8

Learning From Examples Via Self-Explanations

Michelene T. H. Chi[1]
Miriam Bassok
Learning Research and Development Center
University of Pittsburgh

Numerous topics are involved in the study described here: learning, problem solving, self-explanations, the role of examples in learning, individual differences, and physics. However, three issues are overarching in our domain of inquiry and in the task and the design we used—namely, how one learns, what one learns, and how one uses what has been learned. This

[1]We view this chapter as an interim report of ongoing work on which Bob Glaser has been a collaborator since 1978. Hence, in describing this work, we are describing a part of Bob's research. In the first half of our decade of collaboration, Bob and I had been investigating the nature of expertise in complex problem solving. Our interest has shifted because it has become clear that we cannot understand the acquisition of expertise without first understanding *how one learns,* even the beginning novice. Most importantly, we need to understand how individuals differ in the way they learn. Any differences that we might detect could contribute to our knowledge of what enables a person to become an expert. Although the study of learning and individual differences represents a new direction for me, these topics have been a long-standing interest of Bob's (see Glaser, 1967).

chapter first describes the rationale for the particular design that we used; it then gives a brief description of the procedure of our learning study. A discussion of these three overarching issues, and of how our study can provide evidence to address them follows. In this way, we give an overview of our current state of understanding of the three issues and the degree to which the findings reported can or cannot elucidate our understanding.

RATIONALES FOR THE DESIGN OF THE STUDY

Two major goals guided the design of this study. The first was to study how one learns to solve problems, rather than to study individual differences in problem solving abilities per se. The latter approach, usually contrasting the problem-solving successes of good and poor solvers, or experts and novices, sheds light only on performance differences. By focusing instead on what students have learned *prior* to solving problems, as well as what they learn *while* solving problems, we can gain further knowledge about the transition mechanism, thus providing insights for a model of competence in problem solving.

A second concern that guided the design of this study was to approach learning as a constructive process. That is, in learning to solve problems, one must declaratively encode and store new knowledge in terms of new concepts and principles, build problem schemata, and attach procedures (or their derivations) to the problem schemata. Our interest was, thus, in the construction of this knowledge, rather than in the proceduralization of already encoded knowledge.

In the sections that follow, we elaborate on how these goals were met by the design of our study. Briefly, in order to study how one learns to solve problems, we examined the way students learned from worked-out examples in the text. The constructive aspect of their learning was then assessed via the explanations that they gave to themselves.

Contrasts With Other Approaches to Problem Solving

A customary approach to the study of problem solving is to observe how people with different skills (such as experts and novices) solve problems by collecting and analyzing their protocols and formulating models to capture their solution processes. The models of performance thus formulated constitute the knowledge of a set of procedures that each of the students possesses and can apply for solving problems. Individual differences are then explained by the differences in the knowledge possessed, as embodied by the sets of production rules or programs. The intention of these models has been to see if one could derive the knowledge of the skilled solver from the knowledge of the unskilled solver. Unfortunately, inferences about the transition to the expert's from the novice's knowledge have not been straightforward. An alternative approach—the one we have chosen—is to see what students acquire from studying, in hope of being able to represent the knowledge underlying the generation of the solution procedures for the skilled and less skilled students.

To contrast our primary interest with other researchers' contributions to the literature on problem solving, we can describe an alternative focus: how the declarative knowledge encoded from text or from the teacher's instruction becomes proceduralized into a skill. This conversion process dominates, for example, Anderson's theory (1987) of skill acquisition.

In that theory, the process of conversion is achieved by using general weak methods that can convert declarative knowledge into domain-specific procedures via the mechanism of compilation. In Anderson's theory, compilation consists of two processes: proceduralization and composition. Composition is analogous to chunking in that separate procedures can be concatenated into one larger procedure. Proceduralization, by contrast, converts declarative knowledge into a procedure. This is accomplished by applying a weak method, which takes as the condition certain declarative knowledge that has been encoded a priori as important and fundamental to the problem at hand. For instance, suppose a student is trying to solve a geometry problem that requires proving that two triangles are congruent. The student might start by applying a weak method, such as means-ends analysis, in which the goal is to transform the current state (the givens) into the goal state. To do so, the weak method sets out to reduce the largest difference between the current state and the goal state. The largest difference that the model notices is that "no corresponding sides have been proven congruent." Therefore, the model attempts to reduce this difference. However, for the model to detect that "there are no corresponding sides that have been proven congruent," it must contain the declarative knowledge that *if the corresponding sides are congruent, then one can prove the triangles to be congruent*. Thus, Anderson's theory assumes that all relevant declarative knowledge has been properly encoded. The aspect of problem solving that Anderson's theory focuses on is the compilation of the declarative knowledge in the context of a derivation of the set of actions that constitute the solution.

We are not concerned with the mechanism of compilation per se. In a complementary way, our research focuses on how students can successfully solve problems as a function of whether or not they have encoded the relevant declarative knowledge and the level to which this declarative knowledge is "understood," as revealed by the way it relates to the procedures that instantiate it. Thus, in our view, problem solving, in the sense of organizing the derivation of a set of actions, does not seem to be guided by such general heuristics as means-ends analysis (at least for the good solver). Rather, we believe that the organization of a derivation is guided by a general domain principle, which then governs the set of actions that should be derived in order to solve the problem. For example, in simple mechanics problems, if the solver knows the problem is a static one, then his set of actions (the solution) is guided by that piece of declarative knowledge and its associated consequence, namely, that the sum of all the forces acting in the system is zero. This knowledge is a specific rule that can be inferred from one of Newton's principles. The rule then guides the derivation of the set of actions, which consists of finding all the forces acting on the body, decomposing them into their respective axes, and summing the ones in the same dimension. Thus, the derivation is an implementation of the knowledge that "All the forces must sum to zero." With such a framework, the critical issues are how and to what extent students understand the declarative principles that govern the derivation of a solution.

LEARNING IN THE DOMAIN OF MECHANICS

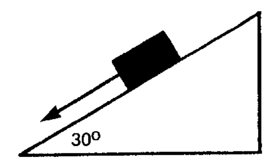
Because we view problem solving as the organization of a derivation of a set of actions, and this derivation is guided by a general principle, we necessarily have to choose a domain in which there is a rich interrelation between the set of actions that are executed as the solution and the principles that guide the selection of the set of actions. This constraint necessarily rules out knowledge-lean puzzle-type problems. Many knowledge-rich domains, on the other hand, are still too syntactic and algorithmic in nature. Factoring algebraic expressions or subtractions, for example, are too algorithmic in that the students simply have to learn a set of procedures. The procedures can often be performed correctly only with syntactic understanding of the problems. There are many knowledge-rich domains that have a less algorithmic character, but some require diagnostic problem solving, such as troubleshooting, which is quite different from the traditional classroom-type problem solving that interests us. Simple mechanics problems, however, seem to fall in-between the two extremes. Mechanics problems satisfy our constraint in that successfully solving them demands the instantiation of principles and concepts.

Our choice of the domain of mechanics was also guided by additional considerations. First of all, mechanics problems are difficult for students to learn to solve. One major source of this difficulty, we surmise, is the fact that a mechanics problem, even the simplest one, requires that certain forces not explicitly described in the physical situation of the problem be added. Furthermore, not only must these forces be recognized, but their interrelationships must be considered. In other words, the equations or

procedures that apply should not relate the entities explicit in the problem statement (the block on the incline, the angle of the incline, etc.). Rather, the equations should relate the interplay of the forces at work. (In the inclined plane case, it would be the forces that act on the block from gravity, from the string pulling on the block, and so forth.) Whether or not the block moves depends on the total sum of the forces acting on it in a given direction. Hence, a representation of the implicit relations among the forces must be constructed in order to solve such problems.

To illustrate, Fig. 8.1 shows a diagram of an inclined plane problem and the content of a *basic representation.* We propose that some students (probably the poorer ones) have only a basic representation of the problem. Such a representation consists mostly of explicit entities (the block, the inclined plane), so that the relations represented are among the explicit objects (i.e., the block is on the inclined plane, the plane is inclined at 30 degrees).

