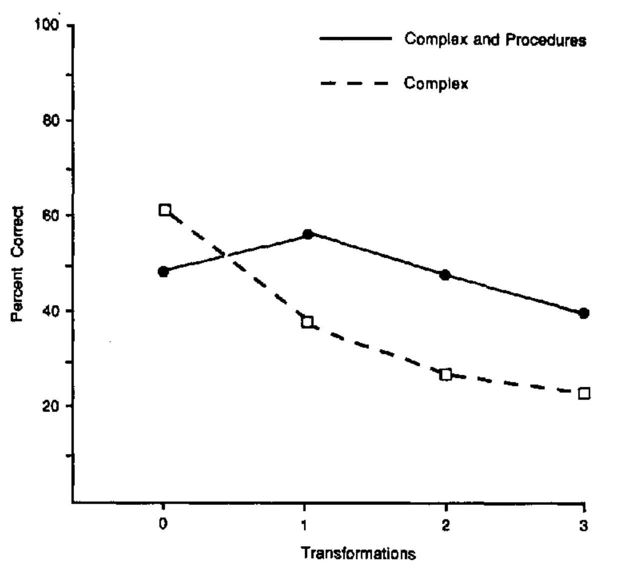

Figure 3. Percentage correct equations for 2 of the instructional groups over the 4 levels of transformation in Experiment 2.

Table 4
Observed and Predicted Values for Six Groups in Experiment 2

Transformations	Procedures			Simple example			Complex example		
	Observed	Predicted	Model	Observed	Predicted	Model	Observed	Predicted	Model
0	14	13	$\boldsymbol{r}^5$	79	85	m^5	62	63	$\boldsymbol{m}^5$
1	12	11	$r^4\boldsymbol{r}$	36	45	$m^4\boldsymbol{g}$	37	44	$\boldsymbol{m}^4g$
2	3	9	$r^3\boldsymbol{r}^2$	24	24	$m^3\boldsymbol{g}^2$	28	31	$\boldsymbol{m}^3g^2$
3	0	8	$r^2\boldsymbol{r}^3$	14	13	$m^2\boldsymbol{g}^3$	24	21	$\boldsymbol{m}^2g^3$
	Simple example and procedures			Complex example and procedures			Simple and complex examples		
0	76	85	m^5	48	63	$\boldsymbol{m}^5$	90	85	m^5
1	48	49	$m^4\boldsymbol{r}$	57	46	$\boldsymbol{m}^4r$	70	79	$m^4\boldsymbol{m}$
2	39	29	$m^3\boldsymbol{r}^2$	48	33	$\boldsymbol{m}^3r^2$	66	67	$\boldsymbol{m}^4m$
3	28	17	$m^2\boldsymbol{r}^3$	41	24	$\boldsymbol{m}^2r^3$	55	63	$\boldsymbol{m}^5$

Note. The predictions are based on parameter estimates of $m = .97$ (matching the simple example), $\boldsymbol{m} = .91$ (matching the complex example), $r = .66$ (applying a rule to a simple transformation), $\boldsymbol{r} = .56$ (applying a rule to a complex transformation), $g = .64$ (applying general knowledge to a simple transformation), and $\boldsymbol{g} = .52$ (applying general knowledge to a complex transformation).

Correspondingly, the rule application parameters ($\boldsymbol{r} = .56$ and $r = .66$) replace the general knowledge parameters to predict the generalization gradients for the simple example and procedures and the complex example and procedures groups.

Although the model provides fairly accurate predictions for the simple and the complex examples, it predicts gradients that are too steep when the examples are combined with the procedures. The problem is that students perform so poorly when they receive only the procedures that the parameter estimates for correctly applying a procedure are too low to predict the improvement that occurs from having both an example and the set of procedures. In this particular case, the rule application and general knowledge parameters are sufficiently similar that the number of parameters could be reduced from six to four without having much effect on the accuracy of the predictions.

The final result is the high level of performance that occurred when students received two examples. The generalization gradient in this case is consistent with the differential performance on the two examples. The model assumes that students use pattern matching to the simple example at zero transformations and pattern matching to the complex example at three transformations. At one transformation four of the quantities can be obtained from the simple example and one quantity can be obtained from the complex example. At two transformations (which is one transformation from the complex example) four of the quantities can be obtained from the complex example and one quantity can be obtained from the simple example. The high level of performance across transformations suggests that students are very efficient at performing pattern-matching operations, even when they must use more than one example.

