EFFECT OF IRRELEVANT ATTRIBUTES AND IRRELEVANT OPERATIONS ON RULE LEARNING†

JOSEPH M. SCANDURA and DONALD J. VOORHIES

University of Pennsylvania

The negative effect of irrelevant attributes on concept learning has been well documented. Rule learning, however, involves both attributes and operations. The purpose of this study was to investigate the effects of both kinds of irrelevancies in rule learning. A 3 × 3 factorial design was employed, with the number of irrelevant attributes (0, 1, 2) and the number of irrelevant operations (0, 1, 2) being the factors. Both factors delayed rule learning. The results were compatible with the notion that subjects select from among the available attributes and operations, testing the various combinations in quasi-systematic fashion, until a rule is found which satisfies all instances.

The inhibiting effects of irrelevant attributes in paired-associate learning (e.g. Scandura, 1965; Underwood, 1963) and in concept learning (e.g. Archer, Bourne, & Brown, 1955; Bulgarella & Archer, 1962) are well known. Corresponding data in the area of rule learning, however, are almost entirely lacking. The reasons for this are not hard to find. For one thing, there has been relatively little research on rule learning. For another, there has been little general agreement on exactly what a rule is, and even less on the relationship between rules, on the one hand, and concepts and associations, on the other (e.g., see Gagné, 1965; Haygood & Bourne, 1965; Scandura, 1965, 1968, 1970).

While there is still no general agreement, the formulations of Gagné and Scandura appear to have much in common. For example, Gagné (1966) has viewed rules as a pair of concepts connected by an “action concept,” and Scandura (1970) has charac-

terized rules in terms of a set of stimulus attributes (determining properties) and an operation which acts on these attributes to generate a set of response attributes.

Scandura (1966, 1967, 1969) has argued that when rules are viewed in this way, concepts become simply special cases of rules in which the operation is degenerate. For example, the concept “red square” can be viewed in terms of the rule, “Say ‘yes’ when shown a stimulus which is a red square, and ‘no’ to all the others.” In the case of associations, the operation is further restricted to a simple linking between one stimulus and its response.

In the case of concept learning, then, knowing the relevant attributes is tantamount to knowing the concept. Since the operation is of such a simple form, the question of which “operations” are to be used rarely arises. In rule learning, however, it is important to know both the relevant attributes and also the relevant operations. One might reasonably expect, therefore, that increasing the number of irrelevant attributes and operations would make it more difficult for a subject to determine which are relevant.

Indirect support for this contention was found in an experiment reported by Scandura (1966, 1969). In that study, the number of relevant and irrelevant attributes and operations remained constant, but cues indicating which attributes and/or operations were to be used were varied. Both types of cue had a facilitating effect on rule learning. This suggests that identifying the
relevant attributes and operations in rule learning is not a trivial task where there are irrelevant attributes and operations from which to select.

The primary purpose of this experiment was to test the dual hypothesis that both irrelevant attributes and irrelevant operations would delay rule learning. A secondary purpose was to obtain information on how subjects select attributes and operations on meaningful tasks. In particular, most theorists (Levine, 1966; Restle, 1962) have assumed that subjects select attributes on a random basis. This assumption seems reasonable in the experiments cited since the tasks were artificial and constructed explicitly so that each attribute appeared to the subject to be an equally likely candidate. It is not clear, however, that this assumption is realistic with meaningful materials with which the subjects have varying degrees of familiarity. In the Scandura (1966, 1969) study cited above, for example, the attributes and operations selected on learning trials appeared to be a function of the particular instances involved.

**Method**

**Materials**

The materials were similar to those used in an experiment reported by Scandura (1966, 1969). The stimuli were strings of three, four, and five single-digit natural numbers (1, 2, 3, 4, ..., 8, 9). These stimuli were called 3-, 4-, and 5-tuples, respectively. The 3-tuples were of the form (A, B, C); 4-tuples, of the form (A, B, C, D); and 5-tuples, of the form (A, B, C, D, E), where in each case A, B, C, D, and E represent arbitrary natural numbers. Sample 3-, 4-, and 5-tuples, respectively, are (3, 2, 7), (5, 1, 4, 8), and (9, 4, 5, 1, 3).

A single natural number response was associated with each stimulus n-tuple according to a rule. Each rule involved the numbers in three positions of the associated stimulus n-tuples and two of the four arithmetic operations (+, −, ×, ÷). In accordance with Scandura's (1970) characterization, then, each rule can be characterized in terms of a set of triples of numbers (the set of determining properties—attributes) and a composite of two arithmetic operations. For example, the set of determining properties might consist of those triples of numbers (of, say, stimulus 5-tuples) appearing in the first, third, and fourth positions. The composite operation, in turn, might involve addition (applied, say, to the numbers in the first and fourth positions) followed by multiplication (applied to the obtained sum and the number in the third position). For example, application of this rule to the stimulus 5-tuple (9, 4, 5, 1, 3) yields (9 + 1)5 = 50 as the response.

