2 questions were asked: (a) can “what is learned” in mathematical discovery be identified and taught by exposition with equivalent results, (b) how does “what is learned” depend on prior learning and on the nature of discovery? The major hypothesis was that discovery Ss may discover derivation rules for deriving classes of solutions but only when the solutions are not initially known. 4 programs, (specific) rule-given (R), discovery (D), guided discovery (G), and exposition of derivation rule (E) were administered to 7 groups. 1 group received program R alone; the others received R with 1 of the other programs. Both orders of presentation were represented: RD, DR; RG, GR; RE, ER. All Ss were required to derive new solutions within the scope of the derivation rule. As hypothesized, Groups R and RD performed at 1 level which was reliably (p < .001) below the common level of the other 5 groups. Theoretical and practical implications were discussed.

One of the fundamental assumptions underlying many of the new mathematics curricula is that discovery methods of teaching and learning increase the students' ability to learn new content (e.g., Beberman, 1958; Davis, 1960; Peak, 1963). The last decade of research on discovery learning, however, has produced only partial and tentative support for this contention. Even where the experiments have been relatively free of methodological defects, the results have often been inconsistent (e.g., see Ausubel, 1961; Kersh & Wittrock, 1962). More particularly, the interpretation of research on discovery learning has been made difficult by differences in terminology, the tendency to compare identical groups on a variety of dependent measures, and vagueness as to what is being taught and discovered.

While most discrepancies due to differences in terminology can be reconciled by a careful analysis of what was actually done in the experiments (e.g., Kersh & Wittrock, 1962) and thus present a relatively minor problem, the failure to equate original learning has often made it difficult to interpret transfer (and retention) results in an unambiguous manner. Thus, several studies (e.g., Craig, 1956; Wittrock, 1963) have shown that rule-given groups perform better on “near” transfer tests than do discovery groups. The obtained differences, however, may have been due to the fact that the discovery groups did not learn the originally presented materials as well as the rule-given groups.

When the degree of original learning was equated, Gagné and Brown (1961) found that their discovery groups were better able to derive new formulas than were their rule- (i.e., formula) given groups. They attributed this result to differences in “what was learned” but added that they were unable to specify precisely what these differences were. On the basis of an analysis of the experimental programs used by Gagné and Brown (1961), Eldredge (1965) hypothesized that the differences found by
Gagné and Brown (1961) were due to uncontrolled factors. Eldredge conjectured that if the treatment differences were limited to the order of presentation of the discovery hints and the to-be-learned formulas, no differences in transfer ability would result. However, Eldredge's results contradicted his hypothesis. In subsequent studies, Gutherie (1967) and Worthen (1967) obtained similar sequence effects.

Using the Set-Function Language (SFL) characterization of a rule as a guide, Scandura (1966) proposed an analysis of discovery learning that seems to be in accord with experimental findings.

In the SFL, the rule is viewed as the basic unit of behavior; associations and concepts are shown to be special cases (of the rule). The denotation of a rule is defined as a set of functionally distinct stimulus-response pairs—the instances of the rule. The rule construct itself is characterized as an ordered triple (D, O, R) where D refers to the set of those stimulus properties which determine the corresponding responses, and O refers to the operation or transformation by which the derived stimulus properties or (internal) responses in the set R are derived from the properties in D (for more details, see Scandura, 1966, 1967a, 1968a, b, and c).

The main point of the analysis was that in order to succeed, discovery Ss must learn to derive solutions (i.e., responses) whereas solution-given Ss need not. In attaining criterion, discovery Ss may discover a derivation rule by which solutions to new, though related, problems may be derived. Under these circumstances, discovery Ss would be expected to perform better than expository Ss on tasks which are within the scope of such a derivation rule. If the new problems presented have solutions beyond the scope of a discovered derivation rule, however, there would be no reason to expect discovery Ss to have any special advantage.

This study was concerned with two major questions. First, can “what is learned” in mathematical discovery be identified and, if so, can it be taught by exposition with equivalent results? Second, how does “what is learned” depend on prior learning and on the nature of the discovery treatment itself?