Evaluation of the parameters. We again analyzed the probability of generating correct values for the different operations and instructional conditions. Table 5 shows these probabilities and the parameter estimates for simple and complex values. For the pattern-matching operation, the simple values are the data from subjects who have the simple example and the complex values are from subjects who have the complex example. Because the rule application and general knowledge operations apply to transformations, the subjects who have the simple example apply these operations to complex values and subjects who have the complex example apply these operations to simple values.

The parameter values are again generally higher than the observed values, particularly for subjects in the procedures group. Although subjects in the example and procedures group were adequate in applying the procedures, subjects who received only the procedures did poorly. We do not have an explanation of why this context effect occurred. Although we expected the modified procedures to be easier to use, our attempt to improve the procedures was clearly not successful.

Table 5
Observed and Predicted Probability of Generating Correct Values

Operation	Instruction	Simple value	Complex value
Pattern matching	Example	.79	.69
	Example and procedures	.79	.79
	Two examples	.90	.87
	Parameter value	*.97*	*.91*
Rule application	Procedures	.31	.28
	Example and procedures	.65	.55
	Parameter value	*.66*	*.56*
General knowledge	Example	.53	.40
	Parameter value	*.64*	*.52*

Note. Parameter values are in italics.

This failure is partially compensated for the by the success of the group that received two examples. We expected that this group would do well because of the success of pattern matching in Experiment 1. However, we still had some reservations because two examples require that students select the appropriate example when the values differ in the two examples. The pattern-matching data in Table 7 show that subjects generated more correct values when they had two examples than when they had only a single example. Having two examples actually improved pattern matching, perhaps because there was greater opportunity to practice using this skill.

Transformations. This analysis extends the data in Table 4 showing how individual subjects perform across transformations. The data in the bottom half of the table are from the four groups shown in Figure 2. Because we are measuring transformations from the simple example, these data are from the same groups of problems analyzed in Experiment 1.

The pattern of data is similar to the pattern obtained in Experiment 1. Again a high percentage of subjects (82 of 97, or 85%) who solved both single-transformation (1t) problems also solved the corresponding double-transformation (2t) problem. Also, as predicted by our model more subjects solved the 2t problem without solving both 1t problems (29) than there were subjects who solved the two 1t problems without solving the 2t problem (15). A similar pattern occurred for the two groups shown in Figure 3 in which transformations are measured from the complex example. Eight subjects solved both 1t problems without solving the 2t problem, 12 subjects solved the 2t problem without solving both 1t problems, and 52 subjects solved all three problems.

These data support the conclusion from Experiment 1 that there is not a cognitive overload for the 2t problems. However, as in Experiment 1, there is a suggestion of an overload on the 3t problem for the four groups in Figure 2. In contrast to our model, there were more subjects (8) who solved all three 1t problems but failed on the 3t problem than there were subjects who solved the 3t problem without solving all the 1t problems (5). Seventeen subjects solved all four problems, but the percentage of subjects who should have been able to solve the 3t problem based on their correct solutions of the 1t problems (17 of 25 or 68%) was lower than the percentage for the 2t problems.

The inclusion of a complex example in Experiment 2 provides the opportunity to examine whether the 3t problem would produce an overload if the transformations produced simple values. We analyzed the two groups in Figure 3 who had the complex example and found no support for the overload hypothesis. Twelve of the 13 subjects (92%) who solved the three 1t problems (Problems 5, 6, and 7) could solve the 3t problem (Problem 1). In addition, there were 7 subjects who solved the 3t problem without solving all three 1t problems. Overload therefore only occurred when subjects solved a 3t problem that had complex values.

Discussion

We discuss the instructional implications of our results in this section and then conclude by evaluating the proposed model in the general discussion. An encouraging aspect of the results is that there was little evidence of an overload with increased transformations. Although performance declined with more transformations, students who could do the appropriate single transformations could usually solve problems that involved combinations of these transformations. The one exception is the 3t problem with complex values, but even here the evidence is not overwhelming in supporting the overload hypothesis.

These findings are encouraging because John Sweller and his colleagues have recently demonstrated numerous instances in which overload is a serious limitation in problem solving. One set of problems included mathematics and science problems in which the cognitive demands of means–end analysis interfered with learning the solutions (Sweller, 1989). Another set of problems involved the demands of integrating text and diagrams when solving geometry problems (Sweller, Chandler, Tierney, & Cooper, 1990). Although we did not find an overload with our measures and materials, it is possible that a closer integration of the rules within the example would enhance performance, as Sweller and his colleagues found when they integrated text within the diagrams.

