

Learning to solve insight problems

MARY K. JACOBS and ROGER L. DOMINOWSKI
University of Illinois at Chicago Circle, Chicago, Illinois 60680

A series of seven insight problems was given to 28 male and 28 female college students. Each problem required that an object be used in a novel manner to reach solution, but both the critical object and its required use varied across problems. Subjects showed moderate improvement in performance on the second half of the problem series. Males solved more quickly than females; relatively stable and predictable individual differences in performance were observed. Results support the conclusions that insight problems form a class involving common abilities and that general transfer effects can occur if sufficient practice is given.

The Gestalt psychologists introduced a number of tasks requiring subjects to use an object in some relatively unusual way to solve a problem. Such tasks are typically called "insight problems" (although the term has been applied to other tasks as well). Object-use insight problems were used successfully to study functional fixedness and problem solving hints (e.g., Burke, Maier, & Hoffman, 1966; Glucksberg & Weisberg, 1966), but interest in the process of solving insight problems has declined. One likely reason for the lowered interest lies in the failure to obtain results that might plausibly be expected on the basis of theoretical analyses.

For example, if solving an insight problem involves thinking of an unusual use for an object, then it might reasonably be expected that fostering the production of unusual ideas should facilitate solution. However, having subjects give unusual free associations (Maltzman, Belloni, & Fishbein, 1964) or unusual uses for objects (Duncan, 1961) has not facilitated performance. Perhaps more critically, Duncan (1961) found that prior work on one problem had no effect on performance on a second problem.

In addition, there have been failures to find correlates of performance on insight problems. Duncan (1961) found no significant interproblem correlations, and Burke and Maier (1965) found no reliable correlations between performance on the hat-rack problem and a number of predictors, including the Gestalt Transformations Test (Wilson & Guilford, 1969), which requires the subject to select objects having parts that can be used to satisfy stated requirements (and thus seems as if it should be related to using objects to solve problems).

These negative findings suggest that solving any particular insight problem is a rather unique behavioral event based on the utilization of specific information. If this analysis is accurate, then one might find specific

transfer effects but not more general transfer (see Weisberg & Alba, in press) and no substantial correlations. For example, in contrast to the above-mentioned failures to find general transfer, DiVesta and Walls (1967) found that giving subjects experience with the specific unusual use required for solution did increase the probability that the critical object would be used to solve a subsequent problem.

In Greeno's (1977) discussion of problem solving abilities, object-use insight problems are included as a subclass of arrangement problems and are assumed to involve abilities such as fluency in generating arrangements, knowledge and use of constraints in evaluating arrangements, flexibility in considering unusual alternatives, and skill in comprehending new relations. While Greeno's analysis is eminently plausible, we currently lack evidence that insight problems constitute a psychologically meaningful subclass of problems involving abilities of some reasonable level of generality.

The existing, negative evidence is, however, open to the criticism that previous efforts have not been sufficient to detect the kinds of relations that might be expected. Duncan's (1961) failure to find interproblem transfer must be considered in light of the fact that, although the transfer task was an object-use insight problem (the two-string problem), the training task for most subjects was not such a problem. More important, Duncan gave subjects only two problems, whereas research with other tasks indicates that general transfer increases with the number and variety of tasks encountered (e.g., Bourne, 1970; Duncan, 1958; Harlow, 1949). Similarly, the failures to find significant correlates of performance on insight problems seem less critical because of the use of insensitive "pass-fail" measurement, a single test problem, and homogeneous groups of subjects. The major purpose of the present study was to make a more substantial attempt to study general transfer effects and consistency of individual differences across a set of object-use insight problems.

In the present study, each subject received a series of seven insight problems, each involving multiple objects and requiring some critical object to be used in a rela-

tively unusual fashion to achieve solution. Across tasks, both the critical object and the particular use needed for solution differed. Under these circumstances, transfer of specific information from one problem to another would not be helpful. If, however, solving such problems involves some more general processes, such as those suggested by Greeno (1977), then one might expect that, after experience with a number of problems, some improvement in performance would be observed.

