

# Learning set formation and transfer in concept learning: A comparison between attribute identification and rule learning

ADAM MILLER

Oregon System of Higher Education, DCE, Dexter, Oreg. 97431

With a selection paradigm, three experimental groups of 10 Ss each learned six conjunctive concepts successively under three learning conditions: attribute identification (AI), rule learning (RL), or their combination (CL). Only for AI did learning sets form and were strategies near optimum efficiency. Experimental Ss learned two additional disjunctive concepts under the three learning conditions, and three control groups learned only the disjunctive concepts. For a measure of response frequency, positive transfer occurred for RL and CL, indicating the importance of rule learning for transfer. For a measure of strategy efficiency, negative transfer occurred for AI and positive transfer for RL.

Bourne (1966) has proposed that in concept development rule learning can be explained as learning set formation. Applying Harlow's (1959) analysis, a learning set would be demonstrated when learning efficiency for successive concepts gradually increased to an asymptote. In a study relevant to Bourne's proposal, Haygood & Bourne (1965) experimentally separated rule learning (RL), attribute identification (AI), and their combination (CL). Ss learned five successive conjunctive (Cj), disjunctive (Dj), joint denial, or conditional (Cd) concepts either under AI, RL, or CL conditions. Since rule learning was involved in RL and CL conditions, the pattern of errors (or trials to criterion) over the five concepts should have approached learning set characteristics for those two conditions. Only RL data did. Two purposes of the present study were to check Haygood & Bourne's (1965) finding for a selection rather than a reception paradigm and to examine how strategy learning related to learning set formation.

In a related aspect of concept learning, studies (Bourne & Guy, 1968a; Lee, 1968) have demonstrated that positive transfer occurs from simpler to more complex concepts under RL conditions. However, the relative contributions to transfer of rule learning, attribute identification, and the use of strategies have not been reported. The present study investigated these issues.

## SUBJECTS AND MATERIALS

Sixty undergraduate volunteers were randomly by evenly distributed among three experimental and three control conditions.

The materials consisted of instructions, a reference sheet, blank note paper, a penny, a wax pencil, a

learning sheet covered by transparent plastic, and an information card.

The typed instructions indicated the following: S's task was to learn a series of concepts, one at a time. There were three kinds of concepts, which were AND, AND/OR, and IF/THEN. For each kind a definition and examples of positive and negative instances were given. All concepts involved two attributes. For AND (Cj) concepts, both attributes were present in a positive instance. For AND/OR (inclusive Dj) concepts, both attributes or only one were in a positive instance. For IF/THEN (Cd) concepts, a positive instance occurred in one of three ways: both attributes were present; the first attribute was absent, and the second was present; both attributes were absent. Also, in the instructions, S was told the particular procedures for manipulating learning materials and the criterion for learning a concept.

The reference sheet, modeled after Bruner et al (1956), consisted of a 9 by 9 matrix of 81 figures systematically organized by attributes. The figures differed on four attributes, each with three variations. The four attributes and their variations were (1) the color of the rectangular border (red, green, or black) that enclosed the figure; and for the figures (2) their shape (square, circle, or triangle), (3) their number (one, two, or three), and (4) their color (red, green, or black). To the left of each figure was a short horizontal line. The figures, their borders, and the line were visible through the plastic sheet. But below each figure was a square area covered by black wax so the portion of the learning sheet below the wax was not visible. On the learning sheet, in the space concealed by the wax, a + was

marked for positive instances. The penny was for scraping wax below selected figures, and the waxed pencil was to mark on the line the order of selecting a figure (1 for first, 2 for second, etc.).

An information card was presented to S at the beginning of each concept. It specified S's task for the particular concept and learning conditions.

To decide upon the concepts to be learned by Ss, eight pairs of attributes were selected randomly from all attributes. Each of the 10 Ss within a single group was presented a different random order of the eight pairs of attributes which were: three squares, green circle(s), green triangle(s), two red figures, one figure with a red border, three black figures, circle(s) with a black border, and black figure(s) with a green border.

## PROCEDURE

In a transfer design, three experimental groups learned six consecutive Cj concepts followed by two Dj concepts. Three control groups learned only the two Dj concepts. One experimental and one control group were AI, a corresponding pair were RL, and the others were CL. This design does not control for nonspecific warm-up and generalized learning to learn for transfer.

Ss were tested together in a large classroom. The particular materials presented to Ss with the distinctive instructions were the major differentiating factors among the three conditions. At the beginning of experimentation, S was given materials for learning the first concept.

