Explicit Heuristic Training as a Variable in Problem-Solving Performance

Article *in* Journal for Research in Mathematics Education · May 1979

DOI: 10.2307/748805

Alan H Schoenfeld
University of California, Berkeley

EXPLICIT HEURISTIC TRAINING AS A VARIABLE IN PROBLEM-SOLVING PERFORMANCE

ALAN H. SCHOENFELD
Hamilton College

Overview and Discussion of the Relevant Literature

This paper describes, in some detail, the problem-solving processes of seven upper-division college students working a series of problems that could be solved by the application of one or more heuristics, or general problem-solving strategies. The primary purposes of the experiment were—

(a) to study the impact of instruction in five heuristics on some students' performance on a series of problems comparable to, but not isomorphic to, the instructional problems;
(b) to see if other students working the same problems (for the same total amount of time and seeing the same solutions), but not receiving the heuristics instruction, would use or intuit the strategies as a result of their problem-solving experience; and
(c) to see, by comparing the two groups, if explicit instruction in heuristics "makes a difference"—both as measured by pretest-to-posttest gains and as indicated by an examination of problem-solving procedures.

E. G. Begle (Note 1) wrote that "problem solving, in my opinion the most important outcome of mathematics education, at the upperclass undergraduate level, to the best of my knowledge, has not been seriously studied before." Thus the body of literature relevant to this study, and dealing with the process rather than the product of problem solving, discusses problem-solving tasks that are easier and call for less mathematical background than the tasks studied here.

The groundwork for explorations in heuristics, of course, was established by Polya in *How to Solve It* (1945) and given a much more extended treatment in his *Mathematical Discovery* (1962, 1965). College-level textbooks on problem solving (Rubenstein, 1975; Wickelgren, 1974), like Polya, described a variety of useful strategies but did not describe research in problem solving. Lipson (1972) studied senior mathematics majors enrolled in a course on heuristics, but the focus of the research was on carry-over effects of heuristics instruction on those students' students (her subjects were student teachers); the work does not bear directly on the material discussed here.

Lucas (1972) studied students' use of heuristics in calculus classes. The results, although suggestive (decreased study time, more checking for heuristics groups), were difficult to interpret clearly—a problem, unfortunately, with much "real world" instruction. There are sufficiently many uncontrolled (uncontrollable?) variables in such experimentation that, for preliminary comparison of some instructional variables, the use of an artificial instructional environment may be appropriate. Thus this study strove to create a replicable and relatively controlled instructional setting, the better to study the impact of specific problem-solving variables. Goldberg's study (1975) may be the most relevant to the work described here. The students, nonmathematics majors taking a course in number theory, provided some evidence that explicit instruction in heuristics did enhance their ability to construct proofs. This study differs in a number of fundamental ways, perhaps the most salient differences being the degree of experimental control and the fact that the problems used here are more difficult than those she used and call for a much larger mathematical knowledge base on the part of the students (see the next section for a more detailed discussion). Smith's study (1973) focused on transfer and provided some evidence that transfer of general heuristics may be less than one would ideally hope for. These results, less optimistic in a sense than those in Wilson (1968) but compatible from different perspectives with arguments advanced by Schoenfeld (1979b) and Resnick (1979), suggest that, at least in short-term, small-scale experiments, we may want some substantial portion of test items to resemble (though not simply copy) instructional problems.

Other studies of mathematical problem solving that are relevant for their focus either on problem-solving processes or on heuristics generally deal with precollege students. Kilpatrick (1978) gave a useful classification of problem-solving variables; Kantowski (1977) gave careful scrutiny to the problem-solving processes of eight ninth graders working on nonroutine geometry problems. Lucas (Note 2) took a detailed look at problem and solution space structure, something we will need to know more about to better explicate process in detail. Landa (1976) discussed explicit training in heuristics for geometry problems.

The other body of relevant literature comes from artificial intelligence and cognitive science. Newell and Simon's *Human Problem Solving* (1972) pioneered the techniques of close observation of the problem-solving process via protocol analysis and demonstrated the utility of that approach. In a survey article, Larkin (1979) speaks in general of the relation between psychological investigations of problem-solving mechanisms in the laboratory and the potential applications of these results to the classroom.