Two variables of secondary interest were included in the design. It has often been found that males perform better than females on insight problems (e.g., Duncan, 1961). Therefore, subjects were stratified according to gender to allow replication of this result, to reduce error variance, and to determine if females might profit more from practice. In addition, although all problems were presented concretely, subjects were either required to or not allowed to physically manipulate the objects while working on a problem. This was done to test the idea that manipulation might provide subjects with a changing perceptual field that could facilitate solution (Gavurin, 1967).

METHOD

Subjects and Design

The subjects were 28 male and 28 female introductory psychology students who participated to fulfill a course requirement. Half of the subjects in each group worked under the manipulation condition, and the other half worked under the nonmanipulation condition. Each subject received seven tasks; within each of the four basic groups (male-female by manipulation-nonmanipulation), each subject received a different order of the seven tasks, with tasks and ordinal positions counterbalanced across the group.

Apparatus

The tasks, all presented concretely, were adapted from those used in prior research in an attempt to construct problems that had only one solution and that might reasonably be solved within 15 min. The tasks were: the box-candle problem (Maier, 1933), flower-stand problem (Duncker, 1945), two-string problem (Maier, 1933), hat-rack problem (Maier, 1933), marble-transfer problem (Saugstad & Raaheim, 1957), candle problem (Saugstad, 1955), and truck-retrieval problem (Saugstad & Raaheim, 1957).

Records of solution times and manipulation periods (when appropriate) were obtained through the use of a multichannel recorder. One channel was used for solution time; for the manipulation condition, two additional channels were used to record periods of manipulation of the critical object and of other objects, respectively. The experimenter operated silent buttons on a hand-held panel to activate the recorder, which was located in a room separate from those in which problems were presented.

Procedure

Each subject was seen individually and was given general instructions prior to the first task. It was explained that the subject would be given several problems that required some sort of physical arrangement of objects for their solution. The subject was told that it was not necessary to complete the physical arrangement, but rather, that the problem would be considered as solved when the subject announced the steps that would be taken to meet the task requirements. Under the manipulation condition, the subject was told that moving the

objects around while working on the task was required, but it was emphasized that physically completing the solution was not required, and that solutions should be announced as soon as they were conceived. In the nonmanipulation condition, the subject was told that none of the objects could be touched. Brief reminders of these instructions were given prior to the second and third problems.

In presenting each task, the experimenter included references to the objects present and their dimensions. The subject was told to announce solutions when they occurred and that it was desirable to describe any solution that came to mind. If the subject offered an incorrect solution, the experimenter indicated that it was wrong, briefly explained why, and encouraged the subject to continue to work on the problem. For each problem, a maximum of 15 min was allowed; the experimenter provided the solution if the subject had not solved the problem within this limit. After each task, the subject was taken to a waiting room while the next task was prepared; the intertask interval was approximately 2 min. After the last task, subjects were given the Gestalt Transformations Test (Wilson & Guilford, 1969). Each item on this test requires the subject to select the object (from four alternatives) that has a part that could serve the function named in the stem of the item.

RESULTS

Each subject provided solution times for seven tasks (in case of nonsolution, a score of 900 sec was recorded). Because solution times tend to be positively skewed, the raw data were subjected to a logarithmic transformation to normalize the scores and reduce the influence of extreme values. The mean solution times reported are geometric means based on the transformed data.

Overall Performance

There were significant overall differences in performance between the sexes and among the seven tasks employed (Table 1). The mean solution time for males (170 sec) was lower than that for females (236 sec) [$F(1,288) = 9.78, p < .01$], and the problems differed considerably in difficulty [$F(6,288) = 48.87, p < .001$].¹ In addition, the difference between males' and females' performance varied somewhat across tasks [$F(6,288) = 3.09, p < .01$] (see Footnote 1). In Table 1, it can be seen that the order of difficulty of the seven tasks was the same for both sexes, with males' performance superior on all but one problem. However, large differences favoring males occurred only for the two-string and marble-transfer problems.