A disappointment from an instructional perspective was the inadequacy of the procedures. However, we believe that it would be premature to argue that we should abandon the procedure approach and always rely solely on carefully selected examples. It is possible that procedures may work better on other tasks or that someone may design a better set of procedures for this task. For instance, rules and examples were equally effective for instructing students about conditional reasoning (Cheng, Holyoak, Nisbett, & Oliver, 1986) and the law of large numbers (Fong, Krantz, & Nisbett, 1986). We do claim that the successful integration of rules and examples is an important issue because we need to find effective ways to increase transfer to test problems that are not equivalent to examples.

A particularly effective method of increasing transfer was the presentation of two examples. The success of this condition is consistent with Sweller and Cooper's (1985) emphasis on examples, but shows that it is not necessary to provide an example for each possible test problem. Our two examples provided enough information to solve each of the test problems, but students had to selectively use information from both examples to solve six of the eight problems. Students' success in using two analogous examples provides empirical support for the claim made by Spiro, Feltovich, Coulson, and Anderson (1989) that multiple analogies are often required to teach complex concepts.

It should be noted that this beneficial effect of two examples is a different effect from that obtained by Gick and Holyoak (1983) in their promotion of transfer from the fortress problem to Duncker's radiation problem. Transfer between these two isomorphic problems represents between-domain transfer in which the difficulty is caused by the lack of physical similarity between the concepts in the two problems. Comparing two examples in their research resulted in the creation of a more abstract convergence schema that facilitated finding corresponding concepts. In contrast, our research required within-domain transfer in which problems shared identical concepts. The difficulty was not in matching concepts but in instantiating quantities when they differed from the example. We consider this last point in greater detail as we examine how our model might be extended to other within-domain analogies.

General Discussion

This research was influenced by the formulation of schema-based models of problem solving in which attached procedures could be used to generate the values for slots in the schema (Bobrow & Winograd, 1977; Larkin et al., 1988; Winograd, 1975). The "slots" in our task refer to the different concepts in an equation that are replaced by values when students solve algebra word problems. We conclude with a final evaluation of the proposed model, including its possible extension to complex statistics and physics problems.

Evaluation of the Model

We designed Experiment 1 to evaluate a model of how students use examples, procedures, and general knowledge to construct equations for test problems that have different values than the example problems. The model assumes that students attempt to match concepts (Rate, Time, and Tasks) in the test problem to concepts in the example. If the values of matching concepts have the same structure, then this structure is copied for the test problem. Otherwise, students search the procedures, if available, or use general knowledge to construct the values. If students have only the procedures, they search the procedures for relevant information.

The stated order of using the different sources of information—example, procedures, and general knowledge—reflects the likely success of each source. According to our parameter estimates in Experiment 1, the probability of correctly constructing a value was .96 when using the example, .65 when using the procedures, and .45 when using general knowledge. The success of the model was demonstrated by the finding that it accounted for 94% of the variance for how three instructional groups would perform on test problems that differed from zero to three transformations from the example. In the second experiment we extended the application of the model to three additional instructional groups that included a complex example by itself or combined with either the procedures or a simple example. The model was again fairly successful in predicting the general pattern of results, accounting for 94% of the variance.

Although there were deviations from the predictions of the model, the deviations can often be useful in evaluating the simplifying assumptions of the model. For instance, the model assumes that the probability of correctly applying each of the three operations is independent of the amount of instructional material. The data from Experiment 1 supported this assumption, but two discrepancies in Experiment 2 suggested that additional material might be beneficial. Subjects did better in applying the rules when they had the example and did better in pattern matching when they had two examples. Both findings are the opposite of an overload prediction that increasing the instructional material would make students less efficient in using that material.

One caution in interpreting the parameter estimates is that our model assumes that students are using the instructional material. An estimate for correctly applying a procedure assumes that students attempted with limited success to apply the rules. If some students ignored the set of rules, the parameter (and level of performance) would underestimate the effectiveness of the rules for students who use the material. This criticism is less applicable to the pattern-matching parameter (because the parameter estimates are so high) and to the general knowledge parameter (because appropriate instructional material is not provided). We also do not want to imply that general knowledge is never used when students receive instructional material. The difference between the general knowledge parameter and the other two parameters reflects the usefulness of the instructional material relative to relying solely on general knowledge.

