Journal of Experimental Psychology:
Learning, Memory, and Cognition
1982, Vol. 8, No. 5, 484-494

0278-7393/82/0805-0484$00.75

Problem Perception and Knowledge Structure in Expert and Novice Mathematical Problem Solvers

Alan H. Schoenfeld
Graduate School of Education and
Department of Mathematics
University of Rochester

Douglas J. Herrmann
Hamilton College

Evidence regarding the relationships between problem perception and expertise has customarily been obtained indirectly, through contrasting-group studies such as expert-novice comparisons. Differences in perception have been attributed to differences in expertise, although the groups compared generally differ on a number of other major attributes (e.g., aptitude). This study explored the relationship between perception and proficiency directly. Students' perceptions of the structure of mathematical problems were examined before and after a month-long intensive course on mathematical problem solving. These perceptions were compared with experts' perceptions. Subjects sorted problems on the basis of similarity. Hierarchical clustering analysis of the sorting data indicated that novices perceive problems on the basis of "surface structure" (i.e., words or objects described in the problem statement). After the course the students perceived problem relatedness more like the experts—according to principles or methods relevant for problem solution. Thus, criteria for problem perception shift as a person's knowledge bases become more richly structured.

Theories of problem solving commonly hold that the mental representation of problems influences how people perceive problems. Moreover, as experience leads to better problem solving, the quality of problem representation is expected to improve with corresponding improvement in problem perception (Chi, Feltovich, & Glaser, 1981; Hayes & Simon, 1974; Heller & Greeno, 1979; Newell & Simon, 1972). At one end of the spectrum, the correct perception of a problem may cue access to a "problem schema" that suggests a straightforward method of solution or a more or less automatic response (Chase & Simon, 1973; Hinsley, Hayes, & Simon, 1977). At the other end, an incorrect perception may send one off on a "wild goose chase." Since problem perception is conceived to be a crucial component of problem solving performance, research on the change in problem perception with the acquisition of expertise has received increasing attention (Larkin, McDermott, Simon & Simon, 1980; Reif, 1980; Simon & Simon, 1978; Eylon & Reif, Note 1).

Early evidence consistent with the hypothesized relationship between expertise and perception was provided in a series of studies by Shavelson (1972, 1974; Shavelson & Stanton, 1975) that indicated that as students learn a discipline, their knowledge of the structural relationships among parts of the discipline become more like that of experts. However, Shavelson's (1972, 1974) procedures did not directly assess how his subjects perceived problems, and therefore his results do not directly address the perception/expertise hypothesis.

More direct evidence about problem perception and expertise has been provided by a series of studies in various domains that contrast the problem perceptions of a group of experts in each domain with the perceptions of a group of novices. For example, expert chess players perceive board positions in terms of patterns or broad arrangements,

We thank Chris Hempel, Mike Stankiewicz, Rob Kantrowitz, and Dorcque Schoenfeld for help on this research.

This material is based upon work supported by the National Science Foundation Grant SED-79 19049 at Hamilton College.

Requests for reprints should be sent to Alan H. Schoenfeld, Graduate School of Education and Department of Mathematics, University of Rochester, Rochester, New York. 14627.

whereas novices do not (Chase & Simon, 1973; de Groot, 1966). Experts in physics perceive problems to be similar if the principles used to solve them, called the "deep structure," coincide. In contrast, novices perceive them as similar if the objects referred to in the problem, or the terms of physics used in the statement, called the "surface structure," coincide (Chi, Feltovich, & Glaser, 1981). Two studies on problem perception in mathematics used algebra as their subject domain (Chartoff, 1977; Silver, 1979). There is a consensus regarding the structural isomorphism of algebra word problems, so in both studies problem structures were assigned a priori by the experimenters, and no experimental data were collected from experts. In both cases, students who were proficient at solving algebra word problems exhibited a greater degree of agreement with the experimenters' perceptions of the problems than did less proficient students.

The evidence regarding the relationship between expertise and perception, although strong, is indirect. Although expert–novice studies do show that experts and novices differ in problem perception, the design of these studies precludes unequivocal conclusions about the origins of these differences. For example, relative to novices, experts are usually older, more trained, more experienced, and most likely possessed of better aptitude for the subject domain.

Presumably, expert–novice differences in perception are rooted in differences in expertise (training and experience), but they may also be influenced by other psychological properties, for example, aptitude. Note that contrasting-group designs involving people of the same age may still confound expertise with aptitude. The ambiguous outcome of contrasting-group design is, of course, not unique to studies of expertise and problem perception; the difficulties of the design are well known, and in some areas of psychology these difficulties are regarded as presenting insurmountable obstacles to inference (Schaie, 1977). The present study sought to investigate the effects of expertise on perception in a design that avoids these difficulties—a design that examines problem perceptions in a group of individuals who, with training and experience, improved in problem-solving proficiency.