Table 1
Mean Solution Times (in Seconds)

| Task | Males | Females |
|-----------------|-------|---------|
| Hat Rack | 724 | 776 |
| Flower Stand | 427 | 372 |
| Box-Candle | 355 | 372 |
| Two String | 141 | 355 |
| Marble Transfer | 91 | 204 |
| Truck Retrieval | 78 | 87 |
| Candle | 34 | 48 |

The mean solution time for the manipulation condition (206 sec) was negligibly higher than that for the nonmanipulation condition (193 sec) ($F < 1$), and there were no significant interactions involving this variable. Overall, subjects in the manipulation condition handled the critical object 23.3% of the time they spent working on the problems; they handled other objects 47.1% of the time. Across tasks, percent manipulation time for the critical object was inversely related to percent manipulation time for other objects ($r = -.43$). However, none of several analyses suggested any systematic relation between the nature of manipulative activities and differences in problem solving performance. Differential manipulative opportunity seemed to have no influence on problem solving efficiency.

Individual Differences

Across the seven tasks, individual differences in performance were moderately stable. The interproblem correlations were all positive, with the exception of that between the flower-stand and box-candle problems ($r = -.04$); the highest correlation occurred between the hat-rack and two-string problems ($r = +.50$). When a single common factor was extracted from the matrix of interproblem correlations (DuBois, 1965), all problems had positive loadings ranging from .16 to .76. In addition, individual differences in overall solution times were significantly related to scores on the Gestalt Transformations Test ($r = -.41$, $p < .01$). These results suggest that some common abilities underlie performance on object-use insight problems.

Performance on the Gestalt Transformations Test was also helpful in understanding the difference between males and females in problem solving efficiency. On the test, the mean score for males (12.2) was higher than that for females (8.0) [$F(1,54) = 23.50$, $p < .001$, $MSe = 10.40$]. The correlation between overall solution times and test scores, when computed separately for females ($r = -.32$) and for males ($r = -.42$), was similar to that obtained for all subjects combined. Within each group, problem solving efficiency was related to test scores, with males tending as a group to have higher test scores and greater efficiency than females.

Practice Effects

For problems in Ordinal Positions 1-7, the mean solution times were 248, 225, 247, 192, 152, 162, and 188 sec, respectively. Performance on the last three problems was better than that on the first three problems [$F(1,288) = 8.43$, $p < .01$] (see Footnote 1). This contrast, which accounts for 82% of the variation among ordinal-position means, reflects moderate improvement in problem solving efficiency over the series of seven problems.² To examine practice effects for individual subjects, each subject's score for each task was standardized relative to the mean and standard deviation for that task for all subjects combined (recall that subjects received the seven tasks in different orders and that tasks differed greatly in difficulty). Analysis

Table 2
Mean Solution Times (in Seconds) for Successive Problems

| | Ordinal Position | | | | | | |
|---------|------------------|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Males | 200 | 174 | 191 | 148 | 145 | 145 | 170 |
| Females | 309 | 282 | 309 | 240 | 162 | 182 | 219 |

of the standardized scores confirmed the result of the initial analysis, namely, better performance on the last three problems compared with the first three problems [$F(1,52) = 8.77$, $p < .01$] (see Footnote 2). Of the 56 subjects, 35 showed better performance on the second half of the problem series. Although females showed numerically greater improvement than males (Table 2), the interaction effect fell far short of significance ($p < .25$).

DISCUSSION

The present research produced two major results: (1) individual differences in performance are relatively stable across tasks and are predictable. (2) Subjects improve over a series of problems. These findings seem to contradict prior results, but they are not actually inconsistent with earlier studies; rather, the use of more problems and more sensitive measurement led to positive findings.

In contrast to earlier reports (Burke & Maier, 1965; Duncan, 1961), the present data show that individual differences in performance are relatively consistent across tasks and are correlated with scores on the Gestalt Transformations Test. Previous research employed just one object-use insight problem and a pass-fail measure, which led to insensitive tests. In the present study, both the correlations between problems and those between the test and problem performance varied over tasks. However, a common factor could be extracted from the intertask correlations, and scores on the Gestalt Transformations Test were clearly related to subjects' average solution times for the seven tasks (the most reliable measure of problem solving efficiency). It seems reasonable to conclude that individual differences in solving insight problems are somewhat stable and predictable, which supports Greeno's (1977) contention that insight problems form a meaningful subclass of arrangement problems. Whether the abilities involved in solving these problems are those suggested by Greeno remains to be determined.