BASIC REPRESENTATION



(1) EXPLICIT CONCRETE OBJECTS (INCLINED PLANE BLOCK)

(2) EXPLICIT RELATIONS AMONG CONCRETE OBJECTS (BLOCK IS ON TOP OF THE PLANE; PLANE SLOPES AT 30° ANGLE)

(3) OPERATORS THAT RELATE EXPLICIT ENTITIES (F=MA)

(4) GOAL (FIND THE MASS)

FIG. 8.1. Content of a basic representation.

Moreover, because such a representation contains only the explicit objects and their relations, the operators necessarily act on these alone, as the following protocol excerpt from a poor solver confirms:

What is the mass of the block?	(The question posed by the problem statement.)

Let me see. I know $F = ma$.	(Her operator is an algebraic one, embodied in the equation that F relates to m and a.)
What's m equal to?	(The solver is trying to find m by deriving it algebraically by dividing F by a.)

The components of her basic representation contain entities (the force, the acceleration, and the mass of the block) that are explicitly described in the problem statement, and the operators that the poor solver has learned (algebraic ones, such as $F = ma$ can be converted to $m = F/a$) can only be applied to the explicit entities and relations.

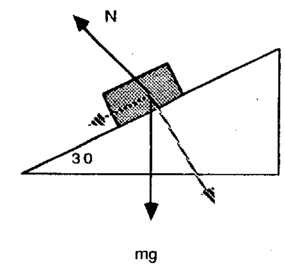
A more successful solver, however, has a *physics representation* that includes, in addition, generated physics entities that are not explicitly described. For the same simple physical situation as depicted in Fig. 8.1, she would have, in addition, inferred implicit entities (gravity force, normal force, etc.) and their relations (the directions of these forces, and whether they oppose and cancel each other out). Figure 8.2 shows the implicit entities and relations (Items 3 and 4), as well as the explicit ones. The *macro-operator* ($\Sigma F = ma$), (Item 5 in Fig. 8.2) in this case, is really a guiding principle, and the actual proceduralized *micro-operators* are the component procedures of decomposing forces, summing forces in the same dimension, and so forth. Hence, one might even argue that the operators in the two representations (the basic and the implicit physics) are different. The necessity of generating implicit entities under the guidance of a domain principle is a fundamental difference between solving problems as algebraic expressions and doing physics. That is, to solve a mechanics problem successfully, a student must generate an implicit representation. We believe this requirement contributes largely to the difficulty in solving physics problems. In tracing the learning processes, we may be able to find out how a transition from explicit (basic) to implicit (physics) representation takes place.

Another reason for studying mechanics problems is that textbooks usually use a general problem-solving procedure as a guideline for teaching problem solving. A general procedure is not only laid out in the textbooks, but it is often taught directly by the instructor as well. It usually encompasses some set of subprocedures, as indicated in Halliday and Resnick (1981, p. 69):

1. draw a sketch of the explicit physical situation;
2. draw a free-body diagram (in other words, add the forces which are operating in the physical situation);
3. resolve the force vectors into their components (that is, calculate their projections onto the coordinate axes, which also requires selecting coordinate axes);
4. select the appropriate equation and substitute the components into the general equation to produce an instantiation corresponding to each coordinate axis;
5. solve the equations.

As one can see, such a general procedure serves only the purpose of setting subgoals, such as "draw a free-body diagram." Setting subgoals using this guideline is not very different from using a general heuristic, such

PHYSICS REPRESENTATION



(1) EXPLICIT CONCRETE OBJECTS (SUCH AS THE INCLINED PLANE AND THE BLOCK)

(2) EXPLICIT RELATIONS AMONG THE EXPLICIT OBJECTS (THE BLOCK IS ON TOP OF THE PLANE)

* (3) IMPLICIT ABSTRACT OBJECTS THAT MUST BE INFERRED (GRAVITY FORCE, NORMAL FORCE)

* (4) RELATIONS AMONG THE IMPLICIT OBJECTS (A COMPONENT OF THE GRAVITY FORCE OPPOSING THE NORMAL FORCE)

* (5) OPERATORS THAT ARE INSTANTIATED ON THE INFERRED OBJECTS AND RELATIONS (ΣF=MA)

(6) GOAL (FIND THE MASS)

FIG. 8.2. Content of a physics representation.

as means-ends analysis, to set up subgoals. But if the problem were a static one, then one would set up the subgoals of seeking all the forces acting on the system and balancing the forces so that there is no resulting net force in any direction. A guideline that is not driven by such a domain principle can only serve the function of a syntactic heuristic that cannot be carried out without knowing the subgoal's purpose in the context of the principle. One cannot simply draw a free-body diagram unless one knows what forces to depict.

Our conjecture that a general procedure alone is not sufficient for solution can be seen in our protocol from a poor solver in our pilot study. The student would often prompt herself with these subgoals. For example, she would often say to herself, "I know I'm supposed to draw a free-body diagram here." She still, however, did not get anywhere with the solution. On the other hand, on occasions where students do succeed in solving the problems, one can consistently see a trace of this general procedure. This means that a successful solution procedure often proceeds through the steps of constructing a free-body diagram, finding and decomposing the forces, and so forth, but these steps do not necessarily serve as a guideline for finding the solution. Hence, if students are not learning problem solving by using this guideline of procedural steps, how exactly are they learning to solve problems?

Simple mechanical problems are suitable in another way. They are very tractable in the following sense: We can discriminate between procedures that are situation-specific to a particular schema (such as procedures for solving an inclined plane problem), versus the principle ($F = ma$) that underlies this solution procedure. That is, all three types of problems that we study (inclined plane, pulley, and rope problems) are instantiations of the application of Newton's Laws of Motion, especially Newton's second law. However, this particular law cannot be used in a direct syntactic way (as the poor solver mentioned earlier did) by substituting the mass of the block for m, the acceleration of the block for a, and the force of the block for F. $F = ma$ applies to the resulting net force of the whole system. Thus, we believe that mechanics problems may be particularly suitable for teasing apart different levels of understanding, both between students and at various stages of the learning process.

Finally, the most intriguing aspect of studying mechanics problems relates to the issue of *understanding*. A strict criterion of understanding requires the depiction of the knowledge that is acquired that can generate the solution procedure from deep domain principles. VanLehn (in press) refers to such knowledge as *teleological understanding*: Students who have teleological understanding could derive a procedure for solving a problem. However, one could conceive of all mechanics solution procedures as a kind of derivation. And yet, knowing (in the sense of having studied) a specific derivation that constitutes the solution to a problem may not indicate that the student understood the problem solution in a teleological sense, that is, in the sense that it is derived from some domain principle. Hence, we clearly need to distinguish teleological understanding from the issue of derivation. Problems in mechanics could conceivably provide insight into the issue.

EXAMPLES AS A SOURCE OF LEARNING

We have chosen to examine how problem-solving success is mediated by what is learned from worked-out examples. There are multiple reasons for focusing on examples. An instructional reason is that examples seem to be the primary tool which textbook writers and instructors rely upon to teach students how to solve problems. Furthermore, students themselves also prefer to rely on examples as a learning tool. In classroom findings, VanLehn (1986) has indirect evidence showing that 85% of the systematic errors collected from several thousand arithmetic students could be explained as deriving from some type of learning from example. In another classroom study, Zhu and Simon (1987) have shown a clear advantage (a 3-year course can be reduced to a 2-year math curriculum) if students are given only examples and problems to work on, as opposed to the standard instruction with a text and instructor's presentations.