The relationship of perception and expertise was studied in a repeated measures design involving the discipline of college mathematics. Problem perception was assessed before and after training by having students sort a set of math problems. One group of students (the experimental group) took a month-long problem-solving course between the sortings. Another group (the control group) took a month-long course in computer programming between the sortings. In addition, a group of mathematics experts also completed the sort once. This study permits clear assessment of the relationship of problem perception and expertise: The influence of mathematical training on problem perception may be assessed by comparing the sorting of experimental and control subjects before and after training. If the experimental subjects show sorting after training different than control subjects, inferences about the mathematical improvements in the experimental subjects may be drawn relative to the sorting of the experts. Evidence showing that training affects problem perception and that training fosters problem perception of the kind that experts employ cannot be attributed in this study to differences in individuals (age, maturity, ability, and attentional levels). Although it is not suggested here that the findings of previous contrasting-group studies were not due to differences in expertise, the present procedure provides a clearer assessment of the relationship of problem perception and expertise.

Method

Subjects

Nineteen freshmen and sophomores at Hamilton College, novices, participated in the experiment. All of the students had 1–3 semesters of college mathematics prior to the experiment. Eleven of the students (the experimental group) served without pay as a condition of enrollment in a problem-solving course, which was the experimental treatment. Eight of the students (the control group) were paid a total of $20 each for participating. In addition, nine mathematics professors from Hamilton College and Colgate University participated without pay.

Materials

Thirty-two problems were chosen for the study. Each was accessible to students with a high school background in mathematics, and dealt with objects familiar from the high school curriculum; none required calculus for its solution. Each problem was assigned an a priori mathematical "deep structure" and a mathematical "surface structure" characterization. The problems used in the study are listed in Appendix A. (The characterizations of the problems may be seen in the cluster diagrams, Figures 1, 2, 3.)

"Deep structure" refers to the mathematical principles necessary for solution, as identified by the first author, who is a mathematician. For example, Problems 15 and 17 are both "uniqueness" arguments to be solved by contradiction, although Problem 15 deals with geometric objects and Problem 17 with functions. These characterizations were independently corroborated by another mathematician. Of the 32 problems, the deep-structure assessments were literally or essentially agreed upon by the other mathematician for all but three problems (which were perceived in a different but not contradictory fashion). This level of agreement on deep-structure assignments is comparable to that recently reported for physics problems (Chi et al., 1981). "Surface structure" represents a naive characterization of a problem, based on the most prominent mathematical objects that appear in it (polynomials, functions, whole numbers) or the general subject area it comes from (plane or solid geometry, limits). Thus Problems 15 and 17 discussed above would be considered a "plane geometry" and a "function" problem, respectively.

In addition, two forms of a mathematical problem-solving test were used in the study. The tests each had five problems worth 20 points, and were matched for mathematical content. These examinations and a predetermined scheme for awarding partial credit had been pilot tested, with the grading scheme achieving interjudge reliability of greater than .90. Form 1 of the test is given in Appendix B.

Procedure

Both the experimental and control groups performed the card sort and took Form 1 of the mathematics test immediately preceding the intensive winter term at Hamilton College. Both groups repeated the card sort and took Form 2 of the mathematics test a month later, immediately following the conclusion of the winter term. The experts performed the sort once, at their convenience.

The sorting procedures were as follows: Each of the 32 problems was typed on a 3- × 5-in. card. Each subject read through the problems in a random order and decided which problems, if any, were "similar mathematically in that they would be solved the same way." A problem that was deemed dissimilar to others was to be placed in a "group" containing one card. Subjects were told that they might return from 1 to 32 groups to the experimenter. All subjects finished the task in approximately 20 min.

Between the first and second sortings, the experimental treatment consisted of enrollment in a course, "Techniques of Problem Solving," taught by the first author. The class met for 2½ hours per day for 18 days, with daily homework assignments that averaged 4–5 hours in length. The course focused on general mathematical problem-solving strategies called "heuristics" (Pólya, 1957) and stressed a systematic, organized approach to solving problems (Schoenfeld, 1979; 1980). Problems studied in the course were similar to, but not identical to, those used in the sort; Appendix B gives five problems similar to those studied in the course. No mention of problem perception was made during the course. However, students were encouraged to make certain that they had a full understanding of the problem statement before proceeding with a solution. They were told to examine the conditions of the problem carefully, to look at examples to get a feel for the problem, to check for consistency of given data and plausibility of the results, etc. These instructions may well foster the development of improved problem perception.

The control treatment consisted of enrollment in a course, "Structured Programming." The course taught a structured, hierarchical, and orderly way to solve non-mathematical problems using the computer. The students in the course had backgrounds comparable to those of the students in the mathematical problem-solving course, and the course made similar demands in terms of time and effort from the students. Thus this course served as a control for the subject-specific knowledge and skills that might be acquired by the experimental group.

Results

For purposes of comparison with the results of the student sortings, we first present the results for the experts. Figure 1 presents a clustering analysis, using Johnson's (1967) method, of the experts' card sort. Collections of problems exhibiting strong agreement (proximity level exceeding .5—a minimum of 16 of 32 possible clusters) are bracketed. A brief inspection of Figure 1 indicates that the strong clusters are consistently homogeneous with regard to deep structure characterizations: In 8 of the 11 strong clusters, all of the elements share a common deep-structure characterization. In contrast, only 4 of the 11 strong clusters are homogeneous with regard to surface structure—and 3 of these with regard to deep structure as well.