The SFL was used as an aid in analyzing the guided discovery programs used by Gagné and Brown (1961) and Eldredge (1965) to determine “what was learned.” As a result of this analysis, an expository statement of the derivation rule was devised. It was possible, in the manner described by Scandura, Woodward, and Lee (1967), to determine on an a priori basis which kinds of transfer item could be solved by using this derivation rule and which could not.

Assuming that transfer depends only on whether or not the derivation rule is learned, then the order in which the formulas (i.e., the solutions) and the derivation rule are presented should have no effect on transfer so long as S actually learns the derivation rule. If, on the other hand, a discovery program simply provides an opportunity to discover and does not guide the learner through the derivation procedure, sequence of presentation might have a large effect on transfer. That is, if a capable and motivated S is given appropriate hints, he might well succeed in discovering the appropriate formulas and in the process discover the derivation rule. It is not likely, however, that he would exert much effort when given an opportunity to discover a formula he already knows. Something analogous may well have been involved in the studies by Eldredge (1965), Gutherie (1967), and Worthen (1967).

In particular, the following hypotheses were made. First, what was learned by guided discovery in the Gagné and Brown (1961) study can be presented by exposition with equivalent results. Second, presentation order is critical when the hints provided during discovery are specific to the respective formulas sought rather than relevant to a general strategy (i.e., derivation rule). Third, presentation order is not critical when the program effectively forces S to learn the derivation rule, regardless of whether the learning takes place by exposition or by discovery.

**Method**

**Materials**

There were seven treatments. Each consisted of a common introductory program followed by

---

2 Copies of the experimental materials used are included in Roughead's (1966) dissertation and in Scandura's (1967) final report.
The three-columed table would look like.

```
| Term number n: | 1 2 3 4 5 ... |
| Term value T_n: | 2 6 10 14 ... |
| Sum \( \sum^n \): | 2 8 18 32 ... |
```

The rule and example (R) program consisted of the three series displays together with the respective summing formulas. The presentation of each summing formula was followed by three application problems—e.g., find the sum of 2 + 6 + 10 (= 2\(3^{nd}\) = 18). The S was also required to write out each formula in both words and symbols, but no rationale for the formula was provided. A test of the three training formulas was included at the end of the R program.

The other three basic programs included differing kinds of directions and/or hints as to how the summing formulas might be determined. The expository (E) and (highly) guided discovery (G) programs were based on a simplified variant of that derivation rule presumably learned by the guided discovery Ss in the Gagné and Brown (1961) study. The identified rule can be stated,

\[ f(n) = \text{a formula for } \sum^n \text{ may be written as the product of an expression involving } n \text{ [i.e., } f(n)] \text{ and } n \text{ itself. The required expression in } n \text{ can be obtained by constructing a three columned table showing: (1) the first few sums } \sum^n, \text{ (2) the corresponding values of } n, \text{ and (3) a column of numbers } f(n) = \sum^n/n \text{ which when multiplied by } n \text{ yields the corresponding values of } \sum^n. \]

Next, determine the expression \( f(n) = \sum^n/n \) by comparing the numbers in the columns labeled \( n \) and \( \sum^n/n \) and uncovering the (linear) relationship between them. The required formula is simply \( \sum^n = n \cdot f(n) \).

As an example, consider the display,

\[
\begin{array}{ccc}
| n | 1 & 2 & 3 & 4 & 5 \ldots \\
| t(n) | 2 & 4 & 6 & 8 & \ldots \\
| \sum^n | 2 & 8 & 18 & 32 & \ldots \\
\end{array}
\]

The three-columned table would look like,

<table>
<thead>
<tr>
<th>( f(n) )</th>
<th>( n )</th>
<th>( \sum^n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>\ldots</td>
<td>\ldots</td>
<td>\ldots</td>
</tr>
<tr>
<td>2n</td>
<td>n</td>
<td>2n^2</td>
</tr>
</tbody>
</table>

The emerging pattern is \( f(n) = 2n \); so, \( \sum^n = 2n \cdot n^2 \).

