PROBLEM SOLVING AND PRIOR LEARNING

JOSEPH M. SCANDURA
University of Pennsylvania

This study was concerned with how learning to solve problems depends on prior learning. In exploring a related problem, Scandura (12, 13) found evidence which suggests that mere presentation of subordinate material is not always sufficient to ensure subsequent learning when the terms (i.e. symbols and words), denoting the subordinate notions, are used to describe the higher order material. The implications of this observation, if borne out experimentally, could be substantial. Thus, rapidly covering preliminary material, so as to get to the main topic more quickly, may in the long run defeat its own purpose.

Of course, good teachers have long recognized that when material is introduced determines to a large extent how well it is learned. Nonetheless, to the author's knowledge, there are no published experimental studies which bear directly on the present problem. This does not mean that related questions have gone unnoticed. Kersh (10) reviewed some earlier studies (4, 16) which were concerned with the effects of familiar and unfamiliar settings on problem solving in arithmetic. These studies were, in effect, concerned with relationships between superordinate material (the problems) and prerequisites, but the results were necessarily in the form of R-R laws. No attempt was made to manipulate familiarity. The recent work of Gagné et al. (6, 7, 8, 9) and Ausubel et al. (1, 2, 3) was also related to this research in that their concern was with the relationships between subordinate and superordinate learning. Gagné, however, was concerned with analyzing complex cognitive skills into component skills which were more readily taught and not with the presentation of continuous discourse in terms of subordinate associations, concepts, and principles. Ausubel dealt with the effects of "advance organizers", which consist of general introductory material on later learning, but not with the situation described.

The purpose of this experiment was to determine the effects on the learning, transfer, and retention of problem solving skills of practice at both the prerequisite and criterion levels in a learning hierarchy.

METHOD

Materials, Apparatus, and Subjects

The materials used in this experiment were based on some abstract materials (12) developed for use in problem solving research. This material was based on an extension of card tasks frequently used to study concept formation (5). Sixty-four cards were used. Each card had two objects on it, six two-valued attributes in all. Corresponding to each of the twelve attribute values was a familiar symbol. Cards were assigned one of several possible labels. A symbol set was said to be associated with a given labeled card if the symbols in the set were related to the card properties in a specified way. A card problem consisted of determining all symbol sets associated with each card comprising the problem. There was an algorithm for solving each of the many possible problem types. Problem types differed in the card labels used and/or the algorithm pattern and number of solution sets.

For the purposes of the present experiment, the material was divided into three levels. The level one material was used to describe the level two and three material and the level two material was used to describe the level three material. At level one were the twelve one-to-one correspondences between the card properties and the symbols which represented those properties. There was a common principle relating each property to the corresponding symbol. For example, S1 and T2 stood for the first object is Small and the second is a Triangle, respectively. Level two consisted of defining the problems (i.e. finding all symbol sets associated with each of three labeled cards) and point-
ing out some of the relationships between an illustrative problem (see R problem below) and its solution. Every symbol used in writing the solution sets corresponded to some property of each of the +1 card and the symbol which corresponded to the +1 card property appeared in each solution set. In no case was information given as to how the solution was obtained. The level one and level two material, taken collectively, was referred to as the prerequisite material. The criterion material consisted of level three.

Three types of problems were used. The routine (R) problems consisted of two +all cards and a +1 card, had θ-θ-, [1], as an algorithm pattern (with 8 solution sets), and were used as practice problems for level three (with knowledge of results) and to test for problem-solving ability. The generalization (RG) problems also consisted of two +all cards and a +1 card, but had θ . θ- [1], and θ θ θ . [1], as algorithm patterns (with eight and six solution sets), and were used to measure transfer. The novel (N) problems each consisted of three cards (+all, -all, and +2) with algorithm patterns θ θ θ 0, [2], and θ θ 0 θ, [2], and 6 solution sets. The N problems were used to test for less specific transfer.

The materials were presented in booklet form and consisted of a one-half to one-and-a-half page double-spaced introduction to each level of the material and a number of practice problems, with answers on the following page. A sufficient number of practice problems were used to keep even the fast Ss busy throughout the corresponding practice period. In addition, a number of routine arithmetic and algebraic factoring problems (filler) were constructed. All materials were presented on dittoed sheets to 80 11th grade college preparatory mathematics Ss in the Niagara Falls, New York, school system. No one of these Ss was in an honors section.