Two measures of the degree of structural homogeneity of Figure 1 are given in Table 1. Measure 1 provides, for surface and deep structure respectively, the proportion of strongly clustered pairs that have the same structural representation.[1] Of the 22 pairs

[1] We wish to thank Jim Greeno for suggesting the measure and strengthening the discussion.

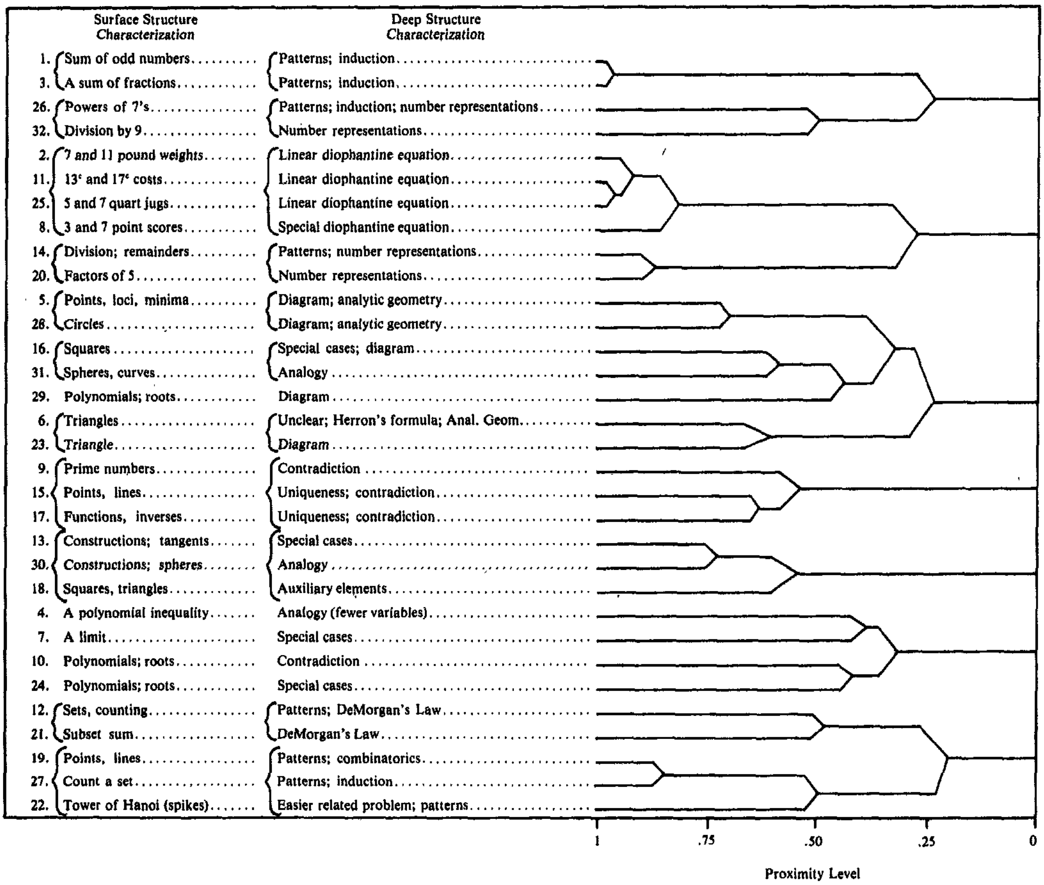


Figure 1. Cluster analysis of experts' card sort.

strongly clustered in Figure 1, 13 (.59) share the same surface structure and 18 (.82) the same deep structure. We should observe, however, that the surface and deep structures coincided in 10 of the 22 pairs used in the computation of Measure 1. To indicate perceptual preference when the two types of structures conflict, these 10 pairs were de-

Table 1
Proportion of Strongly Clustered Pairs in Which Both Problems Share the Same Representation

	Measure 1 (all pairs)				Measure 2 (noncoinciding pairs)			
	Surface structure		Deep structure		Surface structure		Deep structure	
Test group	%	*n*	%	*n*	%	*n*	%	*n*
Experts	.59	22	.82	22	.25	12	.67	12
Experimental, pretest	.81	26	.58	26	.58	12	.08	12
Control, pretest	.91	23	.57	23	.82	11	.09	11
Combined, pretest	.76	21	.62	21	.67	9	.11	9
Experimental, posttest	.58	24	.79	24	.09	11	.55	11
Control, posttest	.83	24	.58	24	.64	11	.09	11