In laboratory studies, there is also some scant evidence showing that students prefer to rely on worked-out examples. Anderson, Greeno, Kline, and Neves (1981) have claimed that students spend a considerable amount of time studying examples and that students often commit knowledge of worked-out examples to memory. Our own informal observation of two students solving problems after having read several chapters from a physics textbook confirm their claims. Upon the students' first encounter with problem solving, over 62% of the time, they either copied directly from the textbook examples or recalled a procedure from a worked-out example and used it as an analogical base. More recently, Pirolli and Anderson (1985) found that novices rely heavily on analogies to examples in the early stages of learning to program recursion. Eighteen of their 19 subjects used the example upon their first programming attempts. Not only do beginning students rely on examples, but they appear not to be able to solve problems without them. Pirolli and Anderson (1985) noted that "over the course of the four hours of protocols obtained from subject JP writing recursive functions, we saw no indication that she could write such functions without a lot of assistance from examples or the experimenter" (p. 262). Reder, Charney, and Morgan (1986) also found that the most effective manual for instructing students on how to use a personal computer are those that contain examples. Finally, LeFevre and Dixon (1986) found that subjects actually preferred to use the example information and ignored the written instruction when learning a procedural task. Hence, from an instructional point of view, worked-out examples seem to be a primary source of information from which students learn and acquire problem solving procedures. Thus, we thought it necessary that we ask what is learned from examples.

Despite the fact that examples have been shown to be a necessary tool for instructors, for textbooks, and for the students to rely upon as an instrument for learning, laboratory studies examining the usefulness of examples found severe limitations in the way that students can generalize what is learned from worked-out solutions. In almost all the empirical work to date, on the role of example solutions (Eylon & Helfman, 1982; Reed, Dempster, & Ettinger, 1985; Sweller & Cooper, 1985), a student who has studied examples often cannot solve problems that deviate slightly from the example solution.

Sweller and Cooper (1985) have clear-cut evidence showing that students who were given the opportunities to study example solutions had advantages over students who were just given opportunities to work the problems. These problems were algebraic expressions in which they had to isolate the unknown variable. The advantages were measured in terms of reduction in errors and speed. It is noteworthy that the group who was given the opportunity to study examples excelled on subsequent test problems, when these test problems were similar to the example problems. However, on dissimilar problems, there were no significant differences. These data suggest that example solutions tend to provide an algorithm for students to use and follow on similar problems; but they really have not acquired any deep understanding of the example. Similarly, in the Reed et al. (1985) study, students were able to solve only 6% of simple algebra word problems (such as mixture problems) that required a slight transformation of the original equation. For example, the mixture problem required the use of the formula:

$$(\text{Percentage}_1 \times \text{Quantity}_1) + (\text{Percentage}_2 \times \text{Quantity}_2) = \text{Percentage}_3 \times \text{Quantity}_3.$$

In the solution that the student studied, the application of this formula required that the third quantity (Quantity_3) be derived from the second quantity plus 2 ($\text{Quantity}_2 + 2$). Thus Quantity_3 can be represented as the sum of Quantity_2 and 2 quarts. In the similar test problem that the students had a great deal of difficulty solving, Quantity_1 had to be expressed in terms of 16 quarts minus Quantity_2. The point of this example is simply to show the triviality of the transformation that had to be derived from the example solution in order to solve the similar problem. The only conclusion from such findings is that the students must have learned a very syntactic structure in which they could only substitute numerical values into the unknown of the formula directly, because they had difficulty making even a slight transformation of the formula.

The pessimistic outcomes of empirical studies examining the role of examples and what can be derived from them direct our attention to the possible limitations of examples per se. One limitation seems to be that examples are inadequate at providing the rationales for the application of each of the procedural steps. That is, the solution procedure depicts a sequence of actions, without providing the specifications of the inputs that will produce such a sequence of actions (VanLehn, 1986). Simon (1980) noted that "the actions of the productions needed to solve problems in the domain of the textbooks are laid out systematically, but they are not securely connected with the conditions that should evoke them" (p. 92). We can provide a concrete example of the lack of specification of the explicit conditions under which actions should be executed by examining a worked-out solution taken from a physics textbook (Halliday & Resnick, 1981), as shown in Fig. 8.3. This is an exact solution example that we have used in our study. We can see clearly that explication of the rationale underlying the sequence of actions is often not given. It is not clear, in Statement 2, why one should "consider the knot at the junction of the three strings to be the body." This is a critical piece of information, because it implies that, at this location (as opposed to the block, for example), the sum of the forces is zero. Statement 2 also does not explain why and how other locations can be ruled out (such as the points of contact between the strings and the ceiling, etc). Such lack of specification of the explicit conditions for actions occurs throughout the example solution. In Statement 6, for example, how does the student know that $\vec{F}_A$, $\vec{F}_B$, and $\vec{F}_C$ are all the forces acting on the body and that there are no others? Statement 7 is essentially a restatement of Newton's first law, but it requires chaining several inferences together, and translating them into an equation (e.g., that because the body is at rest, there are no external forces; therefore the sum of the forces on the body must be equal to zero). Statement 8 is totally unexplained; why are the axes chosen as such? And in Statement 9, why is the vector equation translated into the scalars? (The reason is that one wants to resolve the forces so that they can be added after they are all aligned in the same x- and y- directions.) Our analysis of the example just presented strongly suggests that learning from examples is diminished if the statements in the example solution procedure are not explicit about the conditions under which the actions apply.

If students have learned only a sequence of actions, then they have basically acquired an algorithmic procedure, which is not readily amenable to transformation. This would explain why students in the empirical studies referred to earlier cannot transfer what they have learned from examples, primarily because what they have learned are syntactic or algorithmic

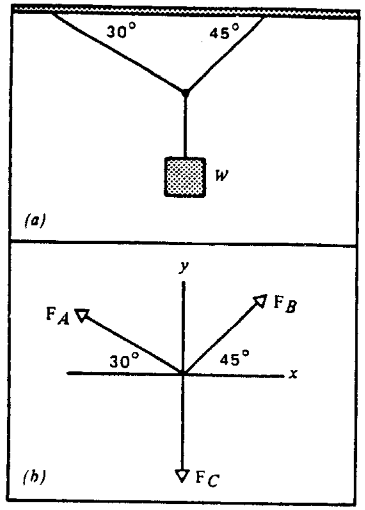


Figure 5-6 Example 5. *(a)* A block of weight *W* is suspended by strings. *(b)* A free-body diagram showing all the forces acting on the knot. The strings are assumed to be weightless.

1. Figure 5-6a shows an object of weight W hung by massless strings.
2. Consider the knot at the junction of the three strings to be "the body".
3. The body remains at rest under the action of the three forces shown Fig. 5-6.
4. Suppose we are given the magnitude of one of these forces.
5. How can we find the magnitude of the other forces?
6. F_A, F_B, and F_C are *all* the forces acting on the body.
7. Since the body is unaccelerated, $F_A + F_B + F_C = 0$
8. Choosing the x- and y-axes as shown, we can write this vector equation as three scalar equations:
9. $F_{Ax} + F_{Bx} = 0,$
10. $F_{Ay} + F_{By} + F_{Cy} = 0$
11. using Eq. 5-2. The third scalar equation for the z-axis is simply:
12. $F_{As} = F_{Bs} = F_{Cs} = 0.$
13. That is, the vectors all lie in the x-y plane so that they have no z components.
14. From the figure we see that
15. $F_{Ax} = -F_A \cos 30^\circ = -0.866F_A,$
16. $F_{Ay} = F_A \sin 30^\circ = 0.500F_A,$
17. and
18. $F_{Bx} = F_B \cos 45^\circ = 0.707F_B,$
19. $F_{By} = F_B \sin 45^\circ = 0.707\ F_B.$

FIG. 8.3. A strings example, taken directly from Halliday and Resnick (1981).

rules, much in the way Neves's program (1981) learns. Neves (1981) constructed a program that could learn how to solve simple algebraic problems by deducing rules from consecutive statements. To illustrate, in the example solution, suppose the following steps were given:

1) $3X + 4 = 0$
2) $3X = -4$
3) $X = -4/3.$

Neves's program can detect or compute the differences between one line and the next line, and deduce the rule (for the differences between lines 1 and 2): IF there is a number on the left hand side of an equation, THEN subtract it from both sides. Such a rule does not explain why or under what conditions one would subtract a number from both sides of the equal sign. The rule is syntactic in that it applies if the structure of the given equation matches the condition of the rule exactly. Thus, in the cases where the given structure of the equation has to be derived or transformed in order for the condition of the rule to be satisfied and the action taken, students will fail, because they have not understood how the conditions of the rule can be derived from other variations of a given equation.