The Experiment: Rationale and Design

The case has already been well made (Kilpatrick, Note 3; Newell & Simon, 1972) for conducting studies examining, in detail, the processes of students as they solve problems. These studies, as opposed to the large-scale statistical tests used for comparing instructional methods, are designed to clarify and elucidate the *mechanisms* used by problem solvers in approaching and working on problems.

The subjects were seven upper-division science and mathematics majors recruited as volunteers from upper-division courses in mathematics at the University of California, Berkeley. The students were told that they were to "work a bunch of specially chosen problems designed to improve your problem solving ability." Of the seven, four were randomly chosen for the experimental group. Two of these four were mathematics majors, as were two of the three control group students; their mathematical backgrounds were comparable.

In order to obtain maximum information about the problem-solving processes of the students, each was treated individually. The students were trained to talk out loud as they solved problems. Then each took a pretest, consisting of the five problems given in Figure 1. They worked on each problem for 20 minutes, or until they were confident that they had solved it. This process was repeated in the posttest (Figure 2). Our analysis of test results will be two-fold: first, a comparison of test scores on the basis of "completely solved" and "almost completely solved" problems; second, a detailed look at the solution processes.

Pretest

1. Let a and b be given real numbers. Suppose that for all *positive* values of c the roots of the equation

$$ax^2 + bx + c = 0$$

are *both* real positive numbers. Present an argument to show that a must equal zero.

2. Ten people are seated around a table. The average income of these ten people is $10,000. Each person's income is the average of the incomes of the people sitting immediately to his left and right. What is the possible range of incomes for each person? (Incomes are given in whole dollar amounts.)
3. Let n be a given whole number. Prove that if the number $(2^n - 1)$ is a prime, then n is also a prime number.
4. You are given the real numbers a, b, c, and d, each of which lies between 0 and 1. Prove the inequality

$$(1-a)(1-b)(1-c)(1-d) \geq 1-a-b-c-d.$$

5. What is the sum of the series

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n\cdot(n+1)}?$$

Prove your answer if you can.

Figure 1.

Between the pretest and the posttest, each of the students sat through five instruction sessions spread over a 2-week period, during which they worked on, and saw the solutions to, 20 problems (including the 5 pretest problems). These problems, and the 5 posttest problems, were chosen from five classes of similar problems—similar in the sense that they could be

Posttest

1. Suppose p, q, r, and s are positive real numbers. Prove the inequality

$$\frac{(p^2+1)(q^2+1)(r^2+1)(s^2+1)}{pqrs} \geq 16.$$

2. For what values of "a" does the system of equations

$$\left\{\begin{array}{c} x^2 - y^2 = 0 \\ (x-a)^2 + y^2 = 1 \end{array}\right\} \text{ have }$$

a) no solutions?
b) 1 solution?
c) 2 solutions?
d) 3 solutions?
e) 4 solutions?

3. Let S be a set which contains n elements. How many different subsets of S are there, including the null set?

4. Prove that the product of any three consecutive whole numbers is divisible by six.

5. Let A and B be two given whole numbers. The *Greatest Common Divisor of A and B* is defined to be the *largest* whole number C which is a factor of both A and B. For example, the G.C.D. of 12 and 39 is 3, and the G.C.D. of 30 and 42 is 6. PROVE that the greatest common divisor of A and B is unique.

Figure 2.

solved by a particular problem-solving approach. There were, for example, 5 problems amenable to an approach by mathematical induction (Figure 3).

One of the questions to be asked in this experiment was the following: Having worked the 20 problems in practice sessions and having seen the 4 problems that were solved by induction, would the control students think to use induction on the posttest problem that called for it? Or more generally, is *explicit* mention of a strategy important, or will students intuit it from their experience? This question applies to each of the problem-solving approaches treated in the experiment, of course. Let us now consider the experimental design in detail.