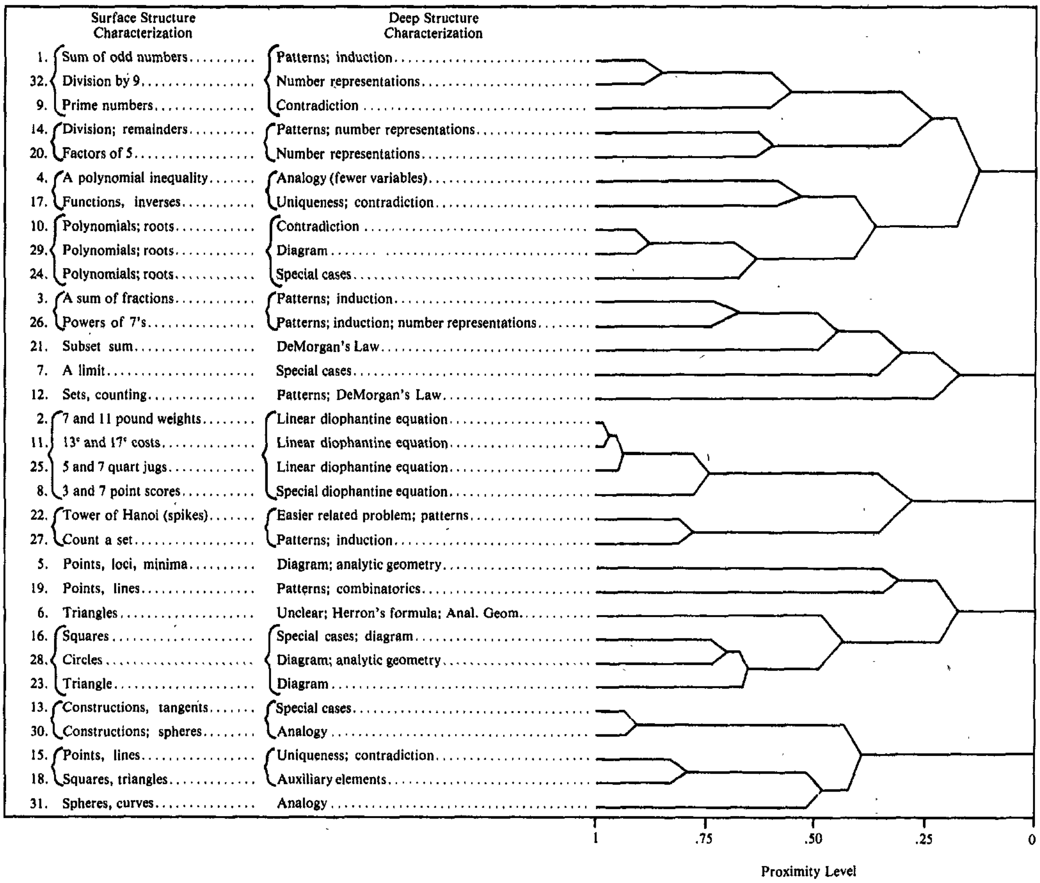


Figure 2. Cluster analysis of combined novices' card sort.

leted from the sample for Measure 2. With noncoinciding pairs, the proportion of surface-homogeneous pairings for the experts is .25 (3 of 12), and the proportion of deep-homogeneous pairs is .67 (8 of 12).

Figure 2 presents the cluster diagram of the sorting performed by the combined group of novices (n = 19) prior to instruction. In the interest of saving space, the cluster diagrams for the separate experimental and control groups are not given.[2] Inspection of Figure 2 indicates a reversal from Figure 1, with emphasis on surface structure as the criterion for sorting problems together: 8 of 10 strong clusters are homogeneous with regard to surface structure, 6 of 10 with regard to deep structure. Of these 6, 5 are also homogeneous with regard to surface structure. The data in Table 1 confirm these impressions. Table 1 also provides the data for the separate experimental and control groups prior to instruction. These data, like those for the combined group, indicate that the deep structural relationships between problems were rarely perceived when they ran in contradiction to perceptions of surface structure.

After training, the students who took the problem-solving course demonstrated a marked improvement in problem-solving performance, whereas those enrolled in the computer course did not. The mean scores on the mathematics test for the experimental subjects were 21 prior to the course and 73

[2] All three diagrams are quite similar. The matrix from which Figure 2 was derived was strongly correlated with both the experimental pretest matrix, $r(496) = .918$, $p < .001$, and the control pretest matrix, $r(496) = .889$, $p < .001$.

afterwards. For the control subjects, the mean scores were 14 before and 24 after the course. Analysis of variance on these means showed that scores increased across the term, $F(1, 17) = 47.5, p < .001$, were greater for experimental rather than for control subjects, $F(1, 17) = 130.6, p < .001$; and that the increase across the term was not equivalent for experimental and control subjects, $F(1, 17) = 48.2, p < .001$. Simple effects tests indicated that the term effect was significant for the experimental subjects ($p < .01$) but not for the control subjects. A detailed description of scoring procedures for this measure and of collateral measures may be found in Schoenfeld (1982).

The effect of instruction on problem perception was measured in the ways described above, and also by correlation with the experts' sorting matrix. Figure 3 presents the cluster analysis of the experimental group's sorting after instruction.

An examination of Figure 3 indicates the shift in the students' perceptions. After training, six of eight strong clusters were homogeneous with regard to deep structure, and only four with regard to surface structure; moreover, surface and deep structures coincided in all four of those clusters.