Extension of the Model

A motivating factor for this research was the previous finding that a single instructional example was insufficient when students had to solve algebra word problems that had values that differed from the example (Reed et al., 1985; Reed & Ettinger, 1987). Before considering whether the proposed model can be extended to other problems, it will be helpful to list the possible limitations of a single analogy. Spiro et al. (1989) described eight ways in which a single analogy can cause misconceptions including (a) the analogy may have missing properties, (b) the analogy may have too many properties leading to the creation of nonexisting slots in the target, and (c) the analogy may have misleading properties, causing an incorrect transfer of an instantiated value. This latter misconception fits our problems.

The first two misconceptions occur because the analogy has either fewer slots or more slots than the target. For example, a work problem in which there is only a single worker would have fewer slots, and a work problem in which there are three workers would have more slots than the two-worker problem. In principle, our model should be able to apply to these variations, in which transformations are defined as either changes in slots or changes in values. Reed, Ackinclose, and Voss (1990) found that students were more successful in applying a more inclusive solution, which has either more slots or a more complex value than the test problem, than in applying a less inclusive solution, which has fewer slots or a less complex value. We may therefore need to distinguish between whether the test problems add or delete slots when estimating parameters, as we distinguished between simple and complex values.

Another way in which within-domain problems can differ from each other is in the subgoals and procedures that are required to solve the problem. Catrambone and Holyoak (1990) defined a subgoal as an unknown entity (numerical or conceptual) that needs to be found to achieve a higher level goal. A method is a series of steps for achieving a particular subgoal. Their example of a subgoal for calculating the Poisson distribution of a random variable is to find the expected value of that variable. This task is similar to our task because it requires entering a value into an equation. However, the method for calculating this value may be considerably more complex than matching or generating values for algebra word problems.

In addition, some amount of search may be required to find the appropriate method. This is illustrated by how FERMI searches for a solution to a complex problem (Larkin et al., 1988, pp. 122–125). To solve an equation it was necessary to find an unknown value for current, which resulted in the subgoal of finding an unknown voltage. FERMI has several methods for finding voltage and searched for the appropriate one by determining which method satisfied the constraints of the problem.

These problems have a more complex level of organization than typical algebra word problems. They fit the general procedural-attachment framework in which procedures, attached to slots in a schema, compute values for those slots. But our model would require additional elaboration to apply to these more challenging problems, in which a procedure consists of a series of steps and requires the evaluation of several constraints before it can be evoked.

Alternative Models

We conclude by comparing the proposed schema model with an alternative model suggested by a reviewer, which we call the *independent sources* model. This alternative approach is designed to predict the performance of subjects who have two sources of information from the performance of subjects who have only a single source of information. Assuming that the probability of a correct solution by using an example is independent of the probability of a correct solution by using the procedures results in the prediction that:

$$Pr(E \text{ or } P) = Pr(E) + Pr(P) - Pr(E)Pr(P), \quad (3)$$

in which $Pr(E \text{ or } P)$ is the probability of generating a correct equation from either the example or the procedures, $Pr(E)$ is the probability of generating a correct solution from the example, and $Pr(P)$ is the probability of generating a correct solution from the procedures.

We used Equation 3 to predict the performance of the example and procedures group in Experiment 1 and the simple example and procedures group, the complex example and procedures group, and the simple and complex examples group in Experiment 2. In each instance we used the observed probabilities from the two groups that had a single source of information to predict the performance of the group that had both sources of information. The last column of Table 6 shows the predictions of the independent sources model (Equation 3), and the preceding column shows the previously described predictions of the schema model. The predictions of the models are of approximately equal accuracy when students combine the example with the procedures. The absolute deviation between predicted and observed values is 8% (schema model) versus 5% (independent sources model) for the first set of data, 8% versus 9% for the second set of data, and 15% versus 15% for the third set of data. The fourth set of data is from combining two examples, and the schema model (6% deviation) is more accurate than the independent sources model (12% deviation) in predicting these results.

An advantage of the independent sources model is that it is a simple model that does not require any parameter estimates. A disadvantage is that it does not make predictions for single sources of information because it uses these results for making predictions about combined information. It is also not a process model, without additional assumptions.

A difficulty for a process interpretation of Equation 3 is that subjects do not produce two solutions, which is implied by a strict interpretation of the equation. If students independently use the example and the procedures to produce two solutions, then they have to decide which one to submit. The simplicity of using Equation 3 to make predictions may therefore be partially overshadowed by complex assumptions regarding how subjects arrive at a final solution.