The bulk of the instruction was carried out through written materials and tape recorded "lectures." In each instruction session the students were given a booklet with four practice problems and a tape recorder. They worked on each problem for up to 15 minutes, or until it was solved. When finished they turned to the next page in the booklet, which presented a solution to the problem. They also turned on the tape recorder to listen to a solution parallel to, but not identical with, the written solution. (This allowed them to hear the solution, which often helps in comprehension.) If, after thinking about a problem solution for 10 minutes, a student had a question, he or she could ask the experimenter. When a question was asked that could be answered by the instructional materials, the student was referred to them. Other questions were generally limited to technical matters such as "the converse is not always true but the contrapositive is, right?" Colleagues of the experimenter listened to the instructional tapes for both the experimental and the control groups and claimed there was no discernible difference in either enthusiasm or clarity of presentation, so that any difference in scores between the two groups should not be attributable to experimenter bias.

In sum then, the students in both groups spent the same amount of time working the same practice problems and saw the same solutions. The differences in treatment between the two groups were as follows:

1. The four heuristics students were told (on tape) at the beginning of their first practice sessions that the experiment would try to show how five specific strategies would help them to solve problems. They were then given the strategies (see Figure 4). The list was placed conspicuously in front of them during all practice sessions and during the posttest. At the first session

Problems Amenable to an Inductive Approach

1. (pretest) What is the sum of the series

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{n\cdot(n+1)}?$$

   Prove your answer if you can.

2. You are given n points in the plane, none of which lie on a straight line. How many straight lines can you draw, if each straight line must pass through two of the n points?

3. Let x be any odd integer. Show that x^2 leaves a remainder of 1 when divided by 8.

4. Determine a formula for the product

$$(1-\tfrac{1}{4})(1-\tfrac{1}{9})(1-\tfrac{1}{16})\cdots(1-\tfrac{1}{n^2}).$$

   Prove it if you can.

5. (posttest) Let S be a set which contains n elements. How many different subsets of S are there, including the null set?

Figure 3.

The Five Problem-Solving Strategies

1. Draw a diagram if at all possible.

   Even if you finally solve the problem by algebraic or other means, a diagram can help give you a "feel" for the problem. It may suggest ideas or plausible answers. You may even solve a problem graphically.

2. If there is an integer parameter, look for an inductive argument.

   Is there an "n" or other parameter in the problem which takes on integer values? If you need to find a formula for $f(n)$, you might try one of these:

   A) Calculate $f(1)$, $f(2)$, $f(3)$, $f(4)$, $f(5)$; list them in order, and see if there's a pattern. If there is, you might verify it by induction.

   B) See what happens as you pass from n objects to $n + 1$. If you can tell how to pass from $f(n)$ to $f(n + 1)$, you may build up $f(n)$ inductively.

3. Consider arguing by contradiction or contrapositive.

   Contrapositive: Instead of proving the statement "If X is true then Y is true," you can prove the equivalent statement "If Y is false then X must be false."

   Contradiction: Assume, for the sake of argument, that the statement you would like to prove is false. Using this assumption, go on to prove either that one of the given conditions in the problem is false, that something you know to be true is false, or that what you wish to prove is true. If you can do any of these, you have proved what you want.

   Both of these techniques are especially useful when you find it difficult to begin a direct argument because you have little to work with. If negating a statement gives you something solid to manipulate, this may be the technique to use.

4. Consider a similar problem with fewer variables.

   If the problem has a large number of variables and is too confusing to deal with comfortably, construct and solve a similar problem with fewer variables. You may then be able to

   A) Adapt the method of solution to the more complex problem.

   B) Take the result of the simpler problem and build up from there.

5. Try to establish subgoals.

   Can you obtain part of the answer, and perhaps go on from there? Can you decompose the problem so that a number of easier results can be combined to give the total result you want?

Figure 4.

they listened to a 10-minute tape describing the strategies. The control students were simply welcomed to the experiment (on tape) and told how to use the materials.

2. Although the solutions to each problem seen by both groups were identical, the students in the heuristics group saw in addition an overlay to each solution that indicated where the strategy had been used. The differences between the two presentations can be seen in Figure 5. There the complete solution is given for one of the induction problems. The entirety of Figure 5 was seen by the experimental group. Only the right-hand side of the page, consisting of the solution, was seen by the nonheuristics group.

