A THEORY OF MATHEMATICAL KNOWLEDGE: CAN RULES ACCOUNT FOR CREATIVE BEHAVIOR?

JOSEPH M. SCANDURA
Associate Professor of Education
University of Pennsylvania
Philadelphia, Pennsylvania

Mathematics is perhaps the most highly organized body of knowledge known to man. Yet, in spite of its clarity of structure, most of the research done on mathematics learning and behavior has been strictly empirical in nature. To be sure, there has been a fair amount of research in the area and the amount seems to be growing rapidly, but there has been no superstructure, no framework within which to view mathematical knowledge and mathematical behavior in a psychologically meaningful way.

A number of psychologists feel that the mechanisms involved in language, mathematical, and other subject-matter behavior may be accounted for within the confines of S-R mediation theory. This may be possible in principle (e.g., see Millenson, 1967; Suppes, 1969a), but the networks of S-R associations required to do the job would almost certainly be so complex as to provide little intuitive guidance in formulating research on complex mathematical learning. For arguments pro and con, see Arbib (1969), Scandura (1968, 1970b, 1970d), and Suppes (1969b).

As a way around these problems, linguists, like Chomsky (1957, 1965), have introduced rules and other generative mechanisms to account for (idealized) language behavior. Although many details still need to be worked out, most generally agree that some sort of analysis in terms of rules will prove adequate to account for most language behavior.

During the past few years, the author has been attempting to develop a similar approach to mathematics learning (Scandura, 1966, 1967a, 1968, 1969b). No comprehensive scheme for classifying mathematical

1. This paper is based on a talk given at a Symposium on Structural Learning in the Subject Matter Disciplines at the AERA Convention in Los Angeles, Thursday, February 6, 1969. The author would like to thank Leon Henkin, Henry Hiz, Dag Prawitz, Marshall Stone, and Patrick Suppes for their helpful comments on an earlier draft and John Durnin for his general assistance in the preparation of this paper.

2. Nonetheless, some linguists (e.g., Hiz, 1967) do not feel that all important aspects of language can be dealt with in this way.

May 1971
behaviors has been proposed, however, and most (but not all) of the experimental research has been based on relatively simple mathematical tasks (cf. Scandura, 1969b). The basic supposition has been that an understanding of what is involved in such tasks will provide a better position for explaining more complex mathematical learning. While there has been increasing support for this contention among behavioral scientists (e.g., Bartlett, 1958; Gagné, 1965; Miller, Galanter, & Pribram, 1960), some mathematics educators have been skeptical. Presumably, the position is that any interpretation of complex mathematical learning in terms of simple rules will surely be inadequate.

In reaction, the author proposes and defends the rather strong thesis that rules are the basic building block of all mathematical knowledge and that, if looked at in the right way, all mathematical behavior is rule-governed. More specifically, it is proposed that the mathematical behavior any given individual is potentially capable of, under ideal conditions of performance, can be accounted for precisely in terms of a finite set of rules.

This statement is clearly meant to imply more than just a post hoc account of a given finite corpus of behaviors. If limited to this, the claim would be trivially true since any given subject during his lifetime is necessarily limited to a finite number of behaviors. (A finite number of behaviors can obviously be generated by a finite number of rules.)

Furthermore, this is not a thesis to be proved since it is basically empirical in nature. The problem is that there is no operational way of determining the behavior potential of a subject independently of the rules used to characterize his knowledge. Unfortunately, it would be extremely difficult and time-consuming to obtain an adequate sample of mathematical behaviors to work with under the ideal conditions envisioned—that is, where the subject is unencumbered by memory or his limited capacity to process information.

To compensate for this difficulty, the author suggests the proposal and evaluation of alternative characterizations of given finite corpora of behavior in terms of their relative powers and/or parsimony. That is, given a large class of behaviors, such as those associated with mastery of a given school curriculum, the idea is not only to come up with a finite set of rules which characterizes the curriculum but to come up with the best possible set. (Loosely speaking, power refers to the diversity of behaviors which the characterization accounts for; parsimony refers to the number

---

3. If there was some way of knowing, then Church's thesis would provide a natural basis for deciding whether or not the behavior (potential) is rule-governed. Church's thesis (Rogers, 1967, 20–21) is that partial recursive functions (which can be defined formally) are precisely those which can be computed by algorithm (which is an informal notion). Thus, the proposal would be true or false depending on whether the class of potential behaviors is or is not partial recursive.
and intuitive simplicity of the rules in the characterizing set.) Such criteria, of course, have been an essential part of formal linguistics ever since Chomsky's (1957) influential *Syntactic Structures* was published.

In order for a characterization to have maximal relevance to psychology, however, these criteria alone are not sufficient. It is also important that a theory of knowledge (i.e., a characterization) be compatible with the mechanisms which govern human learning and performance. Specifically, it is important, in addition to specifying finite rule sets, to also specify how the constituent rules may be combined to generate behavior. It is these "rules of combination" which must find parallels in the way learned rules are put to use in particular situations. This question of relationships between different levels of theorizing is an extremely important one. For further discussion, see Scandura (1970c).