In sum, instructionally, the basic observation is that examples are preferred by both teachers and students as an instrument from which to learn, and yet examples are often poorly constructed (in that the rationales are often not provided for their action sequences). Perhaps because of their poor construction, students sometimes gain very little understanding from examples, other than acquiring syntactic rules, so that the knowledge gained is not transferable to related problems, as reported in the empirical studies.

Besides the value of identifying instructional dilemmas in the use of examples, there are serious theoretical reasons for studying how one learns and what is learned from examples. The important theoretical issue that is addressed by the A.I. literature concerns the ease with which a generalization can be induced from a single example or whether multiple examples are necessary. A similarity-based approach claims that generalizations are developed by inducing a principle (or a set of common features) from multiple examples. Such a principle would embody the essential features shared by all the examples. Thus, it is necessary to provide more than one example in order for such induction to take place. On the other hand, an explanation-based approach (Lewis, 1988; Mitchell, Keller, & Kedar-Cabelli, 1986) claims that generalizations can be obtained from a single example. To do so, the system must possess knowledge of both the domain from which to construct an explanation, as well as a definition of the concept which the example instantiates. That is, the induction is the construction of an explanation that justifies why the example is an instance of the concept. Thus, the method of induction in the two kinds of systems are quite different. Although the majority of psychological theories, including Anderson's ACT*, had assumed a similarity-based approach to learning, empirically, students can often generalize from single examples (Elio & Anderson, 1983; Kieras & Bovair, 1986).

This theoretical debate in the A.I. literature about the mechanism of generalization aside, we conceive of human learning in quite different terms. Students clearly are not learning principles of mechanics, such as Newton's second law, by generalizing from multiple examples, because not only might the students not know a priori the set of critical features to look for across a set of examples but further, the generality that does emerge from a set of examples must necessarily be syntactic. Suppose we mentally simulate what may be similar across a set of worked-out solutions. The kind of features that are similar may be specific procedures such as decomposing forces, imposing a reference frame, summing the forces, and so forth. The summation of these procedural components may not necessarily constitute an understanding of the principle of the second law, even though understanding it does imply that it will generate a solution that embodies these procedural components. Neither can students learn strictly via explanations in the sense of Mitchell et al.'s (1986) model because this would require that students have, by the time they encounter an example, a complete understanding of the principle or concept it instantiates. Yet empirical evidence suggests that many students can learn to solve problems by studying worked-out examples only, without any background text or lectures on the principles and concepts (Zhu & Simon, 1987). Clearly, a combined similarity-based and explanation-based approach seems necessary in order to learn from examples. (An alternative view is to propose that in some domains, such as algebra, learning can be accomplished through similarity-based generalization, whereas in less syntactic domains such as mechanics, a more explanation-based approach is necessary.)

Although our research is not addressed directly to whether generalizations can be induced from a single or multiple examples, our data can provide evidence on whether and how generalizations can emerge from learning from a single or a few examples.

ROLE OF SELF-EXPLANATIONS

We have painted a fairly pessimistic picture of the role and characteristics of textbook examples. We have noted that examples are poorly constructed, in that rationales for each of the actions required are seldom spelled out. This omission results in poor generalization and transfer, as shown in laboratory studies. But then students often prefer to rely on

examples, to the exclusion of using textual elaboration of the relevant procedures and principles. Furthermore, there is also evidence showing that students often tend to ignore rationales that are given, as noted in algebra tutoring experiments (VanLehn, personal communication, December, 1987). It is also the case that some elaborations actually can disrupt students' understanding (Reder et al., 1986).

The only mediating factor that can account for these discrepant findings, we postulate, is self-explanations; which can play a significant role in effecting what can be learned from an example. We view examples as an essential instrument from which to learn because they instantiate the principles that the text aims to introduce. In other words, an example of a worked-out solution presents an interpretation of the principled knowledge presented in the text in terms of the procedural application. An example cannot be more complete without becoming redundant with the text of the chapter. When students fail to generalize from examples, perhaps we should attribute the failures not to the characteristic of the examples, but rather, to the disposition of the learners. In order to optimize learning, the students must actively construct an interpretation of each action in the context of the principles introduced in the text.

The role played by self-explanations can also explain why students sometimes ignore rationales that are provided, and why some rationales actually confuse and hurt the students' understanding. It seems likely that the explanations provided for the examples do not fit with the students' understanding. Self-generated explanations, by contrast are necessarily consistent with the students' own levels of understanding. Of course, then, the usefulness of self-generated explanations depends on the students' initial understanding of the text or principles. In any case, we can analyze students' self-explanations and the degree to which they are adequate in relating to the principles in the text. We should be able to see a relation between the adequacy of the self-explanations and successes at problem solving and transfer.

Thus, we view self-explanation as a mode of studying that can mediate learning. The degree to which self-generated explanations foster learning is a function of the accuracy and completeness of the self-explanations in interpreting the examples in terms of the principles introduced in the text.

A Brief Description of the Learning Study

The study that we conducted, which examined how students learn to solve problems by studying worked-out examples in the domain of physics, analyzed two major components: knowledge acquisition and problem solving (see Fig. 8.4 for a diagram of the experimental procedure). During the first part of the knowledge acquisition phase, students studied the necessary

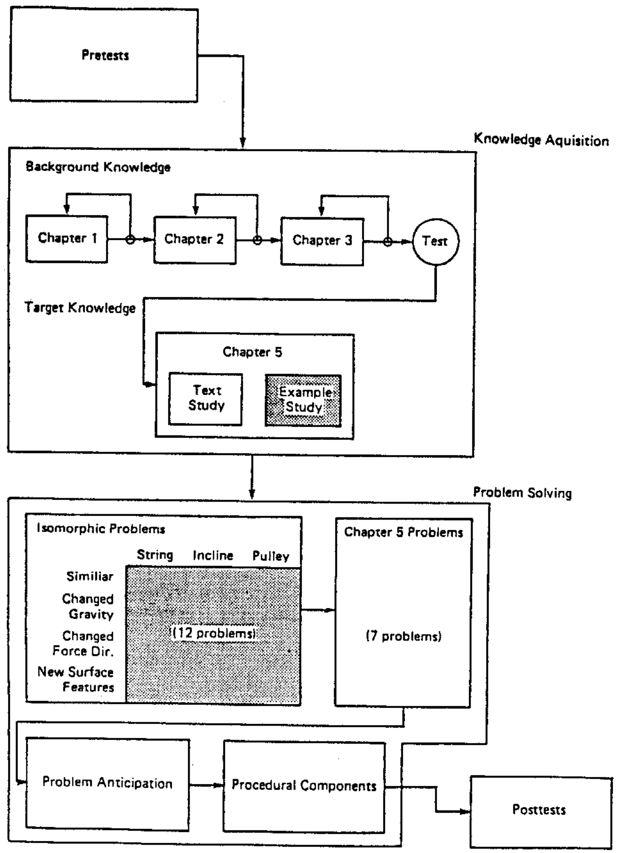


FIG. 8.4. A diagram depicting the design of the study.

background subject matter, covering the topics of measurement, vectors, and motion in one dimension. These materials are covered in the first three chapters of Halliday and Resnick (1981), a text that is adopted by many universities as a fundamental introduction to mechanics. With each of these chapters, students studied until they reached a criterion, which was defined as the ability to correctly solve a set of declarative, qualitative, and quantitative problems.