For the first concept, a S in the AI condition was given an information card which stated that the concept was AND and that his task was to identify the two correct attributes. S was started with a positive instance marked "1," which was the same for all learning conditions for that concept and which could have been correct for a Cj, Dj, or Cd concept. S proceeded to select figures, write their order of selection, and scrape wax from beneath the figure to identify it as a positive or negative instance. The criterion of learning was selection of three positive (including the first) and three negative instances and writing the two correct attributes on the information card. On meeting these criteria, learning sheet and information card for Concept 1 were exchanged for those of Concept 2, and the process was repeated. For Concepts 7 and 8, the only change was that Ss were informed that the rule was AND/OR.

For RL the procedure differed from AI in one way: S was told the two attributes in a specific order and was informed that his task was to identify the rule from the three alternatives.

The CL condition differed from AI and RL in that S was given the task of identifying both the rule and the two attributes for each concept. For the three control groups, Ss learned only the seventh and eighth concepts from the 10 random orders used for the experimental groups.

Two measures of concept learning were used. Choices to criterion was the total number of choices, including the initial and criterion choices. A small number represented efficient learning. The redundancy score was based on the number of choices that provided sufficient information to identify the concept (on logical grounds) regardless of whether or not the criterion had been reached. A redundancy choice was one that failed to provide new information for learning the concept. A zero score represented any optimum strategy. Redundancy was derived from Byers' (1967) work.

#### RESULTS AND DISCUSSION

The present study was designed to check some past findings in the areas of successive concept learning and transfer among concepts. The four major concerns and related overall results were as follows: (1) In comparing the relative difficulty of concept learning among learning conditions, fewer choices to criterion were required for RL as compared to AI. (2) In examining learning set formation, contrary to past findings, learning sets appeared to form in the AI but not in the RL condition. (3) In checking the effects of learning conditions upon transfer from Cj to Dj concepts, the occurrence of positive or negative transfer was found to vary both with the learning conditions (AI or RL) and the response measure (choices to criterion or strategies). (4) In comparing the difficulty of Cj and Dj concepts, fewer choices to criterion were required to learn Cj concepts, regardless of the learning condition. Detailed reports and discussions of these results follow.

##### Relative Ease of RL and AI Learning Conditions

Past studies using reception paradigms have found RL easier than AI (Bourne & Guy, 1968b; Haygood & Bourne, 1965). In order to test for differential effects of learning conditions over Cj concepts, choices to criterion were examined in a 3 by 6 repeated measures analysis of variance. Significant sources of variance were: learning conditions ( $F = 6.90$ ,  $df = 2/27$ ,  $p < .01$ ) and concepts ( $F = 2.29$ ,  $df = 5/135$ ,  $p < .01$ ).

The means for learning conditions combined over concepts were AI = 7.3, RL = 6.2, and CL = 7.9. By  $t$ , the RL mean differed at  $p < .01$  from each of the other means. The means for successive concepts

combined over learning conditions were 7.9, 7.1, 7.0, 7.1, 7.1, and 6.7. The only significant  $t$ s were between the first mean and each of the others at  $p < .01$ .

A comparable analysis of variance for the redundancy score yielded as significant sources of variance learning conditions ( $F = 9.62$ ,  $df = 2/27$ ,  $p < .01$ ) and concepts ( $F = 2.26$ ,  $df = 5/135$ ,  $p < .06$ ). The combined means for the learning conditions were AI = 1.3 and RL = 2.3. CL was not used in this analysis because of unreliable scoring. The means for successive concepts were 2.4, 1.4, 1.6, 1.4, 1.2, and 1.6. None of the  $t$ s between pairs of means among learning conditions and between pairs of means among successive concepts was significant at  $p < .05$ .

In a 3 by 2 by 2 repeated measures analysis of variance on Dj learning for choices to criterion, one significant source of variance was learning conditions ( $F = 24.05$ ,  $df = 2/54$ ,  $p < .01$ ). The relevant means were AI = 14.0, RL = 8.8, and CL = 21.0. By  $t$ , each mean differed significantly from each other at  $p < .01$ . In a comparable analysis for redundancy in control groups, the AI mean of 4.9 differed significantly from the RL mean of 1.0 ( $t = 6.88$ ,  $df = 38$ ,  $p < .01$ ).

Thus, for both Cj and Dj concepts, RL took fewer choices to criterion than AI, consistent with past studies. But for Dj concepts, strategies were more efficient for RL than for AI, and for Cj concepts, though the difference was nonsignificant, strategies tended to be more efficient for AI than for RL. The two conditions in which Ss approached an optimum strategy (zero redundancy score) were learning Cj for AI and learning Dj for RL. Past studies reporting differential difficulty for AI and RL have used reception paradigms rather than the selection paradigm used here.