In contrast, the control group's postinstruction sorting shows little change from preinstruction perceptions. (Again to conserve space, the cluster diagram derived from that sorting, which closely resembles Figure 2, is not given. Of 10 strong clusters in it, 7 are homogeneous with regard to sur-

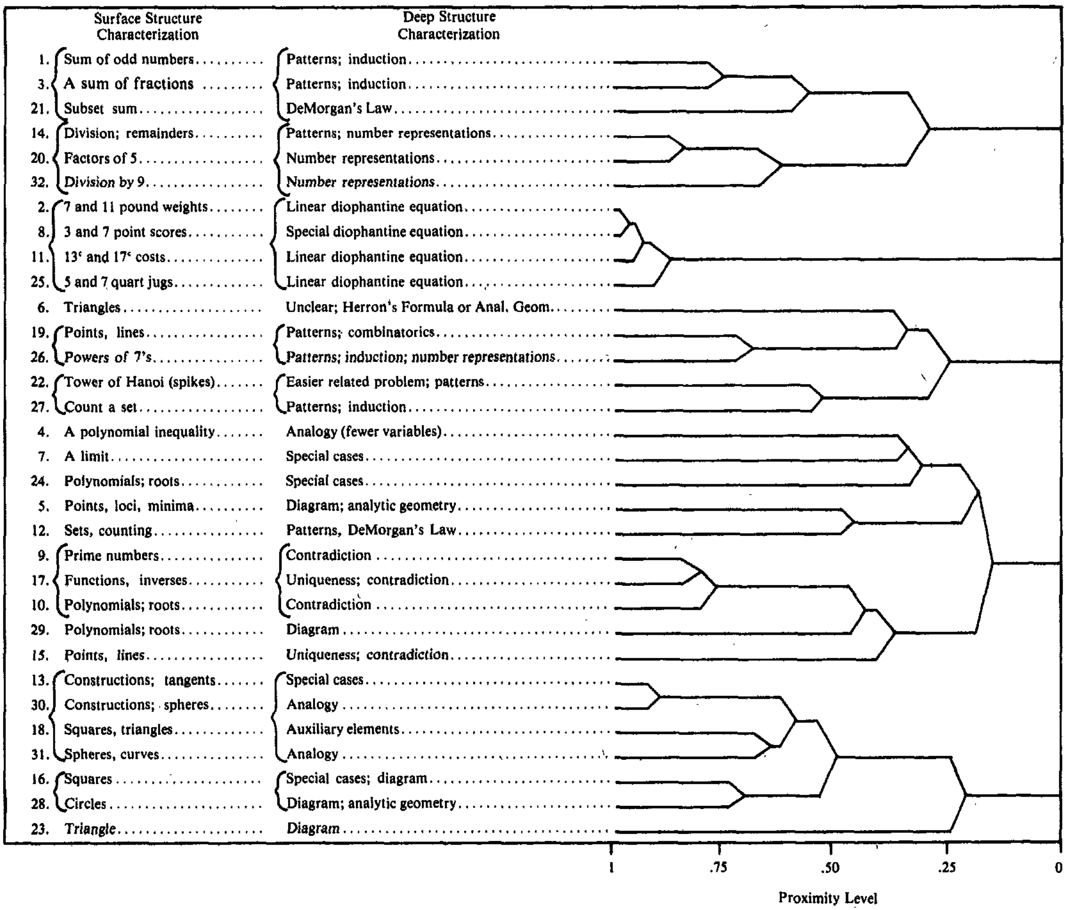


Figure 3. Cluster analysis of experimental group's card sort, after instruction.

face structure and only 4 with regard to deep structure; moreover those 4 share common surface structures as well.) These results, which indicate a strong change towards deep structure perceptions on the part of the experimental group and little or no change on the part of the control group, are given in Table 1. Differences between deep and surface proportions were compared across the various conditions with the t approximation to the binomial. Each of the following comparisons (with one exception noted) was significant to at least the ($p < .05$) level, both in direction and size of the differences. Scores within parentheses are reported first for Measure 1 (all pairs), then for Measure 2 (noncoinciding pairs). The difference for the experts differed, in direction and magnitude, with the difference in the preinstruction proportions from the experimental group, $t(18) = 2.33$; $t(18) = 4.09$; the control group, $t(15) = 2.61$; $t(15) = 4.98$, and the combined novice group, $t(26) = 1.88$, $p < .1$; $t(26) = 4.81$; also with the postinstruction difference from the control group, $t(15) = 2.31$; $t(15) = 3.99$. Similarly, the differences from the experimental group after instruction differed ($p < .05$) from the preinstruction differences from the experimental group, $t(11) = 2.38$; $t(11) = 4.51$, control group, $t(17) = 2.65$; $t(17) = 5.48$, and combined novice group, $t(28) = 2.39$; $t(28) = 5.41$; also from the control group's postinstruction scores, $t(17) = 2.34$; $t(17) = 4.37$.

The comparison of surface and deep-structure proportions given above provides an indirect indication that the experimental group's perceptions became more "expert-like" with instruction, whereas the control group's did not. This relationship was examined more directly by correlating the sorting matrices for each of the treatment groups, before and after instruction, with the sorting matrix obtained from the experts. The correlations were as follows: Control, pretest = .551; Experimental, pretest = .540; Combined, pretest = .602; Control, posttest = .423; Experimental, posttest = .723. With $df = 496$, all correlations are significant. The pretest correlations and the control posttest correlations are significantly less ($p < .01$) than the experimental group's posttest correlation.