During the second part of this phase, students studied chapter 5, the target chapter on particle dynamics. Chapter 4 was skipped because it was not essential to understanding chapter 5. We stopped at chapter 5 because the examples there elicited interesting problem-solving behavior. That is, problem solving of those materials required some understanding of phys-

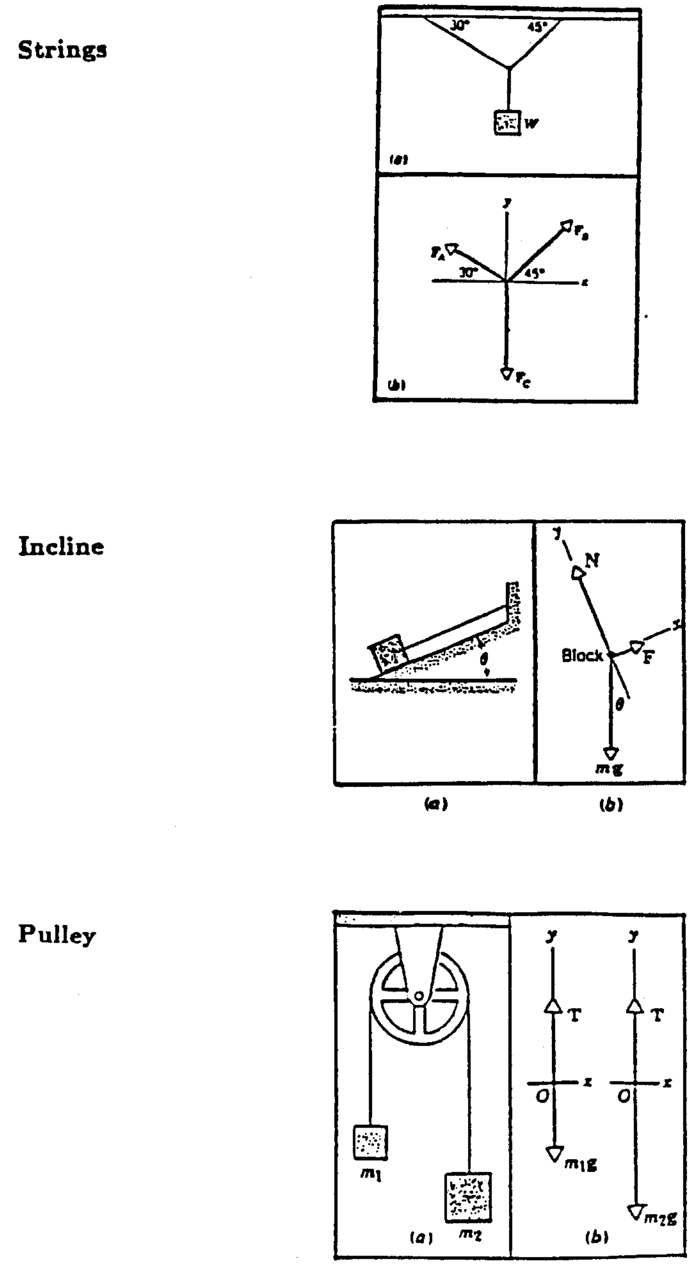


FIG. 8.5. Diagrams of the three worked-out examples used in the study.

ics, rather than merely applying algorithmic procedures. (The term understanding here is used in a general sense.) The major focus of the acquisition phase in their studying of chapter 5 was how students studied 3 worked-out examples taken directly from the text (see Fig. 8.3 for an example of a worked-out solution). Each worked-out solution represented a type of problem. There were three types, which we defined to correspond to the physical situation described: a string problem, an inclined plane problem, and a pulley problem. These are depicted in Fig. 8.5. Students talked aloud while they studied each example, and their protocols were taped. That is, they were asked to utter whatever they were thinking about as they read each statement of the example solution.

At the problem-solving stage, students solved two main sets of problems. The first set consisted of 12 isomorphic problems, with four problems corresponding to each type of example. These problems were specifically designed so that they had different degrees of similarity to the example solutions. For example, the first level of difference (the first set of isomorphs) was a problem that had fundamentally the same description and wording, except a different variable was required for the solution. To solve such a problem, one could basically follow the solution exemplified in the worked-out example, and compute the answer for a different variable than the one calculated in the worked-out example. The second level isomorphs varied only in the values of some of the variables stated in the problem statements. For example, the angle of the incline might be 60 degrees rather than 30. The third level isomorph had some dissimilarity in the surface features. For example, the orientation of the forces could be changed. The fourth level isomorphs had a completely different surface structure. For example, in the case of the strings example, instead of having two of the strings attached to the ceiling, the two strings might be attached to the floor.

The second set of problems to be solved by the students contained seven problems that were taken directly from the end of chapter 5. These problems are considerably more difficult than the isomorphic problems that we designed. Protocols were taken at all problem-solving sessions. A variety of pre- and posttests were also administered, the former before the knowledge acquisition stage and the latter at the very end of the problem-solving stage.

How One Learns From Studying Examples

In spite of the fact that all of our students successfully completed the declarative knowledge test that tapped the information presented in the text, we found substantial individual differences in their subsequent problem solving success, ranging from 28% to 100% correct. These could have

resulted from differences in how students studied the examples of worked-out problems and differences in the mechanisms and techniques they used to learn the procedural instantiation of that declarative knowledge.

Our methodological approach to the study of what and how students learn from examples was to ask them to comment on each of the statements that they read in the example. All the comments that the students made can be described by the generic term "elaborations." However, there are distinct differences between the kinds of elaborations that students generated: About a third of them related to the content of the example they were studying, and we refer to these as *self-explanations*. Another third were *monitoring* statements. A final third were classified as *miscellaneous* and these contained paraphrases and mathematical elaborations, primarily.

We now discuss students' self-explanations as a mechanism of learning by which students generated tacit knowledge that links pieces of explicitly stated knowledge. In the second section we then discuss the role of cognitive monitoring, that is, how students became aware of their failure or success at understanding, and how awareness of misunderstanding led to new attempts at understanding.

THE FUNCTION OF SELF-EXPLANATIONS

Explanations are remarks that students make after reading one or several statements from the worked-out solution example. In particular, these are remarks that have some relevance to the physics content under study. We reasoned that because example-exercises, as we have analyzed, are so incomplete in providing explanations for the action sequences, the students must necessarily construct their own explanations for the sequence of actions in order to understand the material. This idea is not new. Bruner (1966) mentioned it two decades ago and Wickelgren (1974) reiterated it. Indeed, we found that good students (i.e., those who had greater success at solving problems with an average of 82% correct) generated a significantly greater number of elaboration ideas (51 ideas per example) than the poor students (18 ideas). Although the good and poor students both had a similar distribution of types of elaborations generated (explanations, monitoring statements, and miscellaneous others), the good students still generated an average of 15.9 explanations per example, whereas the poor students generated only 4.3 explanations per example. Thus, generating self-explanations seems to be an important mechanism of learning from examples, because it correlated with problem-solving successes. Besides the quantitative difference, a second important finding was that the explanations generated by the good students tended to be qualitatively better than those generated by the poor students. That is, the good students' explanations tended to infer additional tacit knowledge, whereas the poor students' explanations were often paraphrasings of the diagram, with no new information generated.

To give a flavor of what we mean by better explanations, we quote one good and one poor student's explanation of the diagram of the strings example presented in Fig. 8.4. The good student's explanation qualified the nature of the forces involved and their interrelations: "They'd just be the force, the rest mass of the thing holding it up would be the force. It could, well, actually it'd be the force of weight. Cause being upheld by . . . it's the resistance to weight W." The poor student's explanation, on the other hand, basically translated into words the surface information presented in the diagram: "Okay, so three forces are on the two strings and from the string going down to the object."[2] (Further discussions of the nature and the quality of self-explanations is presented in the section dealing with what is learned from studying examples.)