Table 6
Comparison of the Schema and Independent Sources Models

Information	Transformations	Observed (% correct)	Predictions	
			Schema	Independent
Simple example, procedures (Experiment 1)	0	82	81	86
	1	42	55	48
	2	30	37	32
	3	14	25	5
Simple example, procedures (Experiment 2)	0	76	85	82
	1	48	49	44
	2	39	29	26
	3	28	17	14
Complex example, procedures (Experiment 2)	0	48	63	67
	1	57	46	45
	2	48	33	30
	3	41	24	24
Simple and complex examples (Experiment 2)	0	90	85	84
	1	70	79	54
	2	66	67	52
	3	55	63	67

A major difference between the two models is that the schema model assumes that students start by using an example to construct an equation and consult either the procedures or the other example only to generate the transformed values. The model implies that subjects try to integrate the two sources of information whenever a single source is inadequate. In contrast, the independent sources model assumes that subjects sequentially use one source of information followed by the other source. The model suggests that subjects do not integrate knowledge from the two sources, but rather use them independently, and then decide which solution to submit. More detailed data, such as an analysis of search patterns and verbal protocols (see Payne, 1976) should therefore be a valuable addition to the quantitative models explored in this article.

References

Bobrow, D. G., & Winograd, T. (1977). An overview of KRL, a knowledge representation language. *Cognitive Science, 1,* 3–46.

Catrambone, R., & Holyoak, K. J. (1990). Learning goals and methods for solving probability problems. *Memory & Cognition, 18,* 593–603.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293–328.

Falmagne, J., Koppen, M., Villano, M., Doignon, J., & Johannesen, L. (1990). Introduction of knowledge spaces: How to build, test, and search them. *Psychological Review, 97,* 201–224.

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

Greeno, J. G. (1983). Forms of understanding in mathematical problem solving. In S. G. Paris, G. M. Olson, & H. W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 83–111). Hillsdale, NJ: Erlbaum.

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

Larkin, J. H., Reif, F., Carbonell, J., & Gugliotta, A. (1988). FERMI: A flexible expert reasoner with multi-domain inferencing. *Cognitive Science, 12,* 101–138.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction, 3,* 1–30.

Marshall, S. P. (1981). Sequential item selection: Optimal and heuristic policies. *Journal of Mathematical Psychology, 23,* 134–152.

Payne, J. W. (1976). Task complexity and processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16,* 366–387.

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.

Reed, S. K., Ackinclose, C. C., & Voss, A. A. (1990). Selecting analogous problems: Similarity versus inclusiveness. *Memory & Cognition, 18,* 83–98.

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

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

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

Smith, E. E., & Goodman, L. (1984). Understanding written instructions: The role of explanatory schema. *Cognition and Instruction, 1,* 359–396.

Spiro, R. J., Feltovich, P. J., Coulson, R. L., & Anderson, D. K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconceptions in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 498–531). Cambridge, England: Cambridge University Press.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81,* 457–466.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119,* 176–192.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2,* 59–89.

Thorndyke, P. W. (1984). Applications of schema theory in cognitive research. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory* (pp. 167–191). San Francisco: Freeman.

Viscuso, S. R., & Spoehr, K. T. (1986, November). *How does a mental model facilitate comprehension of instructions?* Paper presented at the 27th annual meeting of the Psychonomic Society, New Orleans.

Winograd, T. (1975). Frame representations and the declarative/procedure controversy. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding* (pp. 185–210). San Diego, CA: Academic Press.

Appendix A

Problems Used in Experiments 1 and 2

Simple Example

A. Ann can type a manuscript in 10 hours and Florence can type it in 5 hours. How long will it take them when they can both work together?

Complex Example

B. Jill can complete an audit in 12 hours and Barbara is three times as fast. They both complete 1/8 of the audit before being interrupted. How long will it take Jill to complete the audit if she and Barbara work together but Barbara works 2 hours longer?