The basis of the present argument is that, given suitable rules of combination, much of what normally goes under the rubric of creative behavior can be accounted for in terms of finite rule sets. In order to limit the scope, this paper will deal primarily with those kinds of rules which are more properly associated with mathematical or logical content—specifically, with mathematical systems and axiomatic theories. In each case, one begins with a mathematical characterization and then shows what it means to know the underlying mathematics in a behavioral sense.

Relatively little attention is given to so-called mathematical processes.4 Thus, for example, inference rules are discussed, but relatively little is said about heuristics and other higher order rules by which inference rules may be combined in constructing proofs. This does not imply, however, that such processing skills cannot be formulated in terms of rules. To the contrary, it is basically a simple matter to formulate such heuristics as "organize (arrange) the data" and "work backward from the unknown" (cf. Polya, 1962) as rules. What is hard is to show explicitly how these rules may be combined with other rules to solve problems. Even this problem is not insurmountable, however, and some illustrative analyses of this sort have been worked out (Scandura, 1970b).

**What Is a Rule?**

Before continuing, it is necessary to define what is meant by a *rule*. In spite of an increasing amount of research on the subject, it is perhaps surprising that the term has no clearly defined meaning among behavioral scientists.

As a first step it is necessary to make a sharp distinction between underlying rules—or generative procedures composed of rules—and rule-

---

governed (RG) behavior. Intuitively speaking, a class of behaviors is said to be RG if the behavior can be generated by a common algorithmic (generative) procedure of some sort. This means, in effect, that a person who has mastered any underlying procedure should, ideally speaking, be able to generate each and every response, given any particular stimulus in the class of stimuli.

More specifically, RG behavior involves the ability to give the appropriate response in a class of functionally distinct responses to each stimulus in a class of functionally distinct stimuli. (The term “functionally distinct” refers to the fact that each effective (i.e., functionally distinct) stimulus (response) corresponds to a class of overt and “functionally equivalent” stimuli (responses).) The class of S-R pairs, defined in this way, are called S-R instances. To see what this means, consider simple addition. The proposed definition says that the behavior is RG if each pair of numbers is attached to a unique number called the sum. Thus, for example, any overt representation of the number pair (5, 4) can be paired with any overt representation of the number 9 but not with any representation, say, for the number 6.

Ideally, then, RG behavior corresponds precisely to the notion of a function in the mathematical sense. That is, every stimulus is paired with a unique response. When looked at in this way it is clear that what psychologists call concepts and associations can be viewed as special cases of rules (Scandura, 1968, 1969a, 1969b). Concepts are simply rules in which each stimulus in a class is paired with a common response. Associations are further restricted to a single stimulus-response pair.

In its simplest form, a rule can be viewed as an ordered triple, \((D, O, R)\), where \(D\) is the set of \((n\)-tuples of) stimulus properties which determine the responses, and \(O\) is the operation or generative procedure.

---

5. As indicated above, of course, the behavior of human beings is not always ideal. People make mistakes. There are two conceptually different ways in which errors may occur. First, the rule(s) learned by a subject may only apply to a subclass of S-R instances (of the given RG class). Thus, for example, young children are frequently unable to add numbers which involve “carrying” although they can perform perfectly well on those that do not. In this case, following the notion of partial function in recursion theory, one may refer to such behavior as partial RG behavior. Partial RG behavior is rule-governed but not (necessarily) by rules associated with the given RG class. The other way in which errors may arise is due to the limited capacity of human subjects to process information (Miller, 1956). There is, in effect, an important difference between knowing a rule and being able to use it (Chomsky & Miller, 1963). Thus, a person may know how to add any pair of numbers but be quite unable to perform the necessary operations mentally when the numbers are large. In the present discussion, the author assumes throughout that all rules can be used perfectly.

Note (parenthetically) that the abstract notion of a functor is sufficiently flexible to capture either or both senses of incompleteness. (Roughly, a functor is a structure preserving function between two categories, the categories being analogous to classes of functionally distinct stimuli and responses.) Whether there is any real significance to this fact or not, however, the author cannot say (cf. Scandura, forthcoming).
Concluding Comments

In conclusion, this paper has dealt primarily with what it means to know an existing body of mathematics. Relatively little has been said about intellectual skills of the sort that must inevitably be involved in doing real mathematics. Nonetheless, it has been shown that what appears to be creative behavior might well be accounted for in terms of growing rule sets. The key idea in making this a feasible and rather attractive possibility is that of the higher order rule. Although space limitations have made it necessary to ignore many details, and there obviously are still a good many important questions left unanswered, the author feels that enough has been said to convince the reader that the basic conjecture must be taken seriously: all mathematical behavior is a rule-governed activity and the basic underlying constructs are rules.

REFERENCES


Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81–97.

Journal for Research in Mathematics Education, 1971, 2, 183 - 196