The E program consisted of a simplified statement of the derivation rule as it applied to each of the three training series. To insure that S learned how to use the derivation rule, a vanishing procedure was used which ultimately required S to apply the procedure without any instructions. The G program paralleled the E program in all respects. The only difference was that the G program consisted of questions whereas the E program consisted of yoked direct statements, each followed by a parallel question or completion statement to see whether S had read the original statement correctly. For example, the E statements, "When \( n = 3 \), you can multiply 6 times \( n \) to get \( \sum^n = 18 \). What times \( n \) gives \( \sum^n = 18 ? \)" corresponded to the question, "When \( n = 3 \), what times \( n \) gives \( \sum^n = 18 ? \)" which appeared in the G program. Since the degree of overt responding was held constant, the only difference between the E and G programs was whether the information was acquired by reception or by reacting to a question (i.e., by discovery). The discovery (D) program, on the other hand, simply provided S with an opportunity to discover the respective summing formulas. The S was guided by questions and hints which were specific to the formulas involved (e.g., "the formula has a 2 in it") rather than relevant to any general strategy or derivation rule. The questions and hints were interspersed with liberal amounts of encouragement (e.g., "Good try," "you can do it," etc.) to provide motivation.

There were two transfer tests. The within-scope transfer test consisted of two new series displays which could be solved by the identified derivation rule. These series and their respective summing formulas were \( 3 + 5 + 7 + \ldots + (2n + 1) \rightarrow (n + 2)\cdot n \) and \( 4 + 10 + 16 + \ldots + (6n - 2) \rightarrow (3n + 1)\cdot n \). The extra-scope transfer test involved the series, \( 2 + 4 + 8 + \ldots + 2^n \rightarrow (2T_{n+1} - 2) = (T_{n+1} - 2) + 1/2 + 1/6 + 1/12 + \ldots + 1/(n(n + 1)) \rightarrow n/(n + 1) = n^2T_n \), which, strictly speaking, were beyond the scope of the identified derivation rule. A series of hints, paralleling those used in the D program, were constructed to accompany each test series.

The introductory and treatment programs were mimeographed and stapled together into separate 5\( \frac{1}{2} \times 8\frac{1}{2} \) inch booklets. The four transfer series were presented on separate pages in a test booklet in the same three-row form used in the learning programs. The hints were put on 5\( \times 7 \) inch cards, bound by metal rings.

**Subjects, Design, and Procedure**

The naive Ss were 105 (103 females) junior and senior elementary education majors, enrolled in required mathematics education courses at the Florida State University, who volunteered to participate in the experiment. The data of seven other Ss were discarded because they failed to meet a major premise on which the hypotheses were based. That is, they were poorly motivated and/or made
a large number of errors on the treatment programs.

The experimental Ss were randomly assigned to the seven treatment groups. In addition to the common introductory program, the R group received only the R program. The other six treatment groups received the R program together with one of the other three basic instructional programs. The RE, RG, and RD groups received the R program followed by the E, G, and D programs, respectively, while the ER, GR, and DR groups received these same respective programs in the reverse order.

The Ss were scheduled to come to the experimental room in groups of four or less and were arranged at the ends of two tables which were partitioned to provide separate study carrels. A brief quiz was used to screen out any Ss who were already familiar with number series and/or formulas for summing them. Then, they were told:

This is an experiment in learning mathematics. You will be given two programmed booklets to study. You are expected to try to learn. You should work at a good pace, but read everything for understanding. . . . If you have an error, don’t change your answer, but write the correct answer under your original answer. If you can not respond to a question within a minute or so, put an “X” in the blank and continue. You should, of course, look back at the question after finding the answer to be sure you understand. . . .

The Ss worked at their own rate. The two Es observed the progress closely, provided general assistance and encouragement where needed, and recorded the times taken on the introductory and treatment booklets.

As soon as all of the Ss in the testing group had completed the treatment programs, they were told to review for a test. After 2 minutes, the booklets were collected and the tests and hint cards were presented. The Ss were instructed:

On this test you will be timed. You also will be provided with hints to aid you when necessary. The less time it takes you and the fewer hints you need on a given problem, the better your score. You will be asked to find the formula for each problem on this test. On each problem, you will have 5 minutes to find the correct summing formula. You should show any necessary work in your booklet. When you get an answer, raise your right hand immediately. Like this! Try it! . . . I’ll tell you whether you are correct or incorrect. If incorrect, continue searching for the answer. Be sure to show me your answer quickly so that you get the best possible time score. . . . When I tell you that the 5 minutes are up, if you have not found the formula, you may begin using hints. You may use as many of the hints as you wish, and when you wish, after the 5 minute period. But remember, the fewer hints you use, the better your score.