The tape recordings heard by each group recapitulated the solutions as seen on the page. The nonheuristics group listened to a tape that said "Let's calculate a few of the sums and see what happens. The first term is 1/2; . . ." The heuristics group heard "Notice that there is an n in the problem statement. When we see an integer parameter, we should calculate the values and see what happens. . . ." Recall that colleagues had listened to all of the tapes and determined that there was no difference in clarity, level of exposition, or enthusiasm of presentation between the tapes that the two groups of students heard.

3. The order of the problems was different. In each session the heuristics group practiced one particular strategy. For the nonheuristics group the order of the problems was scrambled, to see if the control students would intuit the strategies from a more random ordering of the problems.

4. Finally, there was one slight difference in the posttest. At 5-minute intervals the control students were told "You've been working for five minutes now. You may be on the right track; you may not. But stop, take a deep breath, look over your work. Then decide whether you want to continue in that direction." The strategy group heard "and look over the list of strategies" after "look over your work." The reasons for this are discussed in the next section.

Results

In the first section we indicated the three main questions to be analyzed. We shall begin with the third, a comparison of the performances of the two groups of students. The evidence marshaled here will serve as the foundation for our analysis of the first two questions. The bulk of our discussion will be, as suggested earlier, a treatment of the students' protocols, or transcripts of their thinking aloud. Surprisingly, however, this experiment with only seven people does provide the fuel for a statistical analysis. We begin with a discussion of two gross measures of problem-solving improvement. Since the problem-solving abilities of the students can vary substantially at the onset of an experiment, the appropriate measure of the instructional materials' effect is the difference in scores from pretest to

What is the sum of the series

$$S = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{(n)(n+1)} \ ?$$

Prove your answer if you can.

Notice the integer parameter . . .

Solution

Let's examine a few of the sums.

Calculate the values for $n = 1, 2, 3, 4, \ldots$ 1/2, 2/3, 3/4, 4/5,

The first term is 1/2.

The sum of the first two terms is $1/2 + 1/6 = 2/3$.

The sum of the first three terms is $2/3 + 1/12 = 3/4$.

The sum of the first four terms is $3/4 + 1/20 = 4/5$.

Guess the formula from the pattern →

We can be pretty certain at this point that the sum of the first n terms is $(n/n + 1)$. As usual, we verify a formula like this by induction.

Verify the formula by induction →

For $n = 1$, the sum of the first n terms is 1/2, which checks.

Suppose the sum of the first k terms is $(k/k + 1)$; that is,

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{(k)(k+1)} = \frac{k}{k+1}.$$

Then the sum of the first $(k + 1)$ terms is

$$\begin{aligned}
&\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots \\
&\qquad + \frac{1}{(k)(k+1)} + \frac{1}{(k+1)(k+2)} \\
&= \frac{k}{k+1} + \frac{1}{(k+1)(k+2)} \\
&= \frac{k^2 + 2k + 1}{(k+1)(k+2)} = \frac{(k+1)^2}{(k+1)(k+2)} \\
&= \frac{k+1}{k+2},
\end{aligned}$$

which is the desired formula for $n = k + 1$.

This completes the argument.

Figure 5. A sample solution, with heuristic overlay.

Table 1
Number of Completely Solved and Almost Completely Solved Problems

		Pretest	Posttest
Nonheuristic	S_1	2 (1)[a]	2 (2)
	S_2	1 (1)	2 (1)
	S_3	2 (0)	1 (0)
Heuristic	S_4	2 (1)	5 (0)
	S_5	0 (0)	2 (0)
	S_6	2 (0)	4 (1)
	S_7	0 (0)	2 (0)

[a] Number of almost completely solved problems are indicated in the parentheses.

posttest. Table 1 provides the number of test problems completely solved by each of the students.

At a qualitative level, we note that *all* the students in the heuristics group improved from pretest to posttest, whereas only one in the nonheuristics group did. In fact, the average net gain for the nonheuristics group is 0, and for the heuristics group it is more than 2. More importantly, all four of the students who received heuristics training outscored all three who did not. The probability of this happening randomly is 1/35, so the differences in the students' performance are statistically significant at $p < .05$.