A problem consisted of a sequence of 11 stimulus-response pairs all of a given type (e.g., all 4-tuples). The stimulus-response pairs were constructed (as indicated below) so that the responses could be generated from the associated stimuli by exactly one rule. Since each rule involved the numbers in just three positions, restricting the stimulus n-tuples to a given type had the effect of specifying a number of irrelevant attributes (IA1, IA2, IA3). IA1 refers to 3-tuples, where none of the numbers (positions) were irrelevant. IA2 refers to 4-tuples, where the number in one position was irrelevant. IA3 refers to 5-tuples, where the numbers in two positions were irrelevant. The number of arithmetic operations which could be used on a given problem was also restricted, this time by specifying a subset of the four operations from which the selections could be made. Where the selection set consisted of just two operations, there were no irrelevant operations (IO2). Where the selection set consisted of three operations, one was irrelevant (IO3); and where all four arithmetic operations were in the selection set, there were two irrelevant operations (IO4).

Six rules were used as a basis for constructing the experimental problems (see Table 1). These rules were randomly selected subject to the following constraints (which helped to insure a broad sample of problems). (a) If (X * Y) % Z was used, (X % Y) * Z was not used (where * and % denote arbitrary operations). (b) (X + Y) − Z and (X × Y) ÷ Z were included. (c) (X + Y) × Z and (X − Y) × Z were not both used. (d) (X + Y) ÷ Z and (X − Y) ÷ Z were not both used. (e) The operations on the left were always performed first (e.g., (X * Y) % Z was used but not X * (Y % Z)). (X, Y, and Z refer to arbitrary positions in n-tuples. In one case, X might refer to the second position, Y to the third position, and Z to the first.)

The stimuli for the 5-tuple problems were constructed first. For each of the six operational rules, 11 stimuli were constructed so that the natural numbers (between 1 and 9) assigned to the relevant positions in each n-tuple gave natural num-

**TABLE 1**

**Rules Used to Construct Experimental Problems**

<table>
<thead>
<tr>
<th>3-tuple rules</th>
<th>4-tuple rules</th>
<th>5-tuple rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>[(A × C) + B]</td>
<td>[(A × C) + B]</td>
<td>[(A × C) + B]</td>
</tr>
<tr>
<td>[(C − A) + B]</td>
<td>[(D − A) + B]</td>
<td>[(D − A) + B]</td>
</tr>
</tbody>
</table>
ber responses. The relevant positions and the order of operations at these positions were determined randomly, as were the numbers in the irrelevant positions. The stimuli for the 4- and 3-tuple problems were obtained by suppressing the number(s) furthest to the right in the corresponding 5-tuples.

Each $n$-tuple was printed on the front of a 3 × 5 card; the same $n$-tuple, with the corresponding number response, appeared on the back. The selection set of operations associated with each problem was printed on a separate 3 × 5 card.

**Design and Subjects**

A 3 × 3 factorial design was used, the factors being the number of irrelevant attributes (IA, IA, IAA) and the number of irrelevant operations (IO, IO, IO).

The subjects were 45 apparently motivated volunteer junior and senior students from Abington High School, South Campus, Abington, Pennsylvania. They were randomly assigned to the nine cells, so that five subjects were assigned to each cell.

Each subject received three problems. To help insure comparability of the nine experimental cells, the same pattern of assigning problems to subjects was used in each cell. This pattern of problems was randomly determined subject to two constraints: (a) No subject received the same problem twice. (There were only six problems from which to select.) (b) No problem appeared more than three times. The actual assignment of problems used is given in Table 2. (The $n$-tuples for the selected problems were always given in the same sequence; the starting point of that sequence was varied when the same problem was given more than once in any one cell. Across cells, the same sequence and starting points were used.)

**Procedure**

The subjects were tested individually during 45-minute study hall periods. Each subject was shown a sample $n$-tuple and selection set of operations, and was instructed to discover a rule (of the type described above) which would make it possible to correctly anticipate the responses. Concurrently, the subjects were provided with training in how to describe rules in words. The subject was told that there might be more than one way to generate the response corresponding to any given $n$-tuple, but that only one rule would generate the responses for all 11 $n$-tuples in a problem. (In fact, the chances were practically nil that more than one rule would work with as many as three successive stimulus-response pairs. In 100% of the problems, there was exactly one rule which could be used to generate the correct responses for the first three stimulus-response pairs. Further, only one rule could generate the correct responses for the first two stimulus-response pairs in 92% of the problems.)