Before continuing on to the second problem, each S read all of the hint cards pertaining to the first problem. The four Ss in each testing group began each problem at the same time. If an S solved a problem before the others, he was allowed to read the rest of the hints for that problem and, then, was required to wait for the others to finish. Before being released, Ss were asked not to discuss particulars of the experiment with others who might participate.

Three indexes of performance on the transfer tasks were obtained: (a) time to solution, (b) number of hints prior to solution, and (c) a weighted score similar to that used by Gagné and Brown (1961). The weighted score was equal to the time to solution in minutes plus a penalty of 4, 7, 9, or 10 depending on whether S used 1, 2, 3, or 4 hints, respectively. Theoretically, a range of scores from 0 to 20 was possible on this measure. Standard analysis of variance procedures were used to analyze the data after Cochran’s C test failed to detect heterogeneity of variance.

RESULTS

Treatment Programs

All treatment groups performed at essentially the same level on the introductory program, both in terms of time to completion (F = 1.74, df = 6/98, p > .05) and number of errors (F = 1.35, df = 6/98, p > .05). Since the number of frames varied among the treatment programs, no overall comparisons were warranted.

Performance on Learning and Transfer Tests

The results on the within-scope transfer test conformed to prediction. Irrespective of the transfer measure used, the group (R) given the formula program only and the group (RD) given the formula program followed by the opportunity-to-discover program performed at one level (F < 1, df = 1/28) while the other five groups performed at a common (F < 1, df = 4/70) and significantly higher level (Ftime = 32.66, df = 1/98, p < .001; Fhints = 54.52, df = 1/98, p < .001; Fweighted = 57.99, df = 1/98, p < .001). In particular, only that sequence effect involving Groups RD and DR was significant (p < .01).

While there were no overall treatment differences on the extra-scope transfer test (maximum F = 1.31, df = 6/98, p > .05),
The contrast between Groups R and RD and Groups DR, RG, GR, RE, and ER attained a borderline significance level ($F_{Time} = 3.66, df = 1/98, .05 < p < .10; F_{Hints} = 4.02, df = 1/98, p < .05; F_{Weighted} = 4.61, df = 1/98, p < .05$)\(^3\). There were, however, no reliable performance differences between Groups DR and RD ($F < 1$).

\(^3\)In a study on rule generality, Scandura, Woodward, and Lee (1967) obtained a similar extra-scope transfer effect. While no extra-scope transfer was almost universally the case, one of the rules (i.e., $50 \times 50$) introduced was apparently generalized ($to \ n \times \ n$) and thereby provided an adequate basis for solving an extra-scope problem. A recent study by Scandura and Durbin (1968) has demonstrated that, indeed, the form of a rule statement is an important determiner of generalization.

While not sufficient as presented, the derivation rule statement, introduced in this study, could also be generalized. In particular, the first hint available on the third transfer problem provided a basis for making appropriate modifications. Similarly, although Problem 4 involved fractional term values, the summing formula could be obtained by a relatively simple extension of the derivation rule statement.

For these reasons and because the results on the extra-scope test were subject to possible transfer effects of testing on the within-scope test, caution is advised in interpreting the extra-scope results. The extra-scope test was originally included to obtain experimental hypotheses and not definitive information. These comments, however, in no way apply to the clear results on the within-scope test. These transfer effects cannot be attributed to differences in original learning. A learning test embedded within the common R program, indicated that Ss had well learned the appropriate summing formulas to the three training series before they took the transfer tests. The group means ranged from 5.5 to 6 with a possible maximum of 6 and, minimum of 0. The error rates on the treatment programs were similarly low with an average of between one and two errors per program.