Discussion

The design of this study allows for the direct attribution of the students' changes in problem perception to changes in their problem-solving proficiency. This attribution cannot be made unequivocally from any of the contrasting-group studies conducted to date, for example the standard expert–novice studies. Note that professors or advanced graduate students in a discipline differ from lower-division undergraduates in maturity, cohort group, comfort in testing situations, and most notably, aptitude. A clear understanding of how novices' performance improves in a discipline cannot be obtained by comparing them to a group of experts whose aptitude for the discipline is, in all likelihood, far beyond that of the novices. Similarly, an understanding of expert perception cannot be obtained by taking as the starting point of that development people whose performance alone makes it unlikely that they will ever be expert in that domain. One might obtain experimental confirmation of the relationship between perception and expertise in contrasting-group designs in which the groups had been matched on all variables except expertise (a difficult proposition, and a condition not present in any expert–novice studies with which we are familiar). However, the most direct way to ascertain that relationship is with a repeated measures (longitudinal) design like the one used here.

Two other points should be considered before the specifics of the data are elaborated. First, the nature of deep structure in mathematics is different from that of other domains. For example, elementary physics is strongly principle-driven, and the subject matter is organized and taught according to those principles. Mathematics is not organized and taught that way, however. One talks about methods of solution, rather than principles; and the curriculum is organized around topics rather than around those methods, which are simply the tools used to solve them. Thus, there does not exist an a priori consensus about the structure of the problems used in this study that would lead one to predict with confidence the particular pattern of results repeated in Figure 1. The absence of such a consensus makes the consistency of the present results more impressive. The word "novice" in this study does not mean "rank beginner"; the students in this study had extensive mathematical backgrounds and were, in the sort task, reading problems accessible to them. The surface labels reflect this, for example in the labels for Problems 2 and 11. Surely, one would be surprised if college students could not see that integer combinations of weights and integer combinations of costs called for the same mathematics! (This would not necessarily be the case with fifth graders, for example).

The data in Table 1 provide a strong indication that the experimental group's perceptions of problem structure shifted from a basis in surface structure to a basis in deep structure. An examination of the experimental group's postinstruction cluster (9, 17, 10) illustrates the change in problem perception. Problem 9 deals with whole numbers and, prior to instruction, was sorted with two other whole-number problems in a homogeneous surface structure cluster. Problem 17 deals with abstract functions, and, prior to instruction, was (barely) clustered with a problem that presented a very complex polynomial function for analysis. Problem 10 deals with polynomials, and was placed in a strong cluster all three of whose terms had the surface label "polynomials, roots." Each of these problems is solved by the mathematical techique known as proof by contradiction and, despite their differing surface characterizations, they are all placed in the same cluster after instruction. The broad shift towards expert perceptions is confirmed (see data above), which shows that the correlation between experts' and the experimental group's sorting matrices jumped from .540 (before instruction) to .723 (after instruction)—the only significant ($p < .01$) change in correlation. This rather dramatic shift after a short period of time indicates that instructional treatments that focus on understanding and performance can have a strong impact on perceptions.

Despite the strong shift in the students' sort, the experimental group's performance after instruction cannot be truly called expertlike. The experts' extended knowledge and experience allow them perceptions inaccessible to the novices. Consider, for example, the three braced clusters in Figures 2, 3, and 1 respectively that include Problem 1: novice (1, 32, 9); experimental (1, 3, 21); expert (1, 3). The experimental group drops Problems 32 and 9, which are similar to Problem 1 only in that they deal with whole numbers. Problem 3, which shares the same deep structure as Problem 1, is added. The mimicry of expert perceptions is not exact, however: Problem 21 is added as well. The addition of Problem 21 provides an indication of the "intermediate" status of the experimental group. Problems 12 and 21 were included in the card sort to see if the experts would cluster them together. Underlying the experts' perception of Problem 21 is the observation that multiples of 9 and multiples of 4 both include multiples of 36 (their intersection), and that one must compensate for subtracting the first two sets by adding the third. This is structurally similar to the rule $N(\mathbf{A} \cup \mathbf{B}) = N(\mathbf{A}) + N(\mathbf{B}) - N(\mathbf{A} \cap \mathbf{B})$ upon which Problem 12 is based. This is a rather subtle observation. Although experts' experience with combinatorics problems might make such an observation readily accessible, novices even with training cannot be expected to see such subtleties. In the absence of such knowledge, it is plausible to think that "looking for patterns" will help to solve Problem 21—and thus to sort it with two other "patterns" problems.

The research described here supports and extends previous research on problem perception. The novices' card sort indicated that, in the broad domain of general mathematical problem solving, students with similar backgrounds will perceive problems in similar ways. This is consistent with previous research in mathematics, which had considered only word problems in algebra (Chartoff, 1977; Hinsley et al., 1977; Silver, 1979). Like research in physics (Chi et al., in press), it suggests that surface structure is a primary criterion used by novices in determining problem relatedness. Moreover, it verifies directly that students' problem perceptions change as the students acquire problem-solving expertise. Not only their performance, but their perceptions, become more like experts'.

In general, questions regarding the deep structures in individual disciplines and the nature of experts' perceptions in those disciplines are more complex than those regarding surface structures and novices' perceptions in them. The differences between the structures of mathematics and physics were discussed above. In another discipline, research on chess perception (Chase & Simon, 1973; de Groot, 1966) indicates that experts' perceptions of routine problems (similar in a way to the routine physics and mathematics problems discussed above) may be based on the acquisition of a "vocabulary" of known situations; this "vocabulary" is not necessarily principle based. Further research might profitably be directed toward the elucidation of how deep structures differ across disciplines and how problem perceptions evolve with the acquisition of expertise in different domains.