**Design**

Two variables were independently manipulated: prerequisite practice (P) and criterion practice (C). P was defined in terms of the amount of prerequisite practice. Half of the 80 Ss were given a total of 11 minutes of prerequisite practice (P), including an opportunity to check their answers; the others were given none (NP). The criterion practice variable refers to the time allotted for attacking the R problems. Half of the Ss, at each prerequisite practice level (NP-C, P-C), were given 11 minutes of R-problem practice (with answers); the others were given none (NP-NC, P-NC). In sum, the 80 Ss were assigned randomly, subject to equal cells, to four treatment groups in a 2 x 2 factorial design.

**Procedure**

The Ss were run in five sections of sixteen (four in each treatment group). The time allotted for each part of the instruction period was as follows: level one material, four minutes; level one practice, four minutes; level two-A (+all criterion) material, four minutes; level two-A practice, four minutes; level two-B (+1 criterion) material, three minutes; level two-B practice, three minutes; level three material, six minutes; and level three practice, 11 minutes. The no practice Ss (whether prerequisite, criterion, or both) were given irrelevant filler problems to work (and check) at the appropriate level(s). Not only were identical materials presented to all groups, but the materials were equally spaced. Provision was made for each S to check his answers to all questions and problems by turning to the next page in his booklet.

Immediately after the instruction and approximately five weeks later, the Ss were given three two-problem tests in the order R-RG-N. Reliability estimates of .88, .68, and .66 respectively, were obtained by computing the correlation coefficients between the correct scores on the two problems in each test. The problems were randomly designated “X” or “Y” so that half of the X(Y) scores were from problem one and half from problem two. The experimenter and four assistants recorded the time taken by each S to complete each test. A maximum of ten minutes was allowed for each. The immediate and retention problems had identical algorithm patterns and card types, but the cards, themselves, differed. All Ss were allowed four minutes to study the -all and +2 criteria just prior to the N testing.

The original instruction-test period took about 85 minutes while the retention testing took about 35 minutes.

**RESULTS**

The number of correct solution set scores varied from zero to 16 (100 percent). A summary of the means and standard deviations is given in Table 1. An analysis of variance revealed that R-performance immediately after learning was an increasing function of prerequisite practice (F = 4.71, df = 1/76, p < .05). The effect due to criterion practice was not significant (F = 2.12, df = 1/76, p > .10). There were 8, 7, 4, and 2 Ss (of 20) in the NP-NC, NP-C, P-NC, and P-C groups, respectively, who had zero correct scores. A significantly greater proportion of those Ss who had had prerequisite practice had non-zero correct scores (χ² = 5.23, df = 1, p < .025). Further comparison of the NP-NC and NP-C score distributions suggested that C practice may have merely increased the proficiency level of those Ss who would have achieved some degree of success without such practice. The generalization (RG) effect due to P practice just did attain an acceptable level of significance (F = 3.97, df = 1/76, p = .05), but there was no C practice effect (F < 1). An identical analysis on the N (correct) scores revealed no
TABLE 1

SUMMARY OF MEANS (STANDARD DEVIATIONS)

<table>
<thead>
<tr>
<th></th>
<th>Correct</th>
<th></th>
<th></th>
<th>Time</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Routine</td>
<td>Generalization</td>
<td>Routine</td>
<td>Novel</td>
<td>Routine</td>
<td>Generalization</td>
</tr>
<tr>
<td>NP-NC</td>
<td>2.7</td>
<td>1.4</td>
<td>2.8</td>
<td>1.2</td>
<td>2.6</td>
<td>.8</td>
</tr>
<tr>
<td></td>
<td>(3.6)</td>
<td>(1.6)</td>
<td>(2.9)</td>
<td>(1.8)</td>
<td>(2.7)</td>
<td>(1.6)</td>
</tr>
<tr>
<td>NP-C</td>
<td>3.2</td>
<td>2.2</td>
<td>2.9</td>
<td>1.6</td>
<td>1.8</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>(4.8)</td>
<td>(4.1)</td>
<td>(2.4)</td>
<td>(2.5)</td>
<td>(1.8)</td>
<td>(2.0)</td>
</tr>
<tr>
<td>P-NC</td>
<td>4.0</td>
<td>2.1</td>
<td>4.2</td>
<td>2.2</td>
<td>1.9</td>
<td>.9</td>
</tr>
<tr>
<td></td>
<td>(4.1)</td>
<td>(2.6)</td>
<td>(3.2)</td>
<td>(2.2)</td>
<td>(2.7)</td>
<td>(1.3)</td>
</tr>
<tr>
<td>P-C</td>
<td>6.5</td>
<td>2.7</td>
<td>4.1</td>
<td>2.4</td>
<td>3.2</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>(5.9)</td>
<td>(4.6)</td>
<td>(3.1)</td>
<td>(3.1)</td>
<td>(3.2)</td>
<td>(1.7)</td>
</tr>
</tbody>
</table>

significant treatment effects. Only the P by C interaction resulted in an $F$ greater than one ($F = 2.83$, $df = 1/76$, $0.05 < p < .10$). The P-C Ss performed somewhat better than the others.