**Discussion and Implications**

Two points need to be emphasized. First, "what is learned" during guided discovery can at least sometimes be identified and taught by exposition—with equivalent results. While this conclusion may appear somewhat surprising at first glance, further reflection indicates that we have always known it to be at least partially true. As has been documented in the laboratory (e.g., Kersh, 1958) as well as by innumerable classroom teachers of mathematics, it is equally as possible to give Ss rules for deriving answers as it is to have them derive (i.e., discover) the answers themselves. No one to our knowledge, however, had ever seriously considered identifying "what is learned" in
deriving rules (i.e., formulas) in addition to the rules themselves. In the present study, the authors were apparently successful in identifying a (derivation) rule for deriving a class of more specific rules. No differences in the ability to derive new (within-scope) formulas could be detected between those Ss who discovered a derivation rule and those who were explicitly given one. What was not done in this study was to consider the possibility that the discovery Ss may have acquired a still higher order ability—namely, an ability to derive derivation rules. In any case, there are undoubtedly a large number of situations where, because of the complexity of the situation, “what is learned” by discovery may be difficult, if not impossible, to identify. In these situations, there may be no real alternative to learning by discovery.

Nonetheless, the value to transfer ability of learning by discovery does not appear to exceed the value of learning by some forms of exposition. Before definitive predictions can be made, careful consideration must be given to “what is learned,” the nature of the transfer items, and the relationships between them. As we identify just what it is that is learned by discovery in a greater variety of situations, we shall be in an increasingly better position to impart that same knowledge by exposition.

The second point to be emphasized concerns the sequence effect—if a person already knows the desired responses, then he is not likely to discover another rule by which such responses may be derived, even if he has all of the prerequisites and is given an opportunity to do so. The reverse order of presentation may enhance discovery without making it more difficult to learn more specific rules at a later time. In short, prior knowledge may actually interfere in a very substantial way with later opportunities for discovery. Nonetheless, there may be some advantages inherent in learning more specific rules. Although data are practically nonexistent on this point, it is quite possible that specific rules may result in shorter latencies. Why and how sequence affects “what is learned” is still open to speculation (e.g., Guthrie, 1967; Yonge, 1966). Our interpretation is as follows: When S is presented with one or more stimuli and is required to produce responses (e.g., formulas or specific rules) he does not already know, he necessarily must first turn his attention to deriving a rule (or derivation rule) by which he can generate the appropriate responses. In the process, S may discover a derivation rule, which is adequate for deriving other responses in addition to the ones needed. The kind and amount of guidance given would presumably help to determine the precise nature of the derivation rule so acquired. On the other hand, if S already knows the responses (i.e., has previously mastered more specific rules or “associations” by which the responses can be derived), it is not likely that he will waste much time trying to find another way to derive them. Under these conditions, it would seem that the only way to get S to learn a more general rule would be to change the context. Presumably, the expository and guided discovery Ss in this study learned the derivation rule because this appeared to be the desirable thing to do. The authors believe that any theory based on the rule construct will have to invoke some such mechanism to account for sequence effects (e.g., see Scandura, 1968 a, b, and c).

The obtained sequencing result may also have important practical implications, as will be attested to by any junior high school mathematics teacher who has attempted to teach the “meaning” underlying the various computational algorithms after the children have already learned to compute. The children must effectively say to themselves something like, “I already know how to get the answer. Why should I care why the procedure works?” This is not to say that meaning should be taught first simply out of some sort of dislike for rote learning—for certain purposes rote learning may be quite adequate and the most efficient procedure to follow. The important point is that learning such things as how to multiply, without knowing what
multiplication means, may actually make it more difficult to learn the underlying meaning later on.

REFERENCES


Kersh, B. Y. The adequacy of “meaning” as an explanation for the superiority of learning by independent discovery. Journal of Educational Psychology, 1958, 49, 282-292.


Scandura, J. M. Using the rule to formulate research on meaningful learning: II. Empirical research. Acta Psychologica, 1968, in press. (b)

Scandura, J. M. Using the rule to formulate research on meaningful learning: III. Analyses and theoretical direction. Acta Psychologica, 1968, in press. (c)


(Received July 28, 1967)