A second gross measure of success can be seen by looking at the "almost completely solved" problems. The problem-solving process was cut off at 20 minutes on each problem; we can ask if it is *clear* that with another 5 to 10 minutes, a solution would have been obtained. Table 1 also provides the data for this measure.

Were the differences obtained really due to the heuristics? Certainly the experimental design was such that they should be. The statistics are silent on that question, however. We turn to the protocols of the solutions themselves. We will examine how all seven students worked on posttest Problem 1.

"Suppose p, q, r, and s are positive real numbers. Prove the inequality

$$\frac{(p^2 + 1)(q^2 + 1)(r^2 + 1)(s^2 + 1)}{pqrs} \geq 16."$$

The problem is most easily dispatched by employing the fourth strategy on the list in Figure 4, "consider a similar problem with fewer variables." Noting that the left-hand side of the inequality is the product of four terms of the form $(x^2 + 1)/x$, one need only prove the one-variable inequality $(x^2 + 1)/x \geq 2$, substitute p, q, r, and s for x, and multiply the four resulting inequalities.

None of the three students in the nonheuristics group solved the problem, although one should have and got an "almost." S_2 and S_3 both jumped into messy algebraic computations and persisted, through the third 5-minute warning, in trying to show by some sort of clever factorization that

$$(p^2 + 1)(q^2 + 1)(r^2 + 1)(s^2 + 1) - 16pqrs \geq 0.$$

With less than 5 minutes left, S_2 decided to set $p = q$ and $r = s$, but was unable to do anything with that; S_3 was interested to see that substituting the value 1 for p, q, r, and s led to the minimum value of 16, and ran out of time. Neither of them was anywhere near a solution.

S_1, on the other hand, tried to solve the problem exactly as described in the solution above. Within a few seconds of reading the statement of the problem, she said "OK, I've got to show that $(p^2 + 1)/p$ is always greater than or equal to 2." Then for some reason she got sidetracked: She noticed that it equals 2 when p is 1, and, after noting that the value of $(p^2 + 1)/p$ increased for the next few integer values of p, tried to prove the result by induction! At the third 5-minute warning she gave this up and made a sketch of the graph of $p + 1/p$, which was accurate, but not analytically justified. She finished this at 20 minutes and said yes when asked if she were satisfied with the solution. When asked "What if we weren't satisfied?" she said, "I guess I could always find the minimum of $(p + 1/p)$ by calculus." That got her an "almost." At that point she was complimented on the design of her solution. Her response was "I noticed that whenever a practice problem had lots of variables in it, you tried to do the one- or two-variable problem. So I tried to do that here." More about this in a short while.

The performance of the heuristics group was substantially different. S_4 read the problem statement, said "that's a fewer variable problem," and tried to analyze the two-variable problem $(p^2 + 1)(q^2 + 1)/pq \geq 4$. After about 8 minutes she said "What about the one-variable problem?" and solved it in another 4 minutes.

Unlike S_4, S_5 relied consciously on the *list* of strategies at the very beginning. After reading the problem statement, he turned to the list and checked them off one by one: "Let's see, I don't think I can really draw a diagram for this one, and there isn't an integer parameter so I can't do an inductions, and let's see, what would the contrapositive be? No, that doesn't make sense; OK, how about fewer variables? Yeah. Let me see if the one-variable problem makes any sense. . . ." He solved the problem by the first warning.

S_6, much like S_2 and S_3, jumped right into a complicated algebraic morass. She was happily calculating away at the 5-minute warning. Only then did she stop, look at the strategies, choose the appropriate one, and after a few minutes on the two-variable problem, solve it correctly.

S_7 first tried to disprove the problem statement by showing that the product must be less than 16, and took about 8 minutes to realize that his logic was flawed. After that he wrote out the left-hand side of the inequality as the product of four similar factors, and decided to explore the nature of $(p^2 + 1)/p$ by plugging in a few values of p, and then seeing what happened as p grew infinitely large. He was nowhere near a solution when he ran out of time.