Several factors seem to have contributed to these differences. For RL, Ss had to learn all rules to identify correctly the first concept, while for AI there was a lesser demand in that S had to learn only one rule for the first six concepts and another for the final two concepts. So AI and RL did not represent completely separate processes, since rule learning occurred in both, but in different orders and amounts. It is evident that rules were not completely learned from reading instructions, so RL did involve rule learning, as the condition was named, rather than simply rule identification.

RL was simpler than AI in two ways: S had to select one out of three rules compared to one of six pairs of attributes for AI, and the minimum number of choices possible for RL was

two (not including the initial given instance and the criterion choices) compared to either two or three for AI. It seems unlikely that Ss were more familiar with the logic of eliminating alternative rules compared to testing successive attributes.

Ss appear to have begun at an early point in learning as a result of several factors. They were naive regarding rules and strategies and were not given special experimental training such as the truth table coding of Bourne & Guy (1968b) or the Venn diagram of Haygood & Bourne (1965). The focus of instructions was not on developing strategies, and the selection paradigm was more difficult than the reception paradigm.

In comparing these results to those of other studies, it is important to realize that the criteria for learning have varied among studies and can be assumed to affect results involving the choices to criterion but not the redundancy measure. Generally, the criteria of other studies were more rigorous. For example, Haygood & Bourne (1965) and Bourne & Guy (1968a) used 16 consecutive correct responses as their criteria.

##### Learning Set Formation

Most studies of concept formation that provide data about learning set development have not been directly related to Bourne's (1966) explanation of rule learning as learning set formation (e.g., DiVesta & Walls, 1968). In relevant studies, Haygood & Bourne (1965) found a pattern of change in learning successive Cj concepts under RL, but not under AI or CL, that appeared consistent with learning set formation. And Bourne & Guy (1968a), when data were combined over rules, found a learning set pattern over six concepts for RL. However, neither study used statistical criteria to check for learning sets.

In the present study, the general criteria for a learning set were considered a gradual increase in efficiency over successive concepts to an asymptote. Translated to statistical comparisons, gradual change would occur when adjacent means did not differ significantly, and increase would be shown by a significant increase from the first to the sixth concept. An asymptote would be a nonsignificant difference at least between the fifth and sixth concepts.

To check for learning set formation, the above statistical comparisons on choices to criterion were conducted separately for AI, RL, and CL by  $t$  at  $p < .05$ . For AI, no mean was significantly different from adjacent means, including the fifth and sixth. The first mean of 8.5 differed significantly from the sixth of 6.8. For RL none of the adjacent means

differed significantly, and the first mean of 6.6 did not differ from the sixth of 6.2. For CL no concept differed significantly from another, including the first of 8.6 and the sixth of 7.2. The criteria for learning set formation were met only for AI.

In order to evaluate the relationship of strategy efficiency and learning set formation, the redundancy score was analyzed in the above manner. For AI the only significant difference between means was between the first of 2.2 and the second of 0.7. The first mean did not differ significantly from the sixth of 1.6. For RL none of the pairs of means differed significantly. The first mean was 3.1 and the sixth was 2.4. Scoring of redundancy was too unreliable in CL to justify using it in this analysis.

For RL it appeared that Ss had begun learning at the asymptote, which also seemed to be the case for Haygood & Bourne (1965). But since, for the redundancy score, there was no evidence of learning set formation, and for RL Ss were on the average more than two selections above the optimum strategy, it seems unlikely that an asymptote had been reached. Even after learning six Cj concepts, for both AI and RL, Ss were still in the early stages of mastering strategies. These results were not compatible with either Bourne's (1966) proposal or data from past studies.

Based on data reported to date, learning sets, as defined here, would be expected to form for AI, RL, and CL conditions, but at different rates, depending on factors such as the amount and complexity of rule learning. In the present study, the AI condition favored learning set formation, whereas the RL condition started S near, but not at, the asymptote, so that the possibility of learning set formation was limited.

#### Transfer from Cj to Dj Concepts

Bourne & Guy (1968b) reported that positive transfer occurred from simpler to more complex concepts under RL conditions. In the present study, a second significant source of variance from the analysis of variance on choices to criterion for Dj learning was treatments ( $F = 13.05$ ,  $df = 1/50$ ,  $p < .01$ ). Combined over concepts and learning conditions, the experimental mean was 11.9 and the control mean was 17.1 ( $t = 3.06$ ,  $df = 118$ ,  $p < .01$ ), which indicated overall transfer from Cj to Dj concepts. General warm-up

and learning to learn effects were not controlled for in this experimental design and no doubt contributed to the above results.