Analyses of variance on the retention (correct) scores indicated that all treatment effects had disappeared after five weeks. None of the $F$-ratios approached significance. Repeated measures analyses, including both immediate and retention scores, were also conducted and are shown in Table 2. In order to simplify these analyses, all frequencies were equalized at 19 by discarding (randomly) the data from one $S$ in the NP-C and P-NC groups.

The repeated measures analyses indicated a high degree of forgetting after five weeks on the R, RG, and N tests ($F = 47.57$, 41.29, and 29.91, $df = 1/72$, $p < .001$). A significant interaction between prerequisite practice and retention on the R scores ($F = 8.88$, $df = 1/72$, $p < .01$) substantiated the results of the separate analyses. Prerequisite practice facilitated R performance immediately after learning, but not after five weeks. Since overall performance on the retention tests was extremely poor and probably reflected near minimal performance, the results of these repeated measures, as well as simple retention, analyses must be viewed with caution. The lack of treatment effects after five weeks may merely have reflected relative insensitivity in the lower scale ranges. Such differential sensitivity could have contributed to the retention by P practice interaction (repeated measures analysis), the lack of a P practice main effect on the retention R scores (retention analysis), and the triple interaction on the N scores (repeated measures analysis).

Fewer solution sets (including incorrect sets) were offered on the retention tests. A $\chi^2$ test indicated that a greater percentage of the Ss offered the maximum number of 16 sets (there were 16 spaces provided on the answer sheet) on the immediate, as opposed to the five-week retention Rtest ($\chi^2 = 13.90$, $df = 1$, $p < .001$). There were 13, 14, 13, and 16 in the NP-NC, NP-C, P-NC, and P-C groups, respectively, who did so on the immediate R test, but only 5, 9, 8, and 6, respectively, on the retention R test. Although the chance generation of correct solution sets was possible, the probabilities involved were too low to have greatly affected the degree of forgetting noted.

The time allowed did not appear to affect the correct scores to any great extent. If the Ss were able to solve the problems, they were able to do so in the ten minutes allowed. Although the time scores did not appear to be systematically affected by prerequisite and/or criterion practice, there was a tendency for the P-C Ss to take less time than the NP-NC Ss on the immediate tests and more time on the five-week tests. The reason for these results was...
### TABLE 2

**ANOVAS ON CORRECT SCORES**

<table>
<thead>
<tr>
<th>Source</th>
<th>R</th>
<th></th>
<th>RG</th>
<th></th>
<th>N</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>MS</td>
<td>F</td>
<td>df</td>
<td>MS</td>
<td>F</td>
</tr>
<tr>
<td>Between prerequisites (P)</td>
<td>1</td>
<td>82.52</td>
<td>3.054</td>
<td>45.32</td>
<td>3.755</td>
<td>2.13</td>
</tr>
<tr>
<td>Between criteria (C)</td>
<td>1</td>
<td>50.95</td>
<td>1.886</td>
<td>.32</td>
<td>&lt;1</td>
<td>5.15</td>
</tr>
<tr>
<td>P X C</td>
<td>1</td>
<td>6.74</td>
<td>&lt;1</td>
<td>.32</td>
<td>&lt;1</td>
<td>11.61</td>
</tr>
<tr>
<td>Pooled between Ss (error)</td>
<td>72</td>
<td>27.02</td>
<td>12.07</td>
<td>8.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between retention (R)</td>
<td>1</td>
<td>172.66</td>
<td>47.565***</td>
<td>106.11</td>
<td>41.288***</td>
<td>65.79</td>
</tr>
<tr>
<td>P X R</td>
<td>1</td>
<td>32.24</td>
<td>8.882**</td>
<td>1.90</td>
<td>&lt;1</td>
<td>2.13</td>
</tr>
<tr>
<td>C X R</td>
<td>1</td>
<td>5.91</td>
<td>1.628</td>
<td>1.50</td>
<td>&lt;1</td>
<td>.95</td>
</tr>
<tr>
<td>P X C X R</td>
<td>1</td>
<td>13.93</td>
<td>3.837</td>
<td>.93</td>
<td>&lt;1</td>
<td>9.50</td>
</tr>
<tr>
<td>Pooled R X Ss (error)</td>
<td>72</td>
<td>3.63</td>
<td>2.57</td>
<td>2.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*significant at .05 level  **significant at .01 level  ***significant at .001 level