Reference Note

1. Eylon, B., & Reif, F. *Effects of internal knowledge organization on task performance.* Unpublished manuscript, 1979. (Available from F. Reif, Department of Physics, University of California, Berkeley, Berkeley, California 94720.)

References

Chartoff, B. T. An exploratory investigation utilizing a multidimensional scaling procedure to discover classification criteria for algebra word problems used by students in grades 7-13 (Doctoral dissertation, Northwestern University, 1976). *Dissertation Abstracts International*, 1977, *37*, 7006A. (University Microfilms No. 77-10,012).

Chi, M., Feltovich, P., & Glaser, R. Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 1981, *5*, 121-152.

Chase, W. G., & Simon, H. A. Perception in chess. *Cognitive Psychology*, 1973, *4*, 55-81.

de Groot, A. Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmunz (Ed.), *Problem solving.* New York: Wiley, 1966.

Hayes, J. R., & Simon, H. A. Understanding written problem instructions. In L. W. Gregg (Ed.), *Knowledge and cognition.* Hillsdale, N.J.: Erlbaum, 1974.

Heller, J. I., & Greeno, J. G. Information processing analyses of mathematical problem solving. In R. Lesh, D. Mierkiewicz, & M. Kantowski (Eds.), *Applied mathematical problem solving.* Columbus, Ohio: ERIC, 1979. (ERIC Document Reproduction Service No. ED 180 816)

Hinsley, D. A., Hayes, J. R., & Simon, H. A. From words to equations: Meaning and representation in algebra word problems. In P. A. Carpenter & M. A. Just (Eds.), *Cognitive processes in comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, *32*, 241-54.

Larkin, J., McDermott, J., Simon, D., & Simon, H. A. Expert and novice performance in solving physics problems. *Science*, 1980, *208*, 1335-1342.

Newell, A., & Simon, H. S. *Human problem solving.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Pólya, G. *How to solve it* (2nd ed.). Garden City, N.Y.: Doubleday, 1957.

Reif, F. Theoretical and educational concerns with problem solving: Bridging the gap with human cognitive engineering. In D. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research.* Hillsdale, N.J.: Erlbaum, 1980.

Schaie, K. W. Quasi-experimental research designs in the psychology of aging. In J. E. Birren, and K. W. Schaie (Eds.), *Handbook of the psychology of aging.* New York: Van Nostrand Reinhold, 1977.

Schoenfeld, A. H. Can heuristics be taught? In J. Lochhead & J. Clement (Eds.), *Cognitive process instruction.* Philadelphia, Pa.: Franklin Institute Press, 1979.

Schoenfeld, A. H. Teaching problem solving skills. *American Mathematical Monthly*, 1980, *87*(10), 794-805.

Schoenfeld, A. H. Measures of problem solving performance and of problem solving instruction. *Journal for Research in Mathematics Education*, 1982, *13*, 31-49.

Shavelson, R. J. Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 1972, *63*, 225-234.

Shavelson, R. J. Methods for examining representations of a subject-matter structure in a student's memory. *Journal of Research in Science Teaching*, 1974, *11*, 231-249.

Shavelson, R. J., & Stanton, G. C. Construct validation: Methodology and application of three measures of cognitive structure. *Journal of Educational Measurement*, 1975, *12*, 67-85.

Silver, E. A. Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education*, 1979, *10*, 195-210.

Simon, D. P., & Simon, H. A. Individual differences in solving physics problems. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.

Appendix A: Problems Used in Card Sort

1. Show that the sum of consecutive odd numbers, starting with 1, is always a square. For example,

$$1 + 3 + 5 + 7 = 16 = 4^2.$$

2. You have an unlimited supply of 7-pound weights, 11-pound weights, and a potato that weighs 5 pounds. Can you weigh the potato on a balance scale? A 9-pound potato?

3. Find and and verify the sum,

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

4. Show that if x, y, and z are greater than 0,

$$\frac{(x^2+1)(y^2+1)(z^2+1)}{xyz} > 8.$$

5. Find the smallest positive number m such that the intersection of the set of all points $\{(x, mx)\}$ in the plane, with the set of all points at distance 3 from (0,6), is nonempty.

6. The lengths of the sides of a triangle form an arithmetic progression with difference d. (That is, the sides are a, $a + d$, $a + 2d$.) The area of the triangle is t. Find the sides and angles of this triangle. In particular, solve this problem for the case $d = 1$ and $t = 6$.

7. Given positive numbers a and b, what is

$$\lim_{n\to\infty}(a^n + b^n)^{1/n}?$$

8. In a game of "simplified football," a team can score 3 points for a field goal and 7 points for a touchdown. Notice a team can score 7 but not 8 points. What is the largest score a team *cannot* have?

9. 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.

10. Prove that there are no real solutions to the equation

$$x^{10} + x^8 + x^6 + x^4 + x^2 + 1 = 0$$

11. If Czech currency consists of coins valued 13 cents and 17 cents, can you buy a 20-cent newspaper and receive exact change?