We would like to interpret the more elaborated explanation statements generated by the good students as an overt manifestation of their active processing during learning (as opposed to taking it as evidence of their articulateness). Our interpretation is consistent with the assumptions made with other types of protocol analyses, in which the verbal protocols are supposed to reflect what is being processed in working memory. Thus, longer protocols should simply indicate a greater degree of processing. This assumption is consistent, as well, with those made in analyses of eye movement protocols, in which longer fixations are taken to mean that the student is processing a particular stimulus for a longer period. Although thinking-aloud is not quite the same as elaborating on a statement presented in an example solution, many studies of this variety have shown no differences in performance between those students who were instructed to think aloud and those who were not (Klinger, 1974; Thomas, 1974). Ericsson and Simon (1984) interpreted these findings to suggest that, in response to the instruction to think-aloud or verbalize, the subjects did not change the structure of their processing, but merely expressed overtly what they otherwise would have thought covertly. Thus, we view the greater amount of elaboration produced by the good students to be the manifestations of their internal processes in trying to understand the example solution better.

Our interpretation also discounts the possibility that the good students, by taking our instruction more seriously, engaged more fully in the activities of verbalizing explanations, and so achieved better understanding and thus more successful problem solving. This causality was not apparent when we analyzed the quality of the explanations. Our interpretation that explanations reflect active construction was further supported by the fact that the good students did not generate a greater number of explanations than the poor students during the problem-solving phase of the study (that is, during problem solving, many students re-read several lines of the examples, and they thus had additional opportunities to provide explanations of what they did not understand). Although they were encouraged to elaborate during problem solving as well, the good students probably did not need to. Hence, it does not seem to be the case that the good students were simply more motivated to generate explanations in order to please us.

[2]In subsequent analyses, we have discounted paraphrases of a diagram as an explanation (see Chi et al., in press), and instead, categorized them in the *miscellaneous* category.

The major support for our claim that the overt explanations reflect natural constructive processing, rather than cause more extensive processing, comes from the fact that the quality of the good students' explanations was better than the poor students'. It doesn't seem that one could obtain quality explanations by simply encouraging the poor students to generate them. The importance of the quality of the explanations generated is consistent with the results of a study by Stein and Bransford (1979). These authors asked subjects to generate their own elaborations in order to help them memorize some sentences. They found that students who generated their own elaborations, on the whole, did remember better than the control students (who were not told to elaborate), but worse than the group of students for whom the experimenter provided very precise elaborations. However, if the students were divided into those who generated precise elaborations (which is analogous to our better quality explanations) and those who generated imprecise elaborations, then those who generated precise elaborations remembered the sentences even better than the group of students who were given experimenter-generated elaborations. This result is especially consistent with our interpretation because it implies that asking students to generate elaborations per se does not necessarily produce better memory. Rather, elaborations generated were helpful only if they were precise.

There are other studies, however, that indicate that the very act of verbalization or providing justifications for solution steps while problem solving can lead to better problem solving. Gagné and Smith (1962), for example, found that by asking students to verbalize the reason for each move they made to the 2-, 3-, 4-, and 5-disk Tower of Hanoi task, they improved their performance on a subsequent 6-disk Tower of Hanoi task. However, the difference Gagné and Smith reported was a group difference: that is, the group with the instruction to verbalize performed better on the whole than the group not told to do so. They did not analyze the explanations as a function of how successful the solvers were. If we look at the content of their explanations, as did Gagné and Smith, we see that there were four types of explanations given (p. 15):

1. Those that were oriented toward single moves in the solution of the problems, with explanations such as:
   "only possible move"
   "just to try it"
   "don't know"
2. Those that anticipated to the extent of two moves, with comments such as:
   "to get at the larger disk"
   "to free up a space"
3. Reasons that anticipated sequences of moves, such as:
   "move as with a three-disk sequence"
   "if disk is odd-numbered, move to circle B"
4. Reasons that identified the principles underlying the solution, such as:
   "move odd-numbered disk in the clockwise direction"
   "move even-numbered disks in the counterclockwise direction"

It is possible that the good solvers in the Gagné and Smith study tended to give explanations that had the same characteristics as those generated by our good students, as those illustrated in Type 2 (stated subgoals), Type 3 (induced a goal from a sequence of actions), and Type 4 (relating the solution to the principle). If indeed there were such individual differences in the relation between successes at solving problems and the type of explanations given, then Gagné and Smith's result would be consistent with Stein and Bransford's and with ours. That is, giving explanations or elaborations does facilitate both problem solving and remembering in general (perhaps because it forces a person to process the material more extensively). However, individual differences in problem solving and remembering will remain as a function of the quality of the explanations and elaborations.

Students' Accuracy at Cognitive Monitoring

An important difference that was found between the good and the poor solvers was their ability to monitor their own comprehension and misunderstanding. Because a large portion of the students' elaborations (39% for the good and 42% for the poor) were monitoring statements, it seems that these statements must serve an important purpose. Even though both the

good and poor students had proportionately the same amount of monitoring statements, there was a significant difference between the good and poor students in the relative proportion of statements indicating comprehension failure versus understanding. Monitoring statements that indicate comprehension failure are questions raised, usually about an example line, such as, "Why is mgsinΘ negative?" An example of a monitoring statement that indicates understanding is "Okay" or "I can see now how they did it."

In general, the good students generated a large number of statements that reflected their failure to comprehend (9.3 such statements per example), whereas the poor students claimed that they did not understand with only 1.1 statements per example. Moreover, the poor students not only did not realize that they did not understand, in fact, they thought more often that they did understand. Thus, they generated an average of 6.2 statements per example indicating that they did understand (in contrast to the 1.1 statements indicating that they did not understand). Good students, by contrast, generated an equivalent number of statements indicating that they understood (10.8 statements) and that they misunderstood (9.3 statements). Basically, this suggests that the poor students do not accurately monitor their own comprehension. Not only do they *not* realize that they have misunderstood, they in fact think that they do understand.

We do not yet have a clear conception of the underlying causes of these different monitoring accuracies between the good and the poor students. If we assume that understanding is the instantiation of cognitive structures, then, we could speculate that, when a student has barely any structures at all to correspond to the example exercise, then they might falsely believe that they understand. However, for good students who have some incomplete structure of the example exercise as derived from having read the text, they can monitor their understanding more accurately by assessing the degree to which parts of their cognitive structures are instantiated, revised, and completed. In fact, we can support this interpretation somewhat by looking at the content of what these monitoring statements say. For the good students, their monitoring statements, when referring to misunderstandings, are specific, for example: "I'm wondering whether there would be acceleration due to gravity?" or "Why does the force have to change?" Such specific queries imply that the good student has a specific schema in mind and is trying to instantiate that schema by asking a specific question that pertains to a part of the schema. Poor students, on the other hand, may not have any kind of structures in memory at all. This would predict that they cannot ask specific questions. In fact, they do ask very general questions such as: "Well, what should you do here?"

The advantage of having an accurate monitoring of one's understanding is that the realization that one does not understand should elicit attempts to understand. That is exactly what we found in both our good and poor students. That is, in the majority of the cases (85% of the time for the good students and 60% of the time for the poor students), realizations of comprehension failures triggered episodes of self-explanations. Hence, because one is more likely to resolve one's misunderstanding by engaging in self-explanation, accurate monitoring of one's understanding is crucial for learning.

WHAT ONE LEARNS WHILE STUDYING EXAMPLES

As mentioned, good students generated a significantly greater amount of self-explanations, and their explanations were qualitatively better. We decided upon two ways to further examine the nature of self-explanations that were particularly suited to those generated by the good students. First, we found that the structure of explanations often took the form of specifying the conditions of applicability for the actions, stating the consequences of actions, as well as imposing subgoals on a sequence of actions. For instance, in response to an example line which stated that: "It is convenient to choose the x-axis of our reference frame to be along the incline and the y-axis to be normal to the incline" (Halliday & Resnick, 1981, p. 71), a good student explained the conditions of the choice by saying "and it is very, umm, wise to choose a reference frame that's parallel to the incline, parallel and normal to the incline, because that way, you'll only have to split up mg, the other forces are already, component vectors for you." Thus, we can generalize that, when students explain the example actions to themselves, they basically justify the actions by inferring the additional tacit information that is not explicated in the example statements.