Thus a detailed look at the students' solutions to the test problems supports the results suggested by the statistics: Conscious application of the problem-solving strategy does make a difference. Only one of the three control students solved the problem, and she did so by intuiting and applying the strategy the others had been taught; yet three of four in the experimental group solved it. The picture is still complicated, however; we have more to look at.

First, we return to question (a) in the first section. What was the impact of the instruction on students' posttest performance? There are two components to this question. They are as follows:

(1) Can the students transfer their training from practice problems to posttest problems?

(2) Will they do so in unconstrained circumstances?

What we have seen indicates that some of the strategies are relatively easy for the students to learn and apply and that clear evidence of success, relative to question (1), is obtainable in the short term. The protocols we have just examined show, for example, that three of the four experimental students learned to use the fewer-variables strategy correctly. None of them had used this strategy on pretest Problem 4. For this strategy, then, the effect is large. We have similar results for the induction problems (Figure 3) where all four of the experimental group but none of the controls solved the posttest problem. The cue of the integer parameter n allowed students who recognized it to make progress; those who did not see it made none. The results were similar for the diagram problems (#1 on the pretest, #2 on the posttest); those who made use of the strategy had an easier time with the problem.

For the other two strategies, however, the results were inconclusive. Posttest Problem 5 disturbed some of the students, who either saw nothing to prove or got jumbled in the mechanics of a proof by contradiction; there was comparable performance for both groups. More instruction, or perhaps a different posttest problem, might have given different results.

With hindsight, we see clearly that the instruction for subgoals, tested on pretest Problem 2 and posttest Problem 4, was inadequate. The strategy itself is very complex; subgoals is, in fact, merely a convenient label for a whole class of related but different skills (Schoenfeld, Note 4). The attempt to look for transfer of this skill on the basis of five practice problems was at best naive. We see then, that there is a good deal more work to be done in selecting and teaching the strategies, even to obtain results in an experimental setting.

Let us now turn to (2) above. Unfortunately, the fact that a student knows how to use a strategy is no guarantee whatsoever that the student will indeed use it. Recall, for example, the performance of S_6 on posttest Problem 1. Even though she was taking part in an experiment designed to teach the five strategies, had the list of strategies in front of her at all times, and was fully competent (as her protocols show) at using the fewer-variables strategy, she ignored it completely when given posttest Problem 1. Only at the 5-minute warning did she stop to consider that it might be worthwhile to use the strategy.

This kind of behavior occurred more than once during the posttest. S_4, for example, had immediately begun solving the two simultaneous equations in posttest Problem 2 and was deeply involved in an incorrect solution when asked to take a deep breath and reconsider. That was enough; with an "Oh, sure: draw a diagram" she went on to solve the problem easily.

The implications of this kind of behavior are serious, for they point to major difficulties in taking work like this experiment from the laboratory to the classroom. The students in this experiment were (relatively) decent problem solvers with ingrained behavior patterns that had served them (relatively) well for many years; it would be naive to expect practice sessions on 20 problems to effect substantial changes in their behavior. For a more extended discussion of this, see Schoenfeld (1979a).

Finally, we turn to question (b) of Section 1. Having seen four problem solutions by induction, four solutions aided by the use of diagrams, four solutions by contradiction, four solved by fewer variables, and four by subgoals, would the control students intuit and use the strategies on problems similar to those that they had studied? As a rule, no. There were exceptions, such as S_1's performance on the fewer-variables problems. On pretest Problem 4, which asked students to show that for $0 \leq a, b, c, d \leq 1$,

$$(1 - a)(1 - b)(1 - c)(1 - d) \geq 1 - a - b - c - d,$$

she had multiplied out the terms on the left-hand side and spent 20 minutes in algebraic manipulations—showing no signs of awareness of the fewer-variables heuristic. As her comment after the posttest indicated, she had correctly intuited the strategy and learned to apply it, from the practice problems.