To appraise transfer differences separately for each learning condition, the mean for the experimental treatment combined over concepts was compared to the mean for the control treatment. For AI the experimental mean of 12.6 did not differ significantly from the control mean of 15.3 ( $t = 1.17$ ,  $df = 38$ ,  $p > .05$ ). For RL the respective means were 7.1 and 10.4 ( $t = 2.75$ ,  $df = 38$ ,  $p < .01$ ), and for CL the means were 16.1 and 25.8 ( $t = 3.03$ ,  $df = 38$ ,  $p < .01$ ). In a comparable analysis of the redundancy score, for AI the experimental mean of 4.6 did not differ significantly by  $t$  at  $p < .05$  from the control mean of 4.9, and for RL the same was true for the respective means of 1.1 and 1.0. But for RL the experimental Cj mean was 2.3 and the experimental Dj mean was 1.1 ( $t = 3.75$ ,  $df = 48$ ,  $p < .01$ ), while for AI the corresponding means were 1.3 and 4.6 ( $t = 5.41$ ,  $df = 78$ ,  $p < .01$ ).

Thus, in the present study, for choices to criterion positive transfer occurred for the RL and CL but not for the AI conditions. For the redundancy measure of strategy efficiency, there was also evidence of transfer. So both strategies and rule learning contribute to transfer. Differences between AI and RL were that for AI a different logic (new strategy) was required in the shift from Cj to Dj, and negative transfer occurred. For RL the same logic was required for Cj and Dj, and positive transfer resulted.

#### Relative Difficulty of Cj and Dj Concepts

Another finding of past studies was that Cj concepts were easier to learn than Dj concepts (Bourne & Guy, 1968a; Lee, 1968). To check the relative difficulty of learning Cj and Dj concepts for a selection paradigm, rather than the reception paradigm of past studies, the choices to criterion mean for the first concept (Cj) of the experimental groups, combined over learning conditions, was compared to the corresponding mean for the first concept (Dj) for the control groups. The Cj mean of 7.9 was significantly smaller than the Dj mean of 18.0 ( $t = 4.59$ ,  $df = 58$ ,  $p < .01$ ). Parallel comparisons separately for AI, RL, and CL yielded the same significant differences by  $t$  at  $p < .01$ . The

respective means were: AI, 8.5 and 16.8; RL, 6.6 and 11.8; and CL, 8.6 and 25.6. A comparable analysis of the redundancy score indicated that the first Cj mean for the experimental groups combined over learning conditions was 2.4 and the first control Dj mean was 2.9 ( $t = 0.71$ ,  $df = 58$ ,  $p > .05$ ).

In the present study, for choices to criterion Cj was easier than Dj, not only for combined groups but also separately for AI, RL, and CL. But redundancy did not show Cj and Dj differences either for RL or AI. One explanation for the difference between Cj and Dj proposed by Bourne & Guy (1968b) was the familiarity of Ss with truth table coding and its distinctive applications to Cj compared to Dj concept learning. A tabulation of Ss' selections into the four truth table categories of TT, TF, FT, and FF showed that for Cj Ss averaged about 50% TT and 10% FF, while for Dj the comparable percentages were 25 and 30, consistent across learning conditions. Since naive Ss use information from positive instances more effectively than from negative ones, their differential use may in part explain the easier learning of Cj concepts.

#### REFERENCES

- BOURNE, L. E., JR. *Human conceptual behavior*. Boston: Allyn & Bacon, 1966.
- BOURNE, L. E., JR., & GUY, D. E. Learning conceptual rules: I. Some interrule transfer effects. *Journal of Experimental Psychology*, 1968a, 76, 423-429.
- BOURNE, L. E., JR., & GUY, D. E. Learning conceptual rules: II. The role of positive and negative instances. *Journal of Experimental Psychology*, 1968b, 77, 488-494.
- BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
- BYERS, J. L. A note on the calculation of strategies in concept attainment. *American Educational Research Journal*, 1967, 4, 361-366.
- DIVESTA, F. J., & WALLS, R. T. Multiple-versus single-problem training and variations of solution rules in the formation of learning sets. *Journal of Educational Psychology*, 1968, 191-196.
- HARLOW, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 492-537.
- HAYGOOD, R. C., & BOURNE, L. E., JR. Attribute- and rule-learning aspects of conceptual behavior. *Psychological Review*, 1965, 72, 175-195.
- LEE, S. S. Transfer from lower-level to higher-level concept. *Journal of Verbal Learning & Verbal Behavior*, 1968, 7, 930-937.