What is the content and the source of this tacit knowledge evident in the self-explanations? We found that the good students try to understand the example by relating the example statements to explanations and principles stated in the text. The response to Line 3 of Fig. 8.3, "So that means that they have to cancel out, only the body wouldn't be at rest," for instance, was guided by Newton's first law, that if there is no motion, then the sum of the forces must be zero. Relating the example statement to the principles previously presented in the text served the purpose of providing the rationale for the procedural steps of the example solution. Moreover, we found that through such explanations, students increased their understanding of the fundamental physics principles introduced in the text (Newton's three laws), as well as the meaning and extension of the various concepts involved (such as the relation between weight and force). It seems then, that students are able to learn a great deal from studying even a few examples. To state in a nutshell what students learned from example-

exercises (these results are described in detail in Chi, Bassok, Lewis, Reimann, & Glaser, in press), we can say that in their explanations, by adding many linkages between the explicit statements made in the examples, as well as between the example statements and the principles in the text, they understood the rationale of the solution procedure better. That is, by specifying the conditions and consequences of various actions and inducing the goals for a set of actions, they probably constructed a more coherent representation of the entire solution procedure. The ultimate outcome of a more coherent representation was that the good students increased their understanding of the fundamental physics principles presented in the text. (This increased understanding was assessed by pre- and posttests of the physics principles, administered prior to and subsequent to the example-studying phase of the study.)

There seems to be a discrepancy between our findings, the fact that students can learn from examples and do seem to transfer their knowledge to end-of-chapter problems, and others in the literature which show that examples are of limited use and the acquired procedures seem to be rather syntactic in nature. However, there are three key differences between our study and others that seem to be responsible for this discrepancy. First, our students studied several pages of text, much as they would in a regular classroom, whereas the students in many of these other studies were not supplied with the background material (e.g., Reed, Dempster, & Ettinger, 1985). The text part of the chapters actually supplies knowledge of concepts and principles, which may lead to learning with understanding rather than to simply learning a syntactic procedure. The second, and perhaps the most important difference is that our students were encouraged to give explanations, and it is possible that those students who did generate self-explanations may have induced greater understanding. Finally, the successful learning we have described was characteristic only of our good students. Poor students generated very few and very impoverished explanations and were subsequently unsuccessful in their problem solving. Because we analyzed individual differences, at least differences between the good and the poor solvers, we were able to uncover differences that may have been masked in the other studies. What students learned depended on the kind of explanations they provided.

How One Uses What is Learned

So far, we have addressed the issues of how one learns from studying examples and what one learns. Once the learning is completed, presumably the knowledge gained is used in solving problems. We have done some preliminary analyses that characterize how examples are used in solving problems. For example, we found that in general, *all* students liked to rely on examples in their initial attempts to solve problems, as is consistent with the findings cited earlier (e.g., Pirolli & Anderson, 1985). For example, for the 12 isomorphic problems, the good students referred to the examples in 9 out of the 12 problems, and the poor students referred to the examples in 10 out of the 12 problems. Within each problem solving protocol, however, there were pronounced individual differences in the frequency with which students used examples. Poor students used examples more often than the good students. That is, within each problem that the students had to solve, the poor solver would refer to the example about 6.6 times per problem, whereas the good solver would refer to the example only 2.7 times.

The most interesting differences between the good and poor solvers lie in the characteristics of their example-using behavior. The poor solvers, when reviewing the examples, basically reread them. Therefore, we examined the purpose of each episode of reference to an example to see how the example was being used. There were three kinds of behavior: The students could be rereading the example; they could be mapping something from the example to the problem that they were working on, such as looking to see how the frame of reference was drawn; or they could be checking for the accuracy of an equation, by looking to see what the units of weight were, for instance. The poor students predominantly spent their example-reference episodes in rereading the example. Thus, 4.1 out of their 6.6 example-referencing episodes were spent rereading, whereas only 0.6 out of the good solvers' 2.7 episodes were spent rereading.

We surmise that the poor students were rereading the examples basically because they had not gotten much out of the examples when they studied them initially. We can substantiate this interpretation by looking at the number of lines in the example that were reread, as well as the location of the line where the student began rereading. The poor students, reread, on the average, 13.0 lines for each example-referencing episode, whereas the good students reread an average of 1.6 lines. Furthermore, all the poor students, upon their first encounters with problem solving (in the first set of isomorphic problems), started rereading the examples from the very first line, whereas none of the good students did. We have thus characterized the poor students' use of examples as searching for a solution or a template from which they could map the to-be-solved problem so that they could generate a solution.

Good students, on the other hand, use examples for other purposes. The best characterization is to say that the good students used examples truly as a reference. This can be substantiated by two additional pieces of evidence. First, when they used the examples, they read only an average of 1.6 lines for each episode of use, and the 1.6 lines tended to contain a particular equation or free-body diagram that they were checking on or

mapping from. This suggests that they were using the example to retrieve a particular piece of information. Thus, the good students must already have had a plan for a solution in mind. A second piece of evidence in support of this interpretation was that good students often prefaced their example-referencing episodes with the announcement of a specific goal, such as "I'm looking at the formula here, trying to see how you solve for Force_1, given the angle."

The kind of analyses that have been completed so far characterize, to some extent, the different ways good and poor students use examples. The kind of analyses that we need to pursue now are those that pinpoint exactly how the examples are used by each individual student. That is, in what analogical ways are examples used by the good and poor students? Do the poor students tend to notice only surface similarities, and the good students notice deeper similarities? An example of a surface similarity might be that both the example and the to-be-solved problem involve inclined planes. A deeper similarity might be noticing that both the example and the to-be-solved problem involve acceleration rather than static situations (or that the two situations are different, one involving a static situation and the other involving a dynamic case).

Likewise, we can also ask at what level is the mapping achieved? Mapping can be done at a superficial surface level. For example, a poor student might map and use an equation from the example simply because the equation contains all the variables that are stated in the to-be-solved problem. On the other hand, a good student might map a procedure from an example in a more global but deeper way. For example, a good student might know that the mapped procedure will achieve a particular result, such as decomposing the forces. Such analyses of how examples are used can also reveal the kind of understanding students have of the examples, so that the findings can validate the analyses of students' understanding as extrapolated from their explanations and their monitoring statements while studying.

INSTRUCTIONAL IMPLICATIONS

Constructing Instructional Materials: Limited Potentials

Our analyses of the limitations of example exercises, coupled with our research results, at first glance indicate that one constructive approach to improving learning might be to invest efforts in designing better example exercises. An obvious approach is to design exercises that fill in all the gaps or leaps in the current kinds of example exercises, that were pointed out earlier. For example, our results seem to indicate that example exercises should be designed to provide conditions for the actions, specify the consequences, supply goals for a sequence of actions, and relate the actions to the principles introduced in the text. Our results can further enhance the design of examples because they provide evidence about the locations of misunderstandings, as well as the kind of inferences that are needed in order to understand a particular sequence of actions in an example. In fact, we can utilize the explanations provided by the best students as a guide for constructing the justifications that are implicitly embedded between the statement lines.

Upon closer scrutiny, however, it becomes apparent that there are several reasons for the limited potential of such an approach to improving learning. First of all, it is not possible to supply all the rationales and justifications. Each example will always have some idiosyncrasies that might be perceived as inherent to the solution, and thus should not be elaborated upon, and each finite set of examples will be but a specific sample of the possible set of problems. But a more important limitation is that the justifications and explanations added have to be tailored to the individual student's understanding. For a good student who can generate his own explanations, these additional comments will be redundant. For the poor student who has little understanding, such explanations may actually confuse rather than clarify, and perhaps undercut rather than enhance performance, as Reder et al.'s (1986) results imply. And finally, as noted by VanLehn (personal communication), students can often choose to ignore the explanations provided, especially if these explanations are not tailored to their level of understanding.