12. If N(A) means "The number of elements in A," then $N(A \cup B) = N(A) + N(B) - N(A \cap B)$. Find a formula for $N(A \cup B \cup C)$.

13. Construct, using straightedge and compass, a line tangent to two given circles.

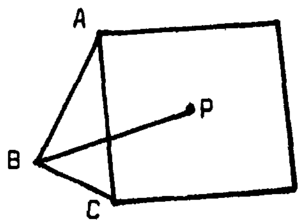
14. Take any odd number; square it; divide by 8. Can the remainder be 3? or 7?

15. You are given the following assumptions: (i) Parallel lines do not intersect; nonparallel lines intersect. (ii) Any two points P and Q in the plane determine an unique line which passes between them.
Prove: Any two distinct nonparallel lines L_1 and L_2 must intersect in an unique point P.

16. Two squares "s" on a side overlap, with the corner of one on the center of the other. What is the maximum area of possible overlap?

17. Show that if a function has an inverse, it has only one.

18. Let P be the center of the square constructed on the hypotenuse AC of the right triangle ABC. Prove that BP bisects angle ABC.



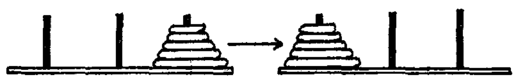
19. How many straight lines can be drawn through 37 points in the plane, if no 3 of them lie on any one straight line?

20. If you add any 5 consecutive whole numbers, must the result have a factor of 5?

21. What is the sum of all numbers from 1 to 200, which are not multiples of 4 and 9? You may use the fact that

$$(1 + 2 + \cdots + n) = 1/2\,(n)(n+1)$$

22. Your goal is to convert the figure on the left to the one on the right. You may move only one disk at a time from one spike to another, and you may never put a larger disk on top of a smaller one. How to?



23. Determine the area of a triangle whose sides are given as 25, 50, and 75.

24. If $P(x)$ and $Q(x)$ have "reversed" coefficients, for example

$$P(x) = x^5 + 3x^4 + 9x^3 + 11x^2 + 6x + 2,$$

$$Q(x) = 2x^5 + 6x^4 + 11x^3 + 9x^2 + 3x + 1,$$

What can you say about the roots of $\mathbf{P}(x)$ and $\mathbf{Q}(x)$?

25. You have 2 unmarked jugs, one whose capacity you know to be 5 quarts, the other 7 quarts. You walk down to the river and hope to come back with precisely 1 quart of water. Can you do it?

26. What is the last digit of $(\cdots((7^7)^7)^7\cdots)^7$, where the 7th power is taken 1,000 times?

27. Consider the following magical configuration. In how many ways can you read the word "ABRACADABRA?"

A
B B
R R R
A A A A
C C C C C
A A A A A A
D D D D D
A A A A
B B B
R R
A

28. A circular table rests in a corner, touching both walls of a room. A point on the rim of the table is eight inches from one wall, nine from the other. Find the diameter of the table.

29. 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.

30. Describe how to construct a sphere that circumscribes a tetrahedron (the 4 corners of the pyramid touch the sphere.)

31. Let S be a sphere of radius 1, A an arc of length less that 2 whose endpoints are on the boundary of S. (The interior of A can be in the interior of S.) Show there is a hemisphere H which does not intersect A.

32. Show that a number is divisible by 9 if and only if the sum of its digits is divisible by 9. For example, consider 12345678: $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 = 36 = 4 \times 9$, so 12345678 is divisible by 9.

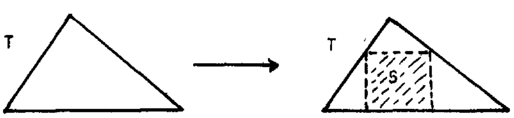
Appendix B: Mathematics Test Form 1

1. If S is any set, we define $O(S)$ to be the number of subsets of S which contain an odd number of elements. For example: the "odd" subsets of $\{A, B, C\}$ are $\{A\}$, $\{B\}$, $\{C\}$, and $\{A, B, C\}$; thus $O(\{A, B, C\}) = 4$. Determine $O(S)$ if S is a set of 26 objects.

2. Suppose you are given the positive numbers p, q, r, and s. Prove that

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

3. Suppose T is the triangle given at the left below. Give a mathematical argument to demonstrate that there is a square, S, such that the 4 corners of S lie on the sides of T, as in the figure to the right.

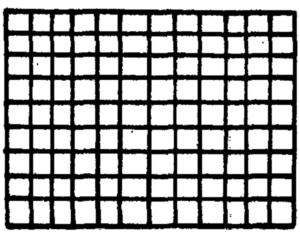


4. Consider the set of equations

$$\begin{Bmatrix} ax + y = a^2 \\ x + ay = 1 \end{Bmatrix}.$$

For what values of a does this system fail to have solutions, and for what values of a are there infinitely many solutions?

5. Let G be a (9×12) rectangular grid, as shown below. How many different rectangles can be drawn on G, if the sides of the rectangles must be grid lines? (Squares are included, as are rectangles whose sides are on the boundaries of G.)



Received December 22, 1980
Revision received December 9, 1981 ■