In contrast to S_1, however, S_2 and S_3 failed to make the connection—and in all honesty, it does not seem a terribly hard one to make, given the nature of the experimental environment. Even more dramatic, we have the students' performance on the induction problems. In many ways the posttest problem is very much like Problem 2 (see Figure 3) and one might well expect some transfer. There was none; not one of the control students tried to discover a pattern by making calculations for small integer values of n. This lack of transfer was not due to lack of memory; after the experiment was over these same students quoted some of the problems nearly verbatim when we discussed what the purpose of the experiment had been.

In sum, then, the results of this experiment point to the following:

1. Even in the enriched environment of the experiment, the degree to which students intuited the heuristics appropriate for the posttest was minimal. Real-life mathematical problem-solving experiences are not nearly as well ordered as they were in this experiment; the likelihood of students' picking up the strategies from their experience is small. (Recall that the

students were upper-division science and mathematics majors; they showed little evidence of using heuristic approaches to the problems on the pretest.) Thus, if we expect students to learn to use such strategies, we should label them explicitly and explicate their use in much the same way we would teach any other mathematical strategies or techniques. We would not rely on implicit instruction for the quadratic formula; can we expect it to work for subgoals?

2. That students *can* master a particular problem-solving technique is no guarantee that they *will* use it. Were it not for the 5-minute warning in the experimental format, a number of the posttest problems solved by the heuristics group would have gone unsolved. Even after a more extended training period, students will not instinctively reach for the strategies that experts may find natural; they must be taught not only how to use the strategies, but when. Little is known about this particular aspect of problem solving, either on the part of experts or on the part of students; much research remains to be done.

3. When problem-solving strategies are identified and taught, and when students think to use them, the impact on the students' problem-solving performance is substantial.

Where Do We Go from Here?

There are two major ways in which the experiment described in this paper can be varied to yield additional information about problem solving. They are as follows:

1. to keep the experimental format largely intact, while varying the heuristics being studied.
2. to "loosen" the format, to approximate real-world conditions.

Let us discuss (1) first. The five heuristics given in Figure 4 are not necessarily the most important, or the most difficult to teach of the strategies that might have been chosen for this experiment. They range from very straightforward (induction) to rather nebulous (subgoals). As we saw in Section 3, the students' posttest performance reflects this. Much more training in subgoals is necessary before instruction in the strategy will have demonstrable impact on students. The experimental environment, which allows for a detailed examination of the students' problem-solving processes before and after instruction in the strategies, is a useful vehicle for determining how much instruction, even under ideal circumstances, students will need before learning to use this or other strategies. It also provides detailed information about what they learn, and what they do not. A revised version of this experiment might contain seven or eight subgoals problems, or a subclass of such problems, or more detailed instructions on how to use the strategy.

In its present form the experiment can be used to study any of the large number of heuristic problem-solving approaches that have been identified in the literature. By systematically varying the strategies to be studied and the amount of training provided in them, we can begin to accumulate a large body of information about them. That information has been lacking to date and would prove extremely useful.

After we have learned more about the use of the heuristics through this kind of experiment, we can still exploit it to examine (in microcosm) various selection mechanisms. Can we find various cues that strongly suggest that particular strategies are appropriate for particular (as yet undefined) types of problems? The integer parameter is one such cue, easily identified; perhaps there are others. We may wish to include such cues in our instruction. The protocols of the posttests will provide clear evidence of whether the students have used them effectively or not.

In this version of the experiment the teaching was done via tape recordings, so that the instruction process could be checked for teacher bias. As much as possible, the goal was to avoid the chance that the difference in performance between experimental and control groups could be attributable to enthusiasm or bias on the part of the experimenter in favor of the heuristics group. Once the point is granted that instruction in the heuristic strategies does have an impact, instruction via tape recordings will no longer be necessary and the experimental format can be changed; the instruction might be given by a teacher to small groups of students, say half a dozen at a time. It might be rerun until, under these still ideal conditions, the students reliably learn the strategies. The instruction might then be moved to a regular classroom setting. With large numbers of students (comparatively), more reliable statistical validation of the experiment than that offered here should be attainable. Statistical analysis should not be the only research method, however. Since the circumstances would be different, it would again be valuable to record (some) pretests and posttests. This would provide information about the changes in instruction that might be necessary to adapt to the new environment. When this experiment is first brought to the classroom, it would probably be appropriate to continue the use of the 5-minute warning during testing; the case has yet to be made that heuristics can make a difference in classroom instruction. Later on one could see how the use of cues or other training in selection mechanisms might make the warning unnecessary.