In the present study, subjects showed modest improvement over a series of seven problems that had in common only the characteristic that each required some object to be used in a relatively novel manner. Since both the critical objects and their required uses differed among problems, the finding demonstrates positive interproblem transfer at the most general level yet reported for insight problems. Duncan (1961) failed to obtain positive transfer, but he gave subjects only two problems, only one of which was an object-use insight problem. In the present study, as in Duncan's, there was no improvement from the first to the second problem: performance did not improve until subjects had experienced several problems. The present study did not provide direct evidence regarding how subjects might have improved with practice. A number of alternatives might be considered, including greater fluency in generating arrangements, better discrimination of critical from irrelevant objects, and realization that relatively simple but unusual object uses are required. These possibilities were sometimes mentioned by subjects in informal postsession interviews, and they bear certain similarities to the processes suggested by Greeno (1977) as relevant to solving insight problems.

It might prove useful to have subjects attempt a series of

problems under thinking-aloud conditions. To the extent that relevant processes can operate at a conscious level, hypothesized bases for improvement could be tested by giving direct instructions that emphasize different aspects of the total solution process. In pursuing such research, it would be necessary to give subjects a number of problems, since the present results indicate that differences might not be apparent with a single test problem.

Finally, it must be noted that, although the majority of subjects improved with practice in the present study, improvement was by no means uniform. Differences in improvement were not systematically related to measures available (i.e., subjects' gender, scores on the Gestalt Transformation Test, the nature of manipulative activities). There is a need to discover both what subjects who do improve and subjects who do not improve do differently while working on the problems and what predictor variables are related to improvement.

REFERENCES

- BOURNE, L. E., JR. Knowing and using concepts. *Psychological Review*, 1970, **77**, 546-556.
- BURKE, R. J., & MAIER, N. R. F. Attempts to predict success on an insight problem. *Psychological Reports*, 1965, **17**, 303-310.
- BURKE, R. J., MAIER, N. R. F., & HOFFMAN, L. R. Functions of hints in individual problem solving. *American Journal of Psychology*, 1966, **79**, 389-399.
- DIVESTA, F. J., & WALLS, R. T. Transfer of object-function in problem solving. *American Educational Research Journal*, 1967, **4**, 596-602.
- DUBOIS, P. H. *An introduction to psychological statistics*. New York: Harper & Row, 1965.
- DUNCAN, C. P. Transfer after training with single versus multiple tasks. *Journal of Experimental Psychology*, 1958, **55**, 63-72.
- DUNCAN, C. P. Attempts to influence performance on an insight problem. *Psychological Reports*, 1961, **9**, 35-42.
- DUNCKER, K. On problem-solving. *Psychological Monographs*, 1945, **58**(5, Whole No. 270).
- GAVURIN, E. I. Anagram solving under conditions of letter order randomization. *Journal of Psychology*, 1967, **65**, 179-182.
- GLUCKSBERG, S., & WEISBERG, R. W. Verbal behavior and problem solving: Some effects of labeling in a functional fixedness problem. *Journal of Experimental Psychology*, 1966, **71**, 659-664.
- GREENO, J. G. Natures of problem-solving abilities. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. New York: Academic Press, 1977.
- HARLOW, H. F. The formation of learning sets. *Psychological Review*, 1949, **56**, 51-65.
- MAIER, N. R. F. An aspect of human reasoning. *British Journal of Psychology*, 1933, **24**, 144-155.
- MALTZMAN, I., BELLONI, M., & FISHBEIN, M. Experimental studies of associative variables in originality. *Psychological Monographs*, 1964, **78**(3, Whole No. 580).
- SAUGSTAD, P. Problem solving as dependent on availability of functions. *British Journal of Psychology*, 1955, **46**, 191-198.
- SAUGSTAD, P., & RAAHEIM, K. Problem solving and availability of functions. *Acta Psychologica*, 1957, **13**, 263-278.
- WEISBERG, R. W., & ALBA, J. W. An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General*, in press.
- WILSON, R., & GUILFORD, J. P. *Gestalt transformations test*. Beverly Hills, Calif: Sheridan Psychological Services, 1969.

NOTES

1. For these tests, the pooled Latin rectangle error term was used; $MSe = .23$ with respect to log solution time.

2. A significant linear trend could also be detected in the data; however, it is felt that the contrast between first-half and second-half performance better describes the results in this case.

(Received for publication March 1, 1981.)