Teaching Constructive Learning

Given that there are limits to the approach of improving learning by designing better—more explanatory—instructional materials, and given that students are continually encountering new domains to learn, a more sensible approach seems to be to focus on the learner, and teach the learner better learning strategies, so that he can use it in a variety of domains and learning situations. On the basis of our results, what kind of learning strategies might be good candidates? One possibility that comes immediately to mind is to teach students to monitor their comprehension more sensitively and accurately. This seems particularly important, at first sight, especially in light of our finding that for both the good and the poor students, when they do detect that they misunderstand, such misunderstanding triggers episodes of explanations. Thus, it seems important for the students to know *when* they do not understand, because such awareness potentially leads to further processing that may result in understanding. However, it is not clear that we can easily teach students to monitor their

own states of comprehension. Our data show clearly that only the good students can monitor their states of comprehension accurately. The poor students are very inaccurate at monitoring their comprehension states. In order to describe what causes good students to know that they understand and poor students not to know, we need to have a better account of the mechanism underlying understanding, which we as yet know little about. One possibility is that accurate monitoring of comprehension can occur only if the statements presented in the example exercises deviate only to a small degree from the students' existing schema (the schema to which they are trying to fit the information in the example).

Regardless of what the mechanism of monitoring is, one promising instructional approach might be simply to alert the poor students to locations that the designer of the example knows might be problematic. A simple probe (such as an asterisk) at critical locations might be sufficient to serve as a prompt for students to engage in explanations. Such a prompt might be useful in eliciting relevant explanations, although not necessarily. As we noted in our results, although the poor students made very few explanation-type remarks, for the few that they did make, their contents were qualitatively poorer as compared to the explanations of the good students. However, because the poor students generated only a few explanations, the finding regarding the quality of their explanations is still tenuous. Thus, whether or not poor students can be encouraged to generate good explanations remains an open empirical question.

Besides teaching students to monitor their own understanding, and/or prodding them to generate explanations, a more promising approach might be to shape students' explanations. Given that we have identified the characteristics of good students' explanations, we can use these characteristics to develop a set of prompts that can shape students' explanations. For example, at critical problematic locations, probes can take the form of asking such questions as: "Under what condition is this action to be taken?"; or "What else may be true if this is true?" or "What is the goal of this set of actions?"; or "How does this action relate to a principle introduced in the text, and which principle does it relate to?" These kinds of more specific probes might succeed in producing explanations that are qualitatively more like those generated by good students.

CONCLUSIONS

Our research on learning from examples has uncovered several interesting phenomena. Most prominently, it has found that students who are successful at solving problems are those who learned the materials in a different way then those students who were less successful at solving problems.

More specifically, the good students' learning from examples was characterized by the generation of self-explanations. Their self-explanations had the characteristic of adding tacit knowledge about the actions of the example-solution. Moreover, these self-explanations induced greater understanding of the principles introduced in the text. It was also found that the production of these self-explanations seemed to be guided by the accuracy with which students monitored their states of comprehension. Their knowing that they misunderstood generally elicited explanations. Furthermore, the good students' self-monitoring yielded specific inquiries through which they could search for an answer, whereas the poor students' self-inquiries were general undirected questions.

There are important implications to be derived from this research. First and foremost, in terms of theories of learning from examples, our data suggest that students can learn, *with understanding,* from a single or a few examples, contrary to the other available empirical evidence. (That is, students can learn more than just syntactic rules and can transfer what they have learned to dissimilar transfer-type problems.) However, only those students who provide adequate explanations during studying are able to see the degree to which they can generalize their problem-solving skills. Hence, not only does our study resolve some of the discrepant findings in the empirical literature, but it also points to limitations in the assumptions of the existing theories of learning from examples. For instance, current theories hold that explanations serve the purpose of justifying an example as an instance of a principle, assuming that the student would have complete knowledge about the principle. Our results suggest, by contrast, that explanations can serve the additional important function of enhancing and completing students' understanding of the principles.

Our research raises more questions than it answers. Some critical ones that we need to address are: What does it mean to understand an example while studying it? How does understanding relate to cognitive monitoring? How does understanding relate to the way that examples are used? Deeper analyses of our results hinge on our potential explication of what understanding entails, how it can be represented, and how it should be assessed.

ACKNOWLEDGMENTS

We are grateful for the continued support of this research from the Office of Naval Research, under Contract No. N00014-84-K-0542. We appreciate Bob Glaser's, Lauren Resnick's, and Kurt VanLehn's comments on this chapter, as well as our lengthy (but unresolved) discussions with VanLehn about the issue of understanding.

REFERENCES

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192–210.

Anderson, J. R., Greeno, J. G., Kline, P. J., & Neves, D. M. (1981). Acquisition of problem-solving skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 191–230). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bruner, J. S. (1966). *Toward a theory of instruction.* Cambridge, MA: Harvard University Press.

Chi, M. T. H., Bassok, M., Lewis, R., Reimann, P., & Glaser, R. (in press). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science.*

Elio, R., & Anderson, J. R. (1983). Effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 113,* 541–555.

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

Eylon, B., & Helfman, J. (1982, February). *Analogical and deductive problem-solving in physics.* Paper presented at the AERA meeting, New York.

Gagné, R. M., & Smith, E. C., Jr. (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology, 63,* 12–18.

Glaser, R. (1967). Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), *Learning and individual differences* (pp. 1–18). Columbus, OH: Merrill.

Halliday, D., & Resnick, R. (1981). *Fundamentals of physics.* New York: Wiley.

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

Klinger, E. (1974). Utterances to evaluate and control attention distinguish operant from respondent thought while thinking out loud. *Bulletin of the Psychonomic Society, 4,* 44–45.

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

Lewis, C. (1988). *Why and how to learn why: Analysis-based generalization of procedures. Cognitive Science, 2,* 211–256.

Mitchell, T. M., Keller, R. M., & Kedar-Cabelli, S. T. (1986). Explanation-based generalization: A unifying view. *Machine Learning, 1,* 48–80.

Neves, D. M. (1981). *Learning procedures from examples.* Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, PA.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology, 32,* 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., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory & Cognition, 11,* 106–125.

Simon, H. A. (1980). Problem solving and education. In D. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research* (pp. 81–86). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stein, B. S., & Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18,* 769–777.

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.
Thomas, J. C. (1974). An analysis of behavior in the Hobbits-Ores problem. *Cognitive Psychology, 6*, 257–269.
VanLehn, K. (1986). Arithmetic procedures are induced from examples. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 133–179). Hillsdale, NJ: Lawrence Erlbaum Associates.
VanLehn, K. (in press). *Mind bogs: The origins of procedural misconception.* Cambridge, MA: MIT Press.
Wickelgren, W. A. (1974). *How to solve problems: Elements of a theory of problems and problem solving.* San Francisco: W. H. Freeman.
Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*, 137–166.

[illegible] ransfers?

Jill H. Larkin
Carnegie-Mellon University

Everyone believes in transfer. We believe that through experience in learning we get better at learning. The second language you learn (human or computer) is supposed to be easier than the first. Professors' lore has it that seniors taking Psychology 101 as a distribution requirement perform "better" than freshmen taking it as an introduction to their major. All these common beliefs reflect the sensible idea that, when one has acquired knowledge in one setting, it should save time and perhaps increase effectiveness for future learning in related settings.

But these are popular concepts of transfer. What is an operational definition of this phenomenon? Transfer does not mean merely applying old knowledge in new situations. That happens often. We frequently use our knowledge of maps to figure out the layout of unknown cities. Transfer means applying old knowledge in a setting sufficiently novel that it also requires learning new knowledge. If there were no transfer, then solving problems in a new domain would require totally mastering a set of neces-