We are still a long way from success. If we can succeed in identifying truly useful problem-solving strategies (and here we have done well); if we can understand and exploit the mechanisms by which experts call these strategies into play; and if we can create efficient means by which this knowledge can be passed on to our students, the potential rewards are great indeed. The obstacles are many, the field vast. This experiment shows that, under appropriate circumstances, explicit instruction in general problem-solving strategies (heuristics) does have an impact on students' problem-solving performance. The rest will, with luck and work, come with time.

REFERENCE NOTES

1. Begle, E. G. Personal communication, March 1977.
2. Lucas, J. *Induction problems and solution state-spaces.* Paper presented at the annual meetings of the National Council of Teachers of Mathematics, San Diego, April 1978.
3. Kilpatrick, J. *Research in the teaching and learning of mathematics.* Paper presented at the meeting of the Mathematical Association of America, Dallas, January 1973.
4. Schoenfeld, A. H. *About heuristics: What they are and how to teach them.* Manuscript to appear in the 1980 *Yearbook* of the National Council of Teachers of Mathematics.

REFERENCES

Goldberg, D. J. The effects of training in heuristic methods on the ability to write proofs in number theory (Doctoral dissertation, Columbia University, 1974). *Dissertation Abstracts International,* 1975, *35,* 4989B. (University Microfilms No. 75-7836)

Kantowski, M. G. Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education,* 1977, *8*, 163–180.

Kilpatrick, J. Variables and methodologies in research on problem solving. In L. L. Hatfield (Ed.), *Mathematical problem solving.* Columbus, Ohio: ERIC Clearinghouse for Science, Mathematics, and Environmental Education, 1978.

Landa, L. M. The ability to think: Can it be taught? *Soviet Education,* 1976, *18.*

Larkin, J. H. Teaching problem solving in physics: The psychological laboratory and the practical classroom. In F. Reif & D. Tuma (Eds.), *Proceedings of the conference on problem solving and education.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Lipson, S. H. The effects of teaching heuristics to student teachers in mathematics (Doctoral dissertation, Columbia University, 1972). *Dissertation Abstracts International,* 1972, *33,* 2221A. (University Microfilms No. 72-30,334)

Lucas, J. F. An exploratory study on the diagnostic teaching of heuristic problem solving strategies in calculus (Doctoral dissertation, University of Wisconsin, 1972). *Dissertation Abstracts International,* 1972, *33*, 6825A. (University Microfilms No. 72-15,368)

Newell, A., & Simon, H. *Human problem solving.* New Jersey: Prentice-Hall, 1972.

Polya, G. *How to solve it.* Princeton, N.J.: Princeton University Press, 1945.

Polya, G. *Mathematical discovery,* (Vols. 1 & 2). New York: Wiley, 1962 & 1965.

Resnick, L. Remarks during panel discussions. In F. Reif & D. Tuma (Eds.), *Proceedings of the conference on problem solving and education.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Rubenstein, M. *Patterns of problem solving.* Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Schoenfeld, A. H. Can heuristics be taught? In J. C. Lochhead (Ed.), *Cognitive process instruction.* Philadelphia: Franklin Institute Press, 1979a.

Schoenfeld, A. H. Remarks during panel discussions. In F. Reif & D. Tuma (Eds.), *Proceedings of the conference on problem solving and education.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979b.

Smith, J. P. The effects of general versus specific heuristics in mathematical problem solving tasks (Doctoral dissertation, Columbia University, 1973). *Dissertation Abstracts International,* 1973, *34,* 2400A. (University Microfilm No. 73-26,637)

Wickelgren, W. *How to solve problems.* San Francisco: Freeman, 1974.

Wilson, J. W. Generality of heuristics as an instructional variable (Doctoral dissertation, Stanford University, 1967). *Dissertation Abstracts,* 1968, *28,* 2575A. (University Microfilms No. 67-17,526)

[Received December 1977; revised January 1979]