Test Problems

1. Bob can paint a house in 12 hours and Jim can paint it in 10 hours. How long will it take them to paint a house if they both work together?
2. Susan can sew a dress in 9 hours and Sherry is three times as fast. How long will it take them to sew a dress if they both work together?
3. An expert can complete a technical task in 5 hours but a novice requires 7 hours to do the same task. When they work together, the novice works 2 hours more than the expert. How long does the expert work?
4. Bill can mow his lawn in 4 hours and his son can mow it in 6 hours. How long will it take both to finish mowing the lawn if they have already mowed 1/3 of it?
5. Jack can build a stereo in 8 hours and Bob is four times as fast. When working together to build a stereo, Bob works 1 hour more than Jack. How long does Jack work?
6. Tom can clean a house in 4 hours and Stan is twice as fast. They clean 1/4 of the house in the morning. How long will it take them to finish cleaning if they continue to work together?
7. A carpenter can build a fence in 7 hours and his assistant can build a fence in 10 hours. On the previous day they built 1/2 of the fence. How long will it take the carpenter to finish the fence if he and his assistant work together, but the assistant works for 3 hours more than the carpenter?
8. John can sort a stack of mail in 6 hours and Paul is twice as fast. They both sort 1/5 of the stack before their break. How long will it take John to sort the remainder if he and Paul work together, but Paul works 1 hour longer?

Appendix B

Procedures for Solving Work Problems

Work problems typically describe a situation in which two people work together to complete a task. The following equation can be used to solve these problems:

$$\text{Rate}_1 \times \text{Time}_1 + \text{Rate}_2 \times \text{Time}_2 = \text{Tasks Completed},$$

where $\text{Rate}_1 \times \text{Time}_1$ is the amount of work completed by the first worker, $\text{Rate}_2 \times \text{Time}_2$ is the amount of work completed by the second worker, and Tasks Completed is the total work completed by both workers.

These rules should be used for entering values into the equation.

Rate

1. The rate specifies how much of a task is completed per unit of time. If this value is known, enter it into the equation.
2. These problems usually state how long it takes to complete a task. The reciprocal of this number is then the rate. For example, if a worker needs 3 hours to complete a task, he will complete 1/3 of the task in 1 hour.
3. If rate is unknown, use a variable to represent it. Be sure to represent the relative rate of the workers. If one worker is 4 times as fast as the other, their rates will be r and $4r$.

Time

1. Time refers to the amount of time each worker contributes to the task. If this value is stated in the problem, enter it into the equation. For example, if one person works for 5 hours, enter 5 hours into the equation for that worker.
2. Time is often the unknown variable in these problems. Be sure to represent the correct relative time among workers if they do not work for the same time. If one worker works 3 hr more than the value (h) you are trying to find, enter $h + 3$ for that worker.

Tasks Completed

1. The number of tasks completed is usually 1 but the number may be greater than 1, or even less than 1 if part of the task is already finished.

Appendix C

Modified Procedures for Solving Work Problems

Work problems typically describe a situation in which two people work together to complete a task. The following equation can be used to solve these problems:

$$\text{Rate}_1 \times \text{Time}_1 + \text{Rate}_2 \times \text{Time}_2 = \text{Tasks Completed},$$

where $\text{Rate}_1 \times \text{Time}_1$ is the amount of work completed by the first worker, $\text{Rate}_2 \times \text{Time}_2$ is the amount of work completed by the second worker, and Tasks Completed is the total work completed by both workers.

These rules should be used for entering values into the equation.

Rate

1. The rate specifies how much of a task is completed per unit of time. If the problem states how long it takes someone to complete a task, the reciprocal of this number is the rate. For example, if a worker needs 3 hours to complete a task, she will complete 1/3 of the task in 1 hour.

2. The rate of one worker is sometimes expressed relative to the rate of another worker. If one person can complete a task in 10 hours, a person who is twice as fast can complete the task in 5 hours. The first worker can therefore complete 1/10 of the task in 1 hour and the second worker can complete 1/5 of the task in 1 hour.

Time

1. Time refers to the amount of time each worker contributes to the task. If this value is stated in the problem, enter it into the equation. If you have to find how long someone works, represent this quantity as a variable, such as h for h hrs.

2. Be sure to represent the correct relative time among workers if they do not work for the same amount of time. If one worker works 3 hours more than the value (h) you are trying to find, then enter ($h + 3$) for this worker. Place parentheses around this quantity because the entire quantity, $h + 3$, is multiplied by the rate.

Tasks Completed

1. The number of tasks completed is usually 1 but the number may be greater than 1, or even less than 1 if part of the task is already finished.

2. If part of the task is already finished, subtract that amount from the total number of tasks. For example, if there is 1 task and 1/4 of it is finished, then there are $1 - 1/4$ tasks to complete.

Received January 8, 1990
Revision received December 19, 1990
Accepted December 27, 1990 ■