To familiarize the subject with the procedure, he was presented with a practice problem prior to the three experimental problems. The following procedure was used throughout. The subject was shown the first $n$-tuple (stimulus) response pair, together with the second stimulus $n$-tuple, and was asked to give a response to the second $n$-tuple. As soon as the subject responded, the second card was turned over revealing the same $n$-tuple with its associated response. If the subject's response was incorrect, the examiner presented a new stimulus $n$-tuple, and asked for a response. If the subject's response was correct, the examiner asked what rule the subject was using. If the subject identified the correct rule, the examiner recorded the total time the subject spent on the problem. In computing the total time, the examiner ignored the time it took the subject to verbalize the rule after he gave correct responses. If not, the examiner told the subject that his rule was incorrect, and presented him with another stimulus $n$-tuple. This process was continued until the subject discovered the correct rule or 8 minutes had elapsed, whichever occurred first.

The subject was given the option of turning over a card to see the correct response at any time he wished. When this occurred, the examiner gave the subject a new stimulus $n$-tuple. If the subject spent 2 minutes with a stimulus $n$-tuple, the examiner urged him to move on to a new $n$-tuple. All previous stimulus-response pairs were visible to the subject throughout the problem.

The subject's score was the average time on the three problems he worked. After testing was completed, some of the subjects were presented with one or two additional problems and were asked to explain how they attacked the problems.

**Results and Discussion**

As hypothesized, both irrelevant attributes and irrelevant operations delayed rule learning ($p < .001$). The group means are presented in Table 3. In spite of the small number of subjects (five in each cell), the pattern of results was quite regular with the exception of cell IO, IA, IA. This suggests that the variables considered were basically quite powerful.

---

**Table 2**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Problem 1</th>
<th>Problem 2</th>
<th>Problem 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
</tr>
<tr>
<td>2</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
</tr>
<tr>
<td>3</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
</tr>
<tr>
<td>4</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
</tr>
<tr>
<td>5</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
<td>(x+y+z)</td>
</tr>
</tbody>
</table>
Presumably, the subjects identified the appropriate rule on each problem by selecting and testing various combinations of attributes and operations until the correct rule was found. If the selection process was random, as many theorists have assumed (e.g., Levine, 1966; Restle, 1962), then the probability of selecting a correct rule can be determined for each cell of the design.

Assuming that the time required to sample (with replacement) and test each combination of attributes and operations is constant, the data of Table 3 should vary inversely with these probabilities. Equivalently, these data should vary directly with the numbers reported in Table 4 which were obtained by computing the number of instances required to attain a probability of .5 (an arbitrary figure) of selecting a correct rule and multiplying by a constant.

Comparison of Tables 3 and 4 indicates that while the general trends are in the same direction, there was an important difference. In Table 3, the difference between columns (rows) IA₁ (IO₁) and IA₀ (IO₀) was substantially greater than the corresponding difference between columns (rows) IA₂ (IO₂) and IA₁ (IO₁). In Table 4, precisely the opposite is true. In effect, the subjects did much better on the IA₂ and IO₂ problems than would be expected on the basis of random selection.

This observation suggests that to some extent, at least, the subjects tended to make their selections in a systematic fashion. Indirect evidence for this contention was obtained from the postexperimental questioning. Most of the subjects indicated that various characteristics of the stimuli and responses caused them to either eliminate or single out for special attention certain attributes and operations. For example, whenever a large response (number) was associated with a stimulus, many of the subjects tended to try multiplication first, and possibly addition second. Similarly, small responses resulted in the use of division, or sometimes subtraction.

It must not be thought from the results of this experiment that the effects of irrelevant attributes and irrelevant operations are uniformly negative. While not dealt with in this study, it is quite possible that the introduction of such irrelevant information may cause subjects to search for new and better ways of discovering rules. To the extent that they succeed (cf. Roughhead & Scandura, 1968), irrelevant information may enhance transfer to new problems. This is an important possibility which should be checked in further research.

### REFERENCES


**TABLE 4**

<table>
<thead>
<tr>
<th>Irrelevant operations</th>
<th>IA₀</th>
<th>IA₁</th>
<th>IA₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>IO₀</td>
<td>1</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>IO₁</td>
<td>3</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>IO₂</td>
<td>6</td>
<td>24</td>
<td>60</td>
</tr>
</tbody>
</table>

Note.—The numbers in this table are meaningful only in relation to one another.

(Received May 24, 1970)