---

not clear, but they probably reflect a confounding of unspecified motivation and learning-retention factors.

**DISCUSSION**

Under the conditions of this study, prerequisite practice facilitated routine (R) problem solving. The effect of this variable was somewhat smaller on the generalization (RG) and smaller still on the novel (N) transfer problems. Since the information presented at the criterion level directly related to the R problems, less to the RG problems, and even less to the N problems, these results suggest that the effects of prerequisite practice were largely via criterion learning. Practice in translating card properties into symbols and writing symbol sets associated with single labeled cards could well have made it easier to learn the criterion material since continual reference was made to relationships between cards and symbol sets at the latter level. Increasing prerequisite learning, through practice, could have effectively reduced the number of new concepts with which the learner had to deal at the criterion level. Having well learned the single card-symbol set relationships, primary attention could have been given to the new notions - such as, "finding the collection of all possible sets associated with each of several cards."

In effect, it was felt that these results primarily reflect the relationships between the prerequisite and criterion material. If these materials had been independent of one another, the effects due to prerequisite practice might not have obtained (11). A limitation of this study was that these relationships were not systematically varied, they were merely identified as potentially critical boundary conditions. More definitive studies await greater sophistication in the analysis of subject matters in terms of their behavioral invariants (7, 8, 14, 15).

Since only one order of testing was used, the lack of any clear cut transfer effects can be given an alternate explanation. The practice provided while solving the R (and RG) test problems may have lessened the effects of the treatment variables on RG (and N) performance, particularly that of criterion practice. Although a weak sequence effect of this sort has been observed (14), it did not seem likely that controlling for this factor would have appreciably affected the critical results due to prerequisite practice. The R problems probably would have reflected the group differences most and the N problems least.

The lack of treatment effects after five weeks did not necessarily reflect differential forgetting rates. Both the previously mentioned asymptotic conditions, noted after five weeks, and testing each S twice would have tended to reduce the treatment effects. The Ss were allowed 30 (plus 4 on the N criteria) minutes of problem-solving practice (without knowledge of results) during the original testing period. Practice during the learning phases of the experiment differed by, at most, 22 minutes.

It is important to note that nothing definitive can be said about the relative effects of prerequisite and criterion practice even though the times allowed for
each were identical. The relative effects probably depend on the precise nature of the relationships between the prerequisite and criterion material, the difficulty of the material, exactly what is practiced, the dependent measures used, and when problem solving is assessed. Nonetheless, the present study demonstrates that there are, at least, some conditions under which prerequisite practice has a greater effect on problem-solving performance, immediately after learning, than an equal amount of practice time at the criterion level. Since no attempt was made to specifically favor prerequisite practice and since the card material (12) was constructed so as to reflect certain kinds of mathematical problems, it would appear that teachers would be ill-advised to gloss over preliminary topics so as to spend more time on the main ideas, particularly when the preliminary topics are relatively unfamiliar to the students. This problem is a very real one in many college classrooms, particularly in the technical and scientific areas.

SUMMARY

A criterion task was defined and illustrated. Two variables were considered: practice with the subordinate material (P) and criterion practice (C). The Ss were tested on criterion (R), generalization (RG), and non-specific transfer (N) tasks immediately after learning and after five weeks. P affected R and, to a lesser extent, RG performance on the immediate tests. There were no significant effects due to C.

FOOTNOTES

1. This research was supported by the Cooperative Research Program of the U. S. Office of Education, Project S-097. Thanks are due to Richard DeMunda who was responsible for obtaining the subjects and assisting with both the conduct of the experiment and scoring the materials.

2. Dashes (-) represent symbols used in solution sets and encircled dashes (θ) represent those symbols that correspond to +1 or +2 card properties. The numbers in brackets indicate the number of encircled dashes (i.e. symbols) in each solution set. For more details see the article by Scandura (12